UNIVERSITY OF SOUTHAMPTON
FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Emerging a Web Science Lexicon

by
Elisabeth Coskun

A thesis submitted for the degree of Doctor of Philosophy

Supervisor: Su White
Second Supervisor: Thanassis Tiropanis
Internal Examiner: Les Carr
External Examiner: Paul De Bra

October, 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES

ELECTRONICS AND COMPUTER SCIENCE

WEB AND INTERNET SCIENCE GROUP

Thesis for Doctor of Philosophy

By Elisabeth Ann Coskun

Web Science is an emerging subject which is fundamentally inter-disciplinary. The term Web Science was first coined in 2006. Subsequently, a growing number of institutions across the world began offering Web Science taught programmes. The majority of Web Science study takes place at Masters or PhD level; however a number of institutions are now also offering Web Science undergraduate programmes. Attempts at creating a formal subject definition for Web Science were made during the early days of formal Web Science teaching, with the Web Science Subject categorisation being released for public use in 2011. However, this was derived using a formal top down, theoretical approach. In practice, Web Science as it is researched and taught differs significantly from the formal model. While it is valuable for educators to have a formal definition of Web Science, the process of manually assembling curriculum guidelines is onerous and protracted. Conventional methods of curriculum design typically involve an international panel of experts meeting, conferring, creating an initial definition and putting it out for consultation with educators. This method is both time consuming and expensive. Within a rapidly evolving field such as Web Science, the definition is also highly likely to become outdated by the time it is formally defined. In order to address this problem, research which identifies from the ground up, the content of Web Science taught programmes, can usefully be used to define the subject.

This is an empirical study, the aim of which is to develop recommendations for a clearer definition of the Web Science field, utilising a ground up approach to develop an overview of Web Science derived from what is being taught, studied and written about Web Science in practice. The strength of such a framework would be its ability to provide academics and students wishing to study Web Science with an updated picture of the Web Science curriculum. Such a definition would also aid with recruitment, University reporting and evaluation, as well as new programme creation. The proposed definition will be achieved utilising a variety of data sources, including a desk survey of Web Science taught programmes and modules, a questionnaire based survey and interviews with Academics and Students in Web Science (or related) fields, as well as a review of existing Web Science conference literature. Data will be analysed using a process of triangulation, involving a mainly qualitative Grounded Theory based approach, utilising thematic analysis of the resources described, also incorporating some mixed methods such as keyword analyses. The final outcome proposed by this project involves the emergence of a framework for a working definition of Web Science with topics drawn from an overview of the Web Science curricula, as well as the current available Web Science literature. This will be supplemented by input from Academics and Students in the Web Science and related fields. The final output of this thesis will comprise of an online resource to be named the 'Web Science Lexicon'; a taxonomy of the key component topics of Web Science identified from the research described.

DECLARATION OF AUTHORSHIP

I, Elisabeth Ann Coskun ....

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Emerging a Web Science Lexicon ....

....

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. [Delete as appropriate] None of this work has been published before submission [or] Parts of this work have been published as: [please list references below]:

Signed: ....

Date: 25th October 2017 ....

Contents

Figures and Tables

Acknowledgements

I would like to thank **Ash Smith (iSolutions)** for his involvement in this project, which comprised of coding a Web Crawler for searching the .ac.uk. domain. The purpose of which was to return links to sites which included a combination of the keywords 'Web' and 'Science'

I would like to thank **Chris Gutteridge (iSolutions)** for his help in setting up a Drupal site for the purposes of hosting the Web Science Lexicon. I would also like to thank him for his input in brainstorming the idea of the Lexicon, including for the idea that it could be hosted as part of the University of Southampton Web and Internet Science Research Group site. I would also like to thank him for contributing the name 'Lexicon'.

I would like to thank **Ash Smith, Melgarejo Heredia, Bartosz Paszcza** and **Caroline Harcow** for their contributions to the 2017 WAIS Web Science Lexicon Fest theme. Special mention goes **to Melgarejo Heredia** for his suggestion of a matrix structure for the Web Science Lexicon.

I would also like to thanks all who contributed their time towards completing the survey of Web Science academics and students, and especially to those who took part in the interviews of Web Science PhD researchers.

1 Glossary

CDT: Centre for Doctoral Training

Component or Constituent Subjects: This refers to subjects which are part of Web Science but may also be considered to be subjects in their own right. For example, Computer Science or Sociology are established displaces in their own right, but they may also be considered as subjects which are a key part of Web Science.

Desk study/survey: "a preliminary investigation and report into something, collating currently available relevant information" [1]

Grounded Theory: "Grounded theory (GT) is a systematic methodology in the social sciences involving the construction of theory through the analysis of data".[2]

Lexicon: The term 'Lexicon' was chosen to describe the proposed taxonomy of key topics for Web Science. The definition of 'Lexicon' is "*a wordbook or dictionary, especially of Greek, Latin, or Hebrew*[3]". Given that the proposed Taxonomy of Web Science topics is a resource presenting the scope and providing a definition of Web Science it was decided that 'Lexicon' was an appropriate title. (Credit to: Chris Gutteridge)

Key Topic: A subject considered significant to Web Science and identified as part of a Web Science taught programme.

Paper Topic: A subject considered significant to Web Science and identified as part of a Web Science Conference publication.

WAIS: This acronym stands for Web and Internet Science; the research group at the university of Southampton to which the author is affiliated.

WAIS Fest: The annual research week held by the University of Southampton's Web and Internet Science Research Group.

Web Science Subject Categorisation (WSSC): The only known attempt to create a set of formal definition for Web Science. (Vafopoulos 2010) See appendix section 13.1 for a full subject listing.

Web Science Cluster Diagram: A visualisation created by Nigel Shadbolt depicting the scope of Web Science.

---

[1] https://www.collinsdictionary.com/dictionary/english/desk-study

[2] https://en.wikipedia.org/wiki/Grounded_theory

[3] http://www.dictionary.com/browse/lexicon

2 Introduction

This thesis aims to explore the scope of the discipline Web Science; an emerging subject which is fundamentally inter-disciplinary. The term Web Science came to prominence in 2006 when used by (Berners-Lee, Weitzner, et al. 2006). It describes a science derived to study the impact of the Web on the world, both from a technological and social perspective. "*Web Science undertakes to study the full scope of socio-technical relationships engaged in the World Wide Web*." (Akkermans et al. 2011). Recognising the importance of this emerging subject, and the need to study the impact of the Web, a growing number of Universities across the world have begun to teach Web Science programmes. There are currently very few definitive subject definitions available for Web Science. Additionally, the nature of the subject is constantly evolving as an increasing number of different disciplines begin to practice what might identifiably be called Web Science. This potentially provides educators and students with a problem; how do you teach or learn about Web Science when there is no clear definition. This is an issue if Web Science is to continue to gain popularity as a taught subject. There are currently few points of reference which allow people to easily gain an overview and basic understanding of the enormous scope of what is a new, challenging and exciting discipline, with far reaching implications for both technology and society.

2.1 The Need for a clearer Definition of Web Science?

A resource outlining the scope of Web Science could be beneficial in a number of areas. The accreditation body for Computer Science and Information Systems (ACM) has provided rigorous guidelines for the development of Computer Science and Information Systems programmes. However, despite the ACM being the accrediting body for Web Science, no such guidelines exist for Web Science. Therefore, there is no current benchmark against which to compare the content delivered within Web Science programmes, other than the content delivered within other existing Web Science programmes. Such a benchmark would be beneficial for universities in terms of exams and assessments of taught content. The proposed Lexicon is potentially useful for universities who wish to begin offering a Web Science taught programme, and therefore are required to provide justification for the provision of a new taught subject. It may help to provide an overview of the scope of the subject as justification for running a new taught programme. For example, a form provided by the university of Sheffield for new programme creation suggests that it is necessary to provide "*a brief, interesting statement, accessible to an informed lay audience, which gives the background and context of the programme and highlights its distinctive and attractive features* [4]".

Deriving learning objectives for programme design and assessment is also an area in which a resource outlining the scope of Web Science could be beneficial. When deriving learning outcomes and programme objectives, Bloom's taxonomy (Bloom 1956) is a framework commonly utilised for the design of learning outcomes. However, in order to derive a set of learning objectives for a programme, it is useful to have a reference for the scope of the subject that is being taught, in order to derive learning objectives. For example, (University of Leicester 2015) describes rigorous standards which must be met for the provision of academic taught content. The specific details of these standards will vary according to different subject areas and accreditation boards. The

[4] https://www.sheffield.ac.uk/polopoly_fs/1.648863!/file/progspec_guidancesep2016.docx

University of Leicester document for the development, approval and modification of taught provision (University of Leicester 2015) states that one of the criteria required for the assessment of a proposed programme is: "*Does the programme design reflect current disciplinary knowledge and best practice*?" If there is little or no definition available for the subject area of the proposed programme, how can this be assessed?

Another key benefit of the proposed Web Science Lexicon, relates to student recruitment. It would be beneficial to have an accessible definition of Web Science, which is understandable by the intelligent layperson, in order to promote Web Science as a desirable subject for prospective students to study. Students interested in enrolling on a Web Science programme could access the resource, in order to gain a better understanding and overview of Web Science.

This thesis employs a 'bottom up', empirical approach to document current Web Science related material, including compiling details of current Web Science taught programmes available, in order to answer the question – 'What is Web Science?' from a taught perspective. The second key data source will include an analysis of Web Science conference material published within the Web Science conference series, which has run since 2009. This will present a view of Web Science according to what is published as part of Web Science in practice. The analyses of taught programmes and conference proceedings will be supplemented with the findings of a questionnaire survey and a series of interviews targeting students and academics in the Web Science field. This will provide an insight into Web Scientists' perceptions of the scope of the subject, as well as gaining an insight into what people study as part of Web Science in practice. The findings of these analyses will contribute to the production of the Web Science Lexicon. The proposed Lexicon is a resource outlining the findings of the study, in the form of a taxonomy of the most frequently occurring topics identified as part of Web Science, as it is taught and published in practice.

2.2 Proposed Outcomes and Research Contribution

The key proposed contribution of this thesis includes:

2.2.1 Web Science Taxonomy of Key Topics (Web Science Lexicon)

The main outcome and contribution of this thesis will be to derive a taxonomy of key topics for Web Science, outlining the scope of the Web Science subject. This will be achieved through a combination of the most frequently occurring 'key topics' taught within Web Science programmes worldwide, and the 'paper topics' published as part of Web Science conference material. The terms 'key topics' and 'paper topics' refer to subjects or topics within Web Science, (such as for example, 'Semantic Web'), which commonly occur as part of Web Science taught programmes (key topics) and Web Science publications (paper topics).

These key topics and paper topics identified as part of Web Science taught programmes and publications will be supplemented by responses to a survey of Web Science Academics and students, and a series of interviews of University of Southampton Web Science PhD students. This will provide a view of Web Science as it is taught, published, and also by how it is perceived and understood by Web Science students in practice.

The proposed taxonomy of Web Science key topics, will be named the 'Web Science Lexicon'. Each key topic will have a dedicated webpage, which will contain the following information:

- A brief description of the topic
- References to Web Science publications which relate to the topic

2.3 Structure of the Thesis

This thesis has the following structure:

- Chapter 1: Glossary
- Chapter 2: Introduction
- Chapter 3: What is Web Science? – Review of the background literature
- Chapter 4: Further Background Research: Subject Comparison and Classification of Disciplines
- Chapter 5: Research Questions
- Chapter 6: Research Approach
- Chapter 7: Data Collection Methods: Introduction
    - Data Collection Methods: The Desk Survey
    - Data Collection Methods: Questionnaire Survey of Web Science Academics and Students
    - Data Collection Methods: The Web Crawler
    - Data Collection Methods: Web Science eBook – Interviews
- Chapter 8: Analysis
- Chapter: 9 Structuring the Lexicon
- Chapter 10: Conclusions and Discussion
- Chapter 11 Future Work
- Chapter 12: References
- Chapter 13: Bibliography
- Chapter 14: Appendices

3 What is Web Science? - An Overview of the Literature

3.1 Introduction

This chapter will include a review of literature which details the purpose and reasons for forming Web Science as a recognised field of study, providing a narrative outlining the history and development of the field. This review will begin by examining the initial creation of the subject in the year 2006, and examine the reasons behind the fact that there is a need for a science dedicated to the study of the Web. It will also review the few existing attempts to formally define Web Science, such as the Web Science Subject Categorization (WSSC). Additionally, the following section will cover some of the problems which exist when trying to define the scope of Web Science; problems which this study aims to address. Finally, this review will also examine the research approach which is suited to studying Web Science, which will have implications for the way in which the research undertaken by this thesis will be conducted.

The following literature review provides an overview of literature which charts the development of Web Science as a discipline. This review has the following structure:

- An overview of Web Science and justification for why it is needed
- The moral need for Web Science
- An examination of the founding publications which outline the purpose of Web Science and its defining characteristics
- Existing attempts to define the scope and curriculum of Web Science
- The rapid evolution of Web Science: A Bottom up Research Approach?
- The future development of Web Science

The purpose of this initial review of background literature is to examine the early publications describing the purpose of Web Science, exploring what they have to say regarding the nature of the subject and its intended purpose. Although there is an entire conference series dedicated to Web Science, this review focuses upon publications which specifically examine Web Science as a subject, in order to scrutinise the background and reasons for the creation of the new science dedicated to studying the Web. The reasoning behind this is to provide an overview of the subject, and also to identify and highlight the problems associated with researching Web Science. The remaining conference publications which are not directly related to the foundation of Web Science are treated as data for study during the analysis stage of this work.

3.2 What is Web Science? – A Summary

In this author's own personal experience, Web Science can often be misconstrued by the uniformed layperson as the mere study of technology. In reality, Web Science has a far greater scope. While the Web is indeed powered by technology, it is far more than mere hardware and algorithms; the Web is driven by its users. Since the advent of Web 2.0, content on the Web has become increasingly user driven. Therefore, any science dedicated to the study of the Web needs to also study the impact and effect of the Web on users, and conversely, the effect of users on the Web, in addition to the technological platform.

The importance of the Web in everyday life, combined with the fact that the Web is constantly evolving, suggests that gaining an understanding of what drives its development is key. Because the Web now impacts the lives of many people across the world; it is not only significant to try to determine how the Web may change, but also to identify what may cause it to change, as this will help to predict *how* it might change. Leading Web Science academics including Tim Berners-Lee and Wendy Hall, suggest that one of the key aims of Web Science, should be not only to chart and study the Web as it is in its current state, but also to ask the question, where is the Web going? How will it develop in the future, and what will influence its development? This is important in order to secure the future of the Web.

A key feature of Web Science is that it is inherently interdisciplinary; in order to study and fully understand the Web as a phenomenon, it is not enough to merely understand the technology behind the Web, but also the social implications. The Web affects all areas of society in which the technology is used. This has huge scope, transcending the traditional Computer Science and Information Technology fields, also encompassing fields such as the Social Sciences, Education, Health, Media, Law, Economics and many more. Because of the inherent multidisciplinary nature of the Web, it is only natural that in order to fully understand the impact of the Web, a multidisciplinary perspective is required. The study of Web Science aims to bring together people from these multiple disciplines, which is reflected in the intake of students for Web Science programmes, and also in the spread of subjects published across the Web Science conferences series.

During the 2015 Web Science conference, a phrase from Tim Berners Lee, which was regularly quoted, was 'the Web we want'. A theme echoed throughout, this suggested that it would be important to understand how the Web is evolving and changing in order to attempt to influence this development in a way that is beneficial and desirable for society. Tim Berners Lee suggested the need for an 'Online Magna Carta' in order to provide standards for acceptable behaviour on the Web. He suggested that it is the job of the Web Science community to monitor the growth of the Web and also to ask the question, 'where is the Web going?'

The following section of this chapter will examine what Web Science literature has to say in answer to the question, “What is Web Science?”

3.3 Comparing Existing Descriptions of Web Science

One approach by which it is possible to gain an understanding of the nature of Web Science, is to examine what different Web Science related institutions have to say when describing the subject. The following table includes a sample of what a number of Web Science related organisations say about the subject; these include the Web Science Trust, and a number of teaching intuitions which offer Web Science taught programmes.

Table 1. Descriptions of Web Science

Descriptions of Web Science	
Location	Description
Eindhoven University of Technology [5]	*Web Science tries to fathom the technology of the internet and users, both from a technical perspective and based on the use of the internet. Why are new services developed and how do they function? Why do people use online services such as Twitter, Flickr, Google Docs, Facebook or Gmail? What can go wrong and how can this be prevented?* *To be able to obtain answers to those questions, Web Science combines computer science with other disciplines such as psychology and sociology. This enables us to understand how online communication is used. After completing a degree in Web Science, you will be able to assist in enabling users to communicate more effectively online and continue the development of the technology behind this communication. In the program, we focus primarily on the stable elements of the internet, rather than hype content.*
Georgia Tech [6]	*"The Web Science initiative at Georgia Tech seeks to understand the impact of the Internet, and help us invent its future. Web technologies are playing an increasingly central role in business, education, entertainment, health, and research. Understanding these phenomena and engineering their future requires a new interdisciplinarity. We aim to strengthen ties among disciplines across campus, explore possible new interdisciplinary educational programs, and build stronger ties between Georgia Tech and the web industry".*
KAIST [7]	*"Web Science and Technology (WebST) is a new emerging discipline that takes the Web as its primary object of study and engineering. Exploring the interactions among the complex technical, engineering, and social aspects of the Web, our graduate-level program rests on computing technologies to build and extend the Web, engineering techniques to develop large-scale applications, and analytical methodologies for understanding its nature and impact on various aspects of human society."*
RPI [8]	*"Web Science is the study of the World Wide Web and its impact on both society and technology, positioning the Web as an object of scientific study unto itself. Web Science recognizes the Web as a transformational, disruptive technology; its practitioners study the Web, its components, facets and characteristics. Ultimately, Web Science is about understanding the Web and anticipating how it might evolve in the future."*
University of Southampton [9]	*"The University of Southampton is pioneering the development of Web Science, a new discipline that aims to provide a thorough understanding of the Web as a social and technical phenomenon. This demands new ways of working, across traditional academic disciplines, to build skills and expertise in the technical underpinnings of the Web as well as the social processes that have shaped its evolution and the impact of the Web on society".*
Web Science Trust [10]	*"Nothing like the Web has ever happened in all of human history. The scale of its impact and the rate of its adoption are unparalleled. This is a great opportunity as well as an obligation. If we are to ensure the Web benefits the human race we must first do our best to understand it. The Web is the largest human information construct in history. The Web is transforming society. In order to understand what the Web is, engineer its future and ensure its social benefit we need a new interdisciplinary field that we call Web Science."*

The key common aspect shared by all these descriptions, is the fact that the Web is interdisciplinary, and is a socio-technical construct. Web Science is the study of the Web as a technology, its users and the impact produced by the technology on the users and wider society.

3.3.1 The Moral Implications of the Web and the need for Web Science

This section compares what several key Web Science publications say when justifying the need for Web Science, including the moral reasoning behind the creation of a completely new science to study the Web. The Web is a radical technology, which has propagated worldwide, and changed the way that the world behaves, conducts business and economics, it has shaped laws, healthcare, education, and the way that people interact socially. (Berners-Lee, Weitzner, et al. 2006) explain that "*The Web started life as an attempt to get people to change their behaviour in an important way*."

---

[5] https://web.archive.org/web/20150709141614/https://www.tue.nl/en/education/tue-bachelor-college/undergraduate-programs/web-science/

[6] https://web.archive.org/web/20161024205846/http://webscience.cc.gatech.edu/

[7] https://web.archive.org/web/20160826113221/http://eng.webst.kaist.ac.kr/content.php?db=_intro

[8] http://tw.rpi.edu/web/WhatIsWebScience

[9] https://web.archive.org/web/20160830082122/http://www.southampton.ac.uk/webscience/index.page

[10] http://webscience.org/web-science/web-science-home/

They go on to explain that this highlights the need for a greater understanding of this phenomenon, with a view to influencing the Web's development in a way that is open and just. In their 2008 work entitled 'Web Science: an Interdisciplinary approach to understanding the Web' (Hendler, et al 2008) also present a moral argument for the need to study the Web. They explain that corporations have a moral duty to ensure their online products do not have negative implications, and that Web Science is one of the answers facilitating a better understanding of these implications. Table 2 includes some examples of moral arguments for Web Science.

Table 2. The moral implications and need for Web Science

The moral Implications and need for Web Science	
Quote	**Reference**
"*The Web's aim was to alter that behaviour radically and provide the technology to do it: people would make their documents available to others by adding links to make them accessible by link following. The rapid growth of the Web, and the way in which this change was quickly adopted in all sectors of Western society have perhaps obscured the radicalism of this step*." "*An important aim of Web Science is to identify the essential aspects of identification, interaction and representation that make the Web work, and to allow the implementation of systems that can support or promote desirable behaviour*." "*So one aspect of Web Science is the investigation of the Web in order to spot threats, opportunities and invariants for its development. Another is the engineering of new, possibly unexpected methods of dealing with information, which create non-conservative extensions of the Web. Such engineering may be research-based, or industry-based*."	(Berners-Lee, Weitzner, et al. 2006)
"*Corporations have a responsibility to ensure that the products and services they develop on the Web don't produce side effects that harm society, and governments and regulators have a responsibility to understand and anticipate the consequences of the laws and policies they enact and enforce. We cannot achieve these aims until we better understand the complex, cross-disciplinary dynamics driving development on the Web—the main aim of Web science*."	(Hendler, et al 2008)
"*The far-reaching impact of the Web on society is widely recognised. The interdisciplinary study of this impact has crystallised in the field of study known as Web Science".*	(White et al., 2011)

One key feature shared by all of these quotes, is the fact that they agree the need to influence the Web in a positive way for the benefit of its users and wider society, and that this should be the goal of Web Science. It is the primary justification for the creation of Web Science.

3.4 Background – The Creation of Web Science: A framework for Web Science

In order to understand Web Science and its importance as a discipline, it is necessary to revisit the inception of the subject, in order to understand the motivation of its creators, and why they felt the need to create a new science dedicated to the study of the Web. The first conclusive publication for Web Science entitled '*A Framework for Web Science: Foundations and Trends in Web Science* '(Berners-Lee, Weitzner, et al. 2006), was published in 2006. This work heralded the introduction of the term Web Science, and argued the need for a science dedicated to the study and understanding of the Web. "*We may paraphrase Web Science as the science of the Web*." (Berners-Lee, Hall, et al. 2006) This is one of statements (Berners-Lee, Weitzner, et al. 2006) make when introducing the concept of Web Science in their 2006 paper: *'A Framework for Web Science*. The authors sought to address the issue that the study of the Web was considered a mere subsidiary to Computer Science, when in fact the Web is a huge phenomenon, which has a massive impact upon not only Computer Science but also the wider world, and hence deserves study in its own right. This publication is widely perceived as the foundation of the Web Science subject; however, the roots of this publication can be traced back to an earlier workshop paper in 2005.

When using the term 'Web Science', (Berners-Lee, Weitzner, et al. 2006) suggest that the term 'Science' is used in two different concepts. They explain that Physical and Biological science derives theory from observing behaviour in the natural world. They analyse the natural world, and attempt to find microscopic laws to explain macroscopic behaviour. By contrast, Computer Science is mainly synthetic and theoretical, and is primarily concerned with the creation of algorithms which govern computer behaviour.

(Berners-Lee, Weitzner, et al. 2006) argue that if the Web is to remain ethical, maintaining social values such as trust and privacy, then the Web must have its own dedicated research agenda. While the Web is an 'engineered space' created by algorithms, it also involves a human element. Humans create and interact via web pages, and this interaction is governed by social conventions. The authors argue that this is why Web Science is fundamentally socio-technical and interdisciplinary. Its goal is to understand the growth of the web, not only technologically, but also socially, and to understand how the technology affects society, and also how society may impact the development of the technology in ways which are beneficial to society.

3.4.1 Further Development: Web Science: a socio–technical construct

This section examines what further supporting Web Science publications have to say about the nature of Web Science as a subject, building upon the initial vision for Web Science proposed by (Berners-Lee, Weitzner, et al. 2006).

(Carr 2010) uses the analogy of the 'Big Bang' to describe the spread and influence that the newly developing World Wide Web had on Society. The development of the Web is influenced by many aspects of life, from economics and politics, to law, health and social movements, which is why standard Computer Science programmes often fail to grasp either the breadth of factors which impact the Web, or the areas upon which the Web itself can also impact. In their work, (Berners-Lee, Hall, et al. 2006) suggest that Web Science is inherently interdisciplinary; attempting to address the need to "breach the embedded binary divide between the natural and engineering sciences on the one hand and the social and human sciences on the other." This is a view supported by (White et al., 2011), who explain the benefit of "sharing approaches" between students and academics of all disciplines, contributing to the Web Science subject.

'Web Science Emerges' (Berners-Lee, Hall, et al. 2006), explains that the key goal of Web Science is to "*understand the growth of the Web*." The inherent heterogeneity of Web Science calls for a wide ranging curriculum which incorporates the knowledge, working methods, skills and understandings intrinsic to the many different participant disciplines and disciplinary cultures. The analysis of Web Science presented by (Halford et al. 2010) in their provocatively titled paper 'A manifesto for Web Science,' confidently asserts that "Computer Science alone, focusing as it does on the engineering and technology of the web cannot "*deliver the ambitions*" of the new agenda, which requires a dedicated science to study the evolution and interdisciplinary nature of the Web.

(Shneiderman 2007) presents Web Science as a new way of thinking about Computer Science. He cites Web Science as a 'provocative' invitation to Computer Scientists to think about the Web and consider its effects upon society, instead of simply focusing on the technologies behind the platform. According to (Shneiderman 2007), Web Science is required to address what users and communities want from the Web. They also suggest a shift in research is required, moving from studying the technology, to looking at what users can do with technology; also looking at the impact of technology on society.

3.4.1.1 Web Science: An Interdisciplinary Approach to Understanding the Web

A notable publication in the evolution of Web Science which builds upon the initial 'Framework' paper (Berners-Lee, Weitzner, et al. 2006), is the 2008 journal article entitled 'Web science: An Interdisciplinary Approach to Understanding the Web', (Hendler et al. 2008). (Hendler et al. 2008) point out that despite the massive impact that the Web has had on the discipline of Computing, the best classification for Web Science that can be found under the ACM Computing Classification Categories[11], is 'miscellaneous'. The scope of their article can be summarised as follows: "Here, we look at some of the technical and social challenges that must be overcome to model the Web as a whole, keep it growing, and understand its continuing social impact."

(Hendler et al. 2008) recognise that it is important to study the technological platform of the Web, as the Computing made significant contributions to technology, without which the Web would not otherwise exist. However, they explain that despite the massive impact that the Web has had on computing, Web Science is still understudied. The topic of the Web is still frequently only addressed as a 'service course', for example, 'Web Design'. They suggest that in part this is due to the fact that the technological principles behind the platform of the Web are addressed in Networking courses, which are a commonly taught topic within Computer Science. (Hendler et al. 2008) explain that such courses and the perspective of many Computer Scientists have a tendency to view the Web as an application running on a platform of technology and protocols, failing to consider the wider picture of the Web as a socio-technical entity. "*the Web is often studied exclusively as the delivery vehicle for content, technical or social, rather than as an object of study in its own right*."

It is the personal experience of this author, as a student having studied such a course, that while Web principles are indeed taught on such networking modules, more could be done to address the 'bigger picture' and practical application of these principles in the context of the web as a whole. Such networking courses have a tendency to explain the principle of technologies such as TCP/IP for example, without expanding on the wider context in which they are used. This is at least partly due to time constraints of the course, however more needs to be done to relate the subject of Web technologies to the wider impact of the Web, and this is the very purpose of Web Science.

Both (Shadbolt et al. 2008) and (Hendler et al. 2008) explain that Web Science is a combination of the study of both the micro and macro; at the micro level, it is based upon a technological platform with a series of programming languages and protocols. At a macro level, the Web facilitates social interaction between human beings and is a communication platform, where content is generated by humans. They explain that the future of Web development must take into account the macro effects produced by Web applications produced at a macro level. This requires the ability to research and

[11] http://dl.acm.org/ccs/ccs_flat.cfm

predict the wider effects that micro technological developments will have upon wider society at a macro level.

(Hendler et al. 2008) describe the process of humans interacting with the Web as creating 'emergent properties', and that these require new research methods in order to study and understand the behaviour. These new research methods involve the use of social science based methodology to understand the effects of technology, a traditional Computer Science methods traditional involve understanding if the technology works correctly at the micro level; not taking into account the wider macro effects that Web has the capability of producing across society as a whole. (Hendler et al. 2008) explain that the combined approach of considering the combined micro and the macro effects requires an interdisciplinary approach, and this is the space that Web Science occupies.

The 'Web Graph' is provided as an example by (Hendler et al. 2008) as one way of looking at the structure of the Web as a whole. "*However, the Web graph is just one abstraction of the Web based on one part of the processing and protocols underlying its function*." While a valuable tool for understanding the structure of the Web, it is still a traditional Computer Science methodology, and it does little to address the motivations and behaviours of users on the Web. While (Hendler et al. 2008) recognise that "Analysing the Web solely as a graph also ignores many of its dynamics" such as user behaviour.

The notion of a 'social machine'; a system comprising of both humans and technology is expounded by (Hendler et al. 2008). The concept of a social machine predates the advent of Web Science and was suggested in the publication (Berners-Lee & Fischetti 1999). Examples of Social machines include sites such as Wikipedia and Facebook. As described by (Hendler et al. 2008): "*The social machine includes the underlying technology (mediaWiki in the case of Wikipedia) but also the rules, policies, and organizational structures used to manage the technology*." Web Science is necessary for understanding these social machines, as they are by nature a social technical construct facilitated by the Web, requiring the understanding of the social and the technical, the micro and the macro.

3.4.1.2 A Manifesto for Web Science

While (Hendler et al. 2008) acknowledge the need to study both aspects of the Web at micro and macro levels, (Halford et al. 2010) highlights the fact, that while stressing the need for interdisciplinarity, (Berners-Lee, Weitzner, et al. 2006)and (Hendler et al. 2008) both approach the problem of defining Web Science from a largely Computer Science perspective. (Halford et al. 2010) therefore build upon the existing work of (Berners-Lee, Weitzner, et al. 2006) and (Hendler et al. 2008), by providing an additional social science insight and perspective into Web Science. The Manifesto paper (Halford et al. 2010), represents an interdisciplinary approach, given the background of the authors; Catherine Pope is a Professor of Medical Sociology, Susan Halford is a Professor of Sociology and a director of the Web Science institute, while Less Carr is Professor of Web Science with a background in Computer Science, Multimedia and Hypertext.

(Halford et al. 2010) further develop the interdisciplinary notion outlined by (Berners-Lee, Weitzner, et al. 2006) and (Hendler et al. 2008), by exploring four core concepts drawn from social theory, which aid in: "*developing the inter-disciplinary thinking across natural, social and human sciences that will be essential for Web Science to fulfil the aspirations of its originators*." They suggest that the combination of these four core concepts provide a foundation for genuinely interdisciplinary research into the Web. These four concepts are summarized in Table 3.

(Halford et al. 2010) begin their paper by outlining the interdisciplinary nature of Web Science; they explain that it has always been clear how disciplines such as Computer Science and Maths can be harnessed and utilised to understand the Web. However (Halford et al. 2010) build upon the prior arguments of (Berners-Lee, Weitzner, et al. 2006) and (Hendler et al. 2008), who's papers both argue that the Web requires a much wider understanding, incorporating expertise and alternative view points from areas such as the social sciences. (Halford et al. 2010) then list a series of disciplines which are also associated with Web Science, these include: "Sociology, Geography, Psychology and Cultural Studies". While (Halford et al. 2010) identify these disciplines as having 'a research interest' in Web Science, this is not an exhaustive list of disciplinary areas which are related to Web Science.

(Halford et al. 2010) highlight the fact that discipline specific expertise relating to the Web has a tendency to remain within the discipline in which it originates, and that there is little sharing between disciplines, especially between the "embedded binary divide" between engineering and the social sciences. One of the key goals of Web Science is to address this. (Halford et al. 2010) emphasise the (Hendler et al. 2008) assertion that Computer Scientists rarely study the "Web as a subject in its own right". They reiterate the importance of the interdisciplinary approach required in order to understand the Web as a socio-technical construct.

The four core concepts outlined in by (Halford et al. 2010) shown in Table 3, provide a way of understanding Web Science according to concepts taken from social science research theory. The first two concepts are self-explanatory, and are core to the nature of the Web; which is a combination of the technological and the social, and is shaped by both humans and technology, two networks which also in turn shape each other. The third concept, termed 'performability', suggests that the standard top down method of research utilised by quantitative researchers, who begin with a preconceived theory and then set out to find supporting evidence in order to prove it, is not necessarily the best approach for understanding the Web. Instead of this standard 'top down', theoretical approach, (Halford et al. 2010) propose that a 'bottom up' evidence driven approach is more useful for understanding the realities of the Web.

Table 3. Core concepts for understanding Web Science (Halford et al. 2010)

Core concepts for understanding Web Science (Halford et al. 2010)
• **Co-constitution: technology vs society** ○ The mutual shaping of technology and society "*The point of co-constitution, then, is to look at how technology and society shape each other. In web science this means we must examine how the web impacts on what people do and how people impact on what the web becomes*."
• **Heterogenous actors: humans vs technology,** ○ The web is a network of humans and non-humans (technology)
• **Performability** "*If we want to understand the web it is not a case of starting with pre-conceived ideas about what the web 'is' or what society 'is' but looking at what socio-technical relations become as they are performed in everyday life*."
• **Immutable mobile** ○ The web may appear static, and it is possible to take a snap shot of the Web at any one time, but it is constantly evolving and changing.

(Halford et al. 2010) conclude their justification for these core concepts, by explaining that the Web is constantly changing, and the therefore, while it is possible to take a snapshot of the current state of the Web at any one time, this will only be a static snapshot, and that the Web constantly evolves. This is concept is referred to as the term 'Immutable mobile'. They state that the use of "immutable mobile" in the context of Web Science refers "*a temporary stabilisation of the networks*."

The consequences of these core concepts identified in Table 3 result in the conclusion of (Halford et al. 2010) that Web Science requires a different research approach to that of standard subjects. They suggest that: "*if we are to follow the all actors implicated in the web we need to adopt both interdisciplinarity and mixed methods and open up web science to the ontological, epistemological and methodological possibilities offered by the social sciences and humanities*." In response to this need for an interdisciplinary approach, (Halford et al. 2010) state five proposals which they suggest Web Science should follow, if it is to fulfil its potential and the aims outlined by the original Berners-lee proposal, which is that Web Science must be genuinely 'pro human'. The goal of Web Science is to engineer "the Web we want", ensuring that the development of the Web benefits wider society, and safeguarding the Web in order to ensure its continued existence and benefits. (Halford et al. 2010) state that the five points proposed, which are shown in Table 4 are intended to build upon the earlier vision of Berners-Lee et al, by expanding on their proposals and offering the benefit of multidisciplinary insight.

Table 4. Five key goals of Web Science according to (Halford et al. 2010)

Five key goals of Web Science according to (Halford et al. 2010)	
1.	Web Science must be the genuine intersection of discipline; i.e. it cannot be allowed to be a sociology or a computer science of the web;
2.	Web Science must look both ways to see how the web is made by humans and how humans are made by the web;
3.	Web Science must follow all the actors (individual, groups and technologies) and trace the networks implicated in the web in the broadest sense and understand the effects of these networks;
4.	Web Science must move beyond narrow epistemologies and methodologies to enable a science which can examine and explain both micro and macro phenomena;
5.	Web Science must be a critical discipline - if it is to speak to the desire for the web to be pro-human – it must develop theoretical thinking and push towards critical, political social theory, to critique the direction of travel, to challenge the web and society

These five proposals are more concerned with outlining the 'features and behaviour' or 'attributes' of the Web Science subject than defining the scope of Web Science in terms of the specific subjects that should be taught within a Web Science curriculum. While this is intentional, as this work is outlining an ethos for general Web Science research, it also highlights the lack of 'component subject' or curriculum content definition for Web Science within the 'founding' Web Science publications.

3.5 Existing Attempts to Define the Web Science Curriculum

The following section examines the few existing attempts to define a curriculum definition for Web Science.

3.5.1 Lack of subject definition

All publications examined during this literature review are agreed about the social-technical nature of Web Science. However, few go into detail about the scope of what specific subjects the study of Web Science should entail. Several publications touch on the 'component subjects' of Web Science, for example, when providing a summary of the Framework paper, (Berners-Lee, Weitzner, et al. 2006), (Riera 2008) list the following attributes "*the science and the engineering*, the *analysis of the web, the social aspects involved and finally the legal, governance, security and standards issues*".

(Halford et al. 2010) state that: "understanding the Web requires knowledge and expertise from the social and the human sciences." They then list the following fields:

"*Sociology, and non-human - as these are constituted in the networks Geography, Psychology and Cultural Studies have long that produce the web.*"

Whilst a number of papers, such as the examples above, touch on the scope of subjects relating to Web Science, there are few detailed attempts to explore and document the 'component subjects' of Web Science. The lack of detailed reference to Web Science's component subjects during the founding Web Science publications is partly due to the fact that when most of these publications were written, Web Science was still in its infancy, and the scope of Web Science was still in the process of being determined. However, more than 10 years on, there has still been very little work into defining the scope of Web Science. What are the specific areas and topics of research and study that together, constitute the subject Web Science?

(Riera 2008) states that there is currently little in the way of a definition for the Web Science subject, and that the founders of Web Science are still attempting to establish the community and provide greater awareness of the subject. According to (Riera 2008): "*Clearly, it is not easy to give a precise definition of what comes under this new area, what is left out, and what interactions there are among the sub-areas*." This was partly addressed by the work of (Vafopoulos 2010) when creating the Web Science Subject Categorization; which is currently the only attempt to provide a subject definition for Web Science. The work of (Hooper, Dix, et al. 2013) analyses the representation of subjects within the first three Web Science conferences. (White et al. 2011) and (White & Vafopoulos 2012) outlined proposals for the creation of a Web Science curriculum repository. However, none of these references refers to current work in progress. This thesis aims to address the gap in current research, by examining the scope of Web Science and presenting these findings as a taxonomy of Web Science topics, to be named the Web Science Lexicon.

3.5.2 The Web Science Subject Categorization (WSSC)

As explained by (Hooper & Dix 2012), the only existing effort to create a formal definition for Web Science, is the Web Science Subject Categorization (WSSC). The WSSC (Vafopoulos, 2010), which was released for public use in 2011, is collaboration between academics, which aims to create a definitive subject definition for Web Science. (See **Appendix 1** for the full listing). (Vafopoulos 2010) describe the purpose of the WSSC as "*first epistemological index for the scientific study of the Web*". The aim of the WSSC was to facilitate interdisciplinary collaboration between Web Science researchers, originating from the multiple sub disciplines of Web Science. The WSSC was developed using a top down, theoretical approach, in a reverse engineering fashion. Academics derived a set of categories by breaking down the Web Science subject into its constitutional parts. E.g. "*computational, mathematical, social, economic and legal"* (Vafopoulos 2010).

The core concept of the WSSC, was to provide a referenceable resource outlining of the scope of Web Science. Plans also included the incorporation of links to papers relating to each topic. This was a constructive idea, which would have provided a much-needed resource for the Web Science community. However, the WSSC is no longer included on the current version of the Web Science Trust website. It is largely acknowledged within the community that this is due to the fact that it is now outdated due to the rapid evolution of Web Science. A contributing factor for this might be that the WSSC was not maintained and developed beyond its initial prototype. If this were the case, an editable resource, facilitating continued input from the wider Web Science community, could provide a sustainable format for a similar definition of Web Science. This could also have the benefit of input from a wider knowledgebase.

3.6 WSSCR: A Proposal for a Crowdsourced Curriculum Repository for Web Science

(White et al. 2011) discuss proposals for a document repository dedicated to Web Science, which they term 'WSCCR'. Based on the topics outlined by the WSSC, this was an attempt to assemble an accredited Web Science curriculum resource in one, easy to access location. As (White et al., 2011) indicate, in a rapidly evolving subject such as Web Science, links to external documents are also necessary, although the relevance of these may be rated against the WSSC. (White et al., 2011) also explain that as the number of items in the proposed Web Science repository increase, it would become more possible to form a picture of the Web Science curriculum, and determine which subjects are more prominent. This proposed repository had the potential to be a resource which could aid in academics when designing a Web Science curriculum. Despite promising proposals, no further information can been found relating to the idea beyond this initial paper, therefore it can be assumed that the proposals never came to fruition.

3.7 The Web Science Cluster Diagram

Rumoured to have been scribbled on scrap paper by Nigel Shadbolt on the way to a meeting, the Web Science Cluster diagram (Shadbolt n.d.) is currently the most widely recognised endeavour to create a visualisation depicting the Web Science curriculum. It is often referred to as 'the butterfly' diagram, due to its shape.

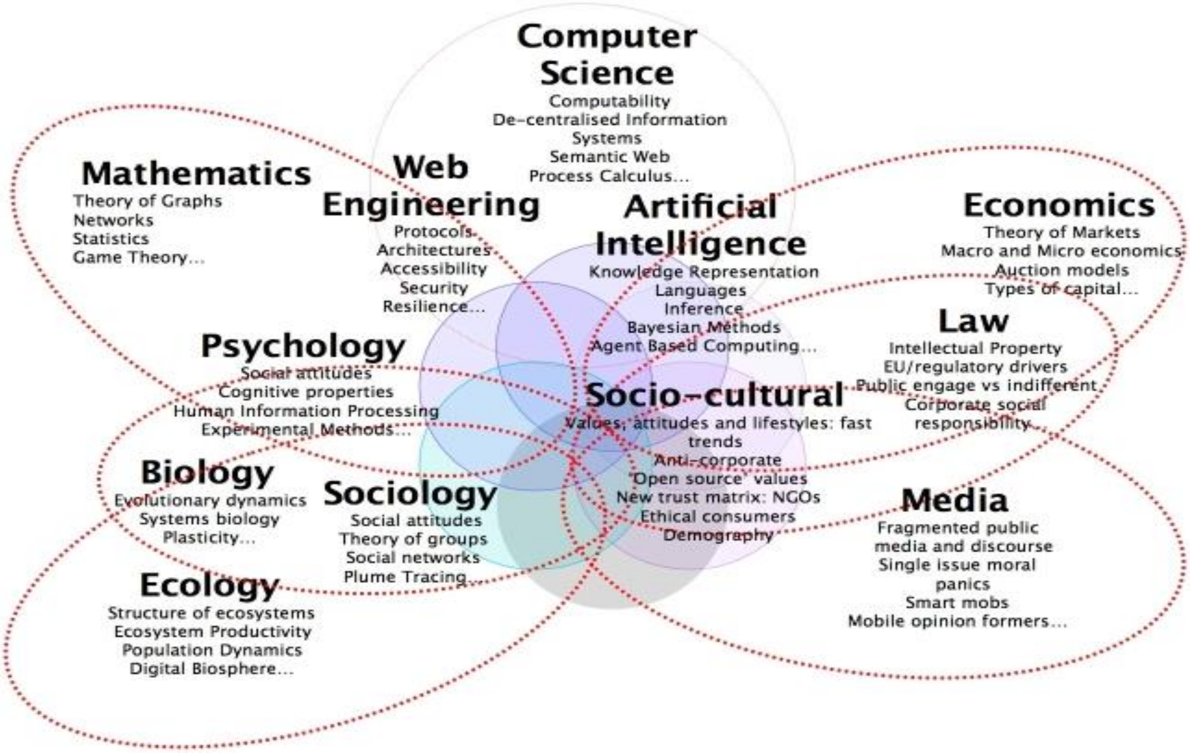


Figure 1: The Web Science Cluster diagram [12]

[12] http://www.edshare.soton.ac.uk/2666/1/Maps.pdf

Neither the WSSC nor the Web Science cluster diagram rate content according to importance; e.g. the amount that each subject within Web Science is represented by the community, or the number of publications in for each subject in the Web Science conference. The following section explores this issue in more detail.

3.7.1 WSSC and Web Science Cluster - Missing Subjects?

Although the WSSC is currently the most definitive definition of Web Science, it is not airtight, and does potentially include some significant gaps. This is likely to be a reason for the fact that it is no longer included on the Web Science Trust Web Site, and is a consequence of the rapid evolution of the Web. When examined, the Web Science conference topics highlight potential gaps in the WSSC curriculum. For example, the first of the Web Science conferences in 2009 contained an entire section on ‘Teaching and Learning’; this is notable, as it is an area which was not initially represented by the WSSC and is still not represented by the Web Science Cluster diagram.

Both the WSSC and the Web Science Cluster [13] are missing a few key subjects, such as for example, health; a key omission, as the Web now plays a huge part in health – e.g. telemedicine. Education is also a prominent omission from the Cluster diagram. Vafopoulos states that:

> "*Introductory school and university courses in Computer Science are often out of date, remaining stuck on the utilitarian prospect of Web technologies. In this aspect, Web Science education should address the complex techno-social issues of the Web in an attractive and compatible way to modern real phenomena and other scientific approaches.*"

Other than the work of (Vafopoulos 2010), a study conducted by Clare Hooper, (Hooper et al. 2012), (Hooper, Bordea, et al. 2013), (Hooper, Dix, et al. 2013) and (Hooper 2014), is one of very few examples of research into the scope of Web Science. (Hooper 2012) begins with details of initial results of an analysis determining the disciplinary coverage of publications at the Web Science conference series. (Hooper et al. 2012) and (Hooper, Bordea, et al. 2013)(Hooper et al. 2012) begin by recognising that there is little work defining the scope of Web Science. They critique the two current examples of attempts to depict the Web Science curriculum; the Web Science Cluster diagram (which they term the ‘Web Science Butterfly’) and the Web Science subject categorisation. They state that the Web Science Cluster diagram represents more of a “vision” of Web Science than a full and comprehensive reality. They also evaluate the Web Science Subject Categorization as only providing “*vision and structure, not information on subjects’ prevalence within the community*”.

According to (Hooper, Dix, et al. 2013) there has been ongoing discussion within the Web Science community regarding which disciplines are represented within Web Science and how Web Science should be defined as a subject; however, there is a “dearth of empirical work in this area.” They go on to explain that it is important to identify the spread and representation of disciplines within Web Science, in order to “identify problems that need addressing” with regards to defining the curriculum, and ensuring that all disciplines within Web Science are accurately represented. (Hooper et al. 2012) explores the notion of discovering and representing the spread of Web Science resources across Web Science, most notably, at the Web Science conferences, in an effort to determine which subjects are key to the Web Science curriculum, and which are currently

[13] Shadbolt, N., What Is Web Science? talk: http://webscience.org/professor-nigel-shadbolt-explains-web-science/

underrepresented. (Hooper et al. 2012) claim that they have not found any other previous research in this specific area, highlighting the need for additional research.

(Hooper et al. 2012) and (Hooper, Bordea, et al. 2013), utilise the Web Science cluster diagram as an example for illustrating the coverage of conference topics, providing an updated version of the diagram, presented as a 'heatmap' showing which subjects within Web Science are well represented within the Web Science conference, and which are not. (See Figure 2.) Unlike the original diagram, the (Hooper et al. 2012) version presents a more detailed picture of Web Science, by colour coding each of the subject areas according to their importance. This was achieved by determining the number of papers in each subject area. In this instance (Hooper et al. 2012) use natural language processing or 'bibliometric mapping' to analyse the keywords defined the first three Web science conferences. The concept of bibliometric mapping suggests that a given subject may defined by a list of important keywords, (Hooper et al. 2015). This process undertaken by (Hooper et al. 2012) accords with the suggestion made by (White et al. 2011), who also propose in their paper that the Web Science curriculum may be determined by examining conference topics. Once the results were gathered by (Hooper et al, 2012), these were then compared to the 'Web Science butterfly' diagrams and the Web Science Subject Categorization, which are the two representations of the Web Science curriculum currently available.

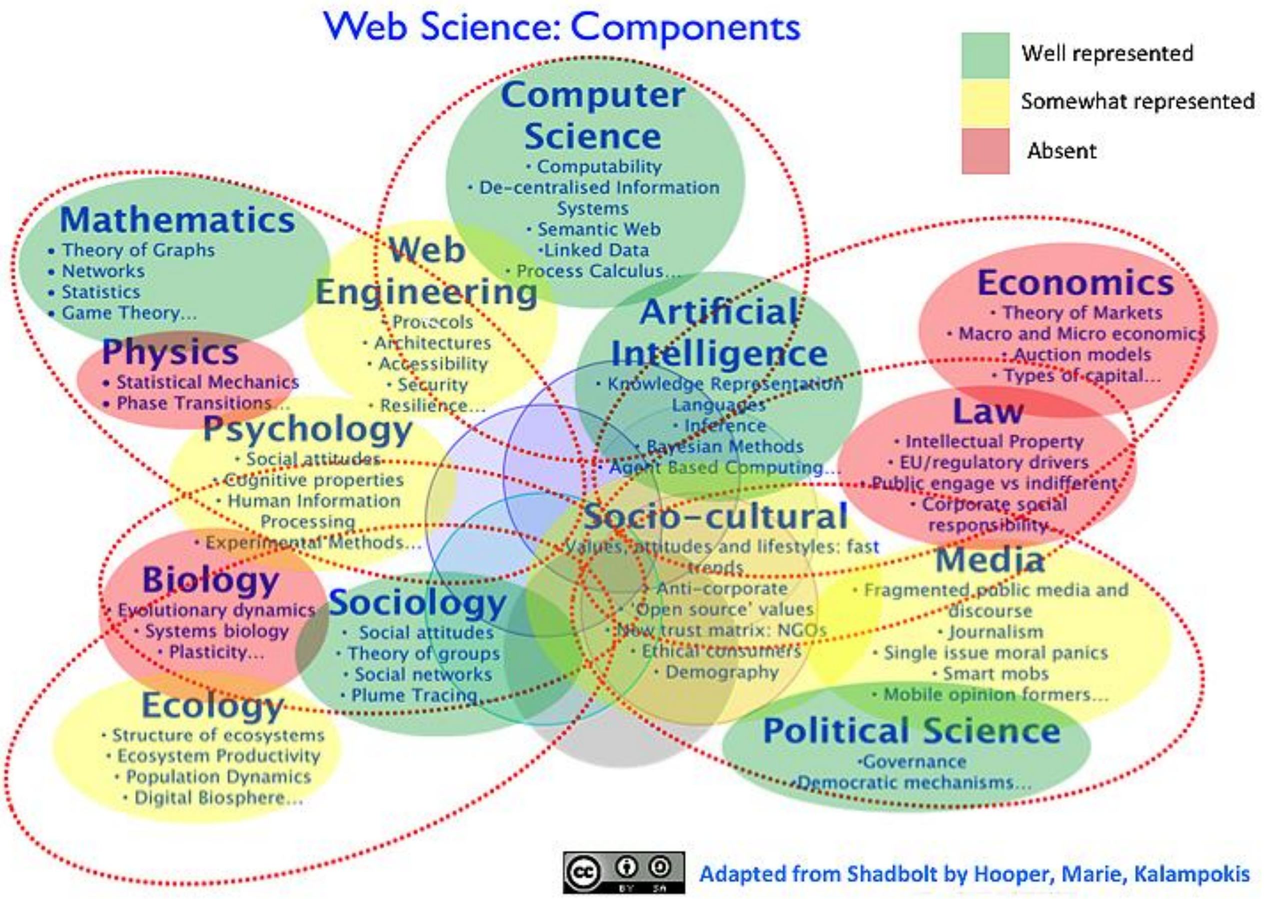


Figure 2: Revised Web Science Cluster Diagram – (C. J. Hooper, Marie, & Kalampokis, 2012)

An interesting point worthy of note, is that Hooper's study shows economics to be "more or less absent" from their version of the Web Science Cluster diagram, therefore indicating it to be poorly represented within the early Web Science conferences. Despite this absence, (Bazan & Vafopoulos 2013) chose to base the entire curriculum of their MSc in Web Science around Economics, indicating its relevance to Web Science. This could be due to the fact that Bazan and Vafopoulos publish papers outside the ACM Web Science conference. This reinforces the need to create a wider definition of Web Science, not only including the WSSC and Web Science conferences, perhaps also increasing the scope of research to include Web Science taught content. It is inevitable in an interdisciplinary field of study that authors would publish beyond a single community, and might still find their primary allegiance outside the emerging Web Science Community. This type of behaviour and allegiance is explored extensively by (Becher & Trowler 2001) in their book: ‘Academic Tribes and Territories’. (Hooper, Dix, et al. 2013) provide a key example of this in their 2013 paper, when during their quest to find Web Science related material to include in their analysis; they discovered six key Web Science papers which were not present within: journal.webscience.org. They explain that these "papers were drawn from the recommended reading list of a forthcoming encyclopaedia article on Web Science."

(Hooper et al. 2012) describe some problems that they encountered when mapping the Web Science discipline, explaining that "topics do not necessarily directly map to disciplines." Another interesting argument that they present, is the fact that the number of papers present is not necessarily a good measure of 'impact'. For example, a conference may include several papers on social networking and only one paper on 'Internet law'; if the paper on Internet law happens to be more influential and noteworthy, it may be considered of greater importance than all the other examples put together. This highlights the fact that it is very difficult to rate content using an algorithm. This is due to the fact that automated methods of scanning content lack human interpretation, and therefore are incapable of making distinctions between relevant and irrelevant publications. Unfortunately, there is no immediately obvious solution to this; manual searches and rating of documents, is by nature very time consuming. This is presumably partly why the Web Science curriculum repository proposed by (White et al. 2010) involves crowdsourcing of content and includes the suggestion that users may be encouraged to rate external content according to the WSSC terms; although this in itself may be limited by the fact that the WSSC could potentially be missing key subjects. This is clearly an area which requires further study; projects such as WSCCR and the research by Hooper et al. show promising inroads into what Hooper et al. describe as a previously untouched avenue of research; however, it is clear that more needs to be done to promote and encourage research in this field.

Hooper et al. expanded the scope of the initial (Hooper et al. 2012) research in their 2013 paper (Hooper, Dix, et al. 2013). The initial (Hooper et al. 2012) paper included an analysis of only the 69 papers included at the Web Science conferences from 2009 to 2011. Within the 2013 paper, (Hooper, Dix, et al. 2013) this was expanded to include 469+ articles, including the 2012 Web Science conference. As with the previous paper, this example also used natural language processing to gather data from this wider range of sources, before analysing and presenting the data using graphing and visualisations. The study uses a tool called ‘Saffron’ which “uses information extracted from unstructured documents using Natural Language Processing techniques" and also uses network graph tool ‘Gephi’ to “build a graph showing links between terms: nodes are extracted terms and arcs are papers that link them." Although it does provide numerous visualisations showing the importance of key terms, regrettably, this paper does not include a list of the actual frequency of key phrases as seen in the previous example, and therefore does not allow for a direct comparison with their previous year’s results. Another interesting observation regarding Web Science conferences made by Hooper et al. (Hooper et al. 2012), is the notion that "collocating with ACM WWW in 2010 may have resulted in more technical submissions." This sentiment also echoed by the later (Hooper, Dix, et al. 2013) paper, which explains that the Web Science conference was collocated with "WWW in 2010, Hypertext in 2011, NetSci in 2012, and CHI this year, 2013)" This lead the authors to theorize that "*co-location with other events influences what disciplines (as measured by terms) are present at WebSci*."

3.8 The rapid evolution of Web Science: A Bottom up Approach Required?

A major problem faced when studying Web Science, is the fact that the Web is constantly evolving, hence it is very difficult to track its development, and any static curriculum soon becomes outdated. "*The Web is different from most previously studied systems in that it is changing at a rate that may be of the same order as, or perhaps greater than, even the most knowledgeable researcher's ability to observe it*." (Hendler, et al 2008) They also go on to state that "*Web scientists need new methodologies for gathering evidence and finding ways to anticipate how human behaviour will affect development of a system that is evolving at such an amazing rate*." Despite the rapid evolution of Web Science, the core concepts will arguably remain the same for several years, as suggested by the concept of 'immutable mobiles', (Halford et al. 2010) and curriculum resources would still be a useful point of reference; although would be of greater relevance were they updated and maintained.

A traditional top down research and development approach was used as the basis for evolving the Web Science Subject Categorization (Vafopoulos 2010). Traditional methods of curriculum development typically involve multiple iterations; a panel of experts however, this approach can lead to the subsequent curriculum becoming quickly outdated. "*Curriculum development, maintenance and management are time -consuming and labour-intensive activities* resulting from countless feedback-rework cycles". (Tang n.d.) A study by (White et al. 2011) proposes the development of a different type of subject definition for Web Science. As detailed previously, they present proposals for the creation of a Web Science curriculum repository, modelled upon the categories of the WSSC, and hosted utilising the University of Southampton ePrints document repository platform. This repository would effectively reverse engineer a Web Science curriculum from the actual teaching materials used.

As with most other work describing Web Science, the (White et al. 2011) paper begins by stating that Web Science is inherently interdisciplinary. The authors then list what they believe are the key attributes of Web Science: "*Understanding the Web growth mechanisms*" and "*Creating approaches that allow new powerful and more beneficial mechanisms to occur*." They explain that teaching Web Science incorporates two distinct features; the microscopic and macroscopic. (White et al. 2011) echo the sentiments of the earlier (Halford et al. 2010) paper, stating that in order to fully understand the Web, it is necessary to consider the technological platform and applications at a microscopic level. It is also equally vital to consider the wider macroscopic implications of the microscopic developments on the macroscopic level of its impact on wider society. What consequences do microscopic developments in technology have upon wider society, and how can these be better influenced in order to benefit society? This echoes the sentiments of the original vision set out by (Berners-Lee, Weitzner, et al. 2006) in their Framework paper; approaching the issue from a teaching perspective.

(White et al. 2011) stress the need for interdisciplinary and multinational cooperation within Web Science. As a means of realizing this collaboration, they propose the creation of a curriculum repository for Web Science, which they term 'WSCR.' The aim of WSCR was to facilitate interdisciplinary and multinational cooperation between Web Science teaching institutions across the globe, by providing a shared curriculum repository in which all Web Science partner institutions can contribute and deposit resources relating to Web Science. This suggestion also echoes the sentiments (Shadbolt & Berners-Lee 2008), who in their paper, suggested that "*part of the new discipline should be to find the most powerful concepts that will help the science itself grow. Perhaps insights will come from the work's interdisciplinary nature."*

A later publication by (White & Vafopoulos 2012) provided additional explanation and justification for the approach proposed to create the WSCR repository outlined in (White et al. 2011). They explain that because of the rapid evolution of Web Science and the lack of existing definitions, it is beneficial to adopt a bottom up or ground up approach to defining the curriculum. In a departure from traditional methods of curriculum creation, they proposed that the actual teaching content be utilised to create the curriculum, in a bottom up, reverse engineering fashion. The "bottom up approach" which is required is a consequence of the fact that Web Science is such a rapidly evolving subject; the curriculum is shaped by the subject itself, as opposed to the subject evolving from the curriculum. The rapid evolution of Web Science is why work undertaken during this thesis utilises a similar bottom up approach, in order to compile an overview of the subject by looking at the material available in order to define the curriculum. This recommended bottom up approach is outlined in more detail in chapter 6.

3.8.1 The Future of the Web and Web Science

One of the most important requirements of Web Science, is not only to outline and study the Web as it is in its current state, but also to ask the question, where is the Web going? How will it develop in the future, and what will influence its development? The fact that the Web is constantly evolving, makes this all the more important. Because the Web now impacts the lives of so many people across the world, it is not only important to try to determine how the Web may change, but also to identify *what* may cause it to change, as this will help facilitate predict as to *how* it might change. During the 2015 Web Science conference, (See Appendix 16.3.1 Web Science 2015 - for trip notes), a key theme reflected throughout, stemmed from a quote from Tim Berners Lee '*the web we want'*[14] in which Tim Berners Lee argues for the need for what he describes as an '*online Magna Carta'*. It was suggested that it is the task of the Web Science community to ask the question of '*what is the Web we want*?' In order to positively influence the development of the Web, and a lot of emphasis was placed upon asking what we would like the Web to be, and how can we achieve this? This discussion was further promoted during the session for the paper 'Developing the Pro Human Web', (Day et al. 2015). The authors suggest that that educating users in the Web and its use and dynamics is vital to creating the Web we want. (Day et al. 2015) point out that programming is being introduced in schools from an early age in order to promote digital literacy, and suggests that the same should be true for Web Education. They state that as with any subject, education is key, and starting at a young age promotes greater awareness

[14] http://www.theguardian.com/technology/2014/mar/12/online-magna-carta-berners-lee-web

During the 2015 Web Science conference, there was also a significant amount of content dedicated to predicting the future development of the Web. An overall summary would suggest that there will be more personalization of web results, more location based sharing websites, more specialized services offered privatized browsing, as well as the continued evolution of Web Technologies and Social Networking. The most widely suggested course of Web Science evolution, is the greater access from mobile devices or Mobile Web and consequently a greater integration of the Web with everyday lives. A common conclusion is the fact that the Web will continue to become more integrated with other media platforms such as interactive television. This is referred to as the 'Internet of things', which suggests that ultimately all household appliances could be connected to the Web so that they can be used anywhere.

As seen through this literature review, the Web is by nature not a static phenomenon. It is always evolving and changing. As explained by (Halford et al. 2010), one factor that influences the Web and the way that it works are the people. Although the Web is referred to as the 'World Wide Web'; it is not currently evenly distributed across the entire world. The Web is currently heavily dominated by the West, and is therefore culturally influenced and biased towards the West. It is arguable therefore; that an important question to be asked by Web Science should be how new countries adopting the Web will affect the cultural bias of the Web. It is also suggested that the Web could evolve from a 'Web of documents' to a 'Web of Data'; (Halford et al. 2010) instead of searching through documents which match a key search term, people might be able to ask a question about a specific thing or person, and be provided with an answer provided by machine reasoning, as opposed to having to manually search through a list of related documents produced in a list of search result.

3.9 Conclusions/Future Work

The following section will look at the key conclusions drawn from examining the available Web Science literature, which scrutinises the creation and definition of Web Science as a subject.

3.9.1 The Defining Features of Web Science

All publications relating to the definition of Web Science as a subject are agreed that Web Science is by nature a Socio-Technical subject. The purpose of Web Science is to engineer the 'Web we want', it should focus on understanding how the Web impacts society, and how society impacts the Web. Web Science should include the study of the micro and the macro realms; with the micro representing the study of technology, and the macro representing the effect of that technology on wider society into which the technology is introduced.

3.9.2 Lack of definition

As seen from this literature review, there are very few existing attempts to define the scope of the Web Science subject. The only two current attempts to create subject representations outlining the scope of Web Science, are the Web Science Subject Categorisation, (WSSC) and the Web Science Cluster diagram. The Cluster diagram is merely a visual overview of Web Science, and while it is a good resource for quickly gaining a visual understanding of Web Science, it lacks detail, and is also missing subjects. The Web Science subject categorization is currently the only detailed attempt to provide a detailed outline of Web Science, however, it too is missing subjects and is now considered outdated, and is no longer hosted by the Web Science Trust. WSSCR, (White et al. 2011) was another attempt to define Web Science, by emerging a Web Science curriculum utilising a ground up approach, however no further news of developments can be found beyond the initial paper.

Both the WSSC and WSSCR would have provided valuable resources for the Web Science community had they been maintained (WSSC) or completed (WSSCR). Both projects propose useful concepts which may be drawn on. The notion of a referenceable taxonomy outlining the scope Web Science and providing links to Web Science related content is a valuable concept which still has scope for development. Were such a concept to utilise with the bottom up approach prosed for WSSCR (White et al. 2011), the concept of the taxonomy might be realised more rapidly. The idea of (White et al. 2011) to crowdsource content would also provide sustainability, were the resulting taxonomy presented in a format which could be edited beyond the initial project lifespan. Any such resource would not be a watertight and completely definitive representation of Web Science. However, it would serve as a platform for encouraging further contribution and discussion. It would also provide a starting point for anyone wishing to gain an overview of the Web Science subject, and a useful reference resource for those studying or teaching Web Science.

Despite the lack of sources detailing the scope of Web Science, (White & Vafopoulos 2012) highlight that there are an increasing number of Web Science taught programmes being offered worldwide, and therefore these can be utilised as a data source for identifying the constituent subjects which are taught as part of Web Science. This combined with an analysis of Web Science publications could provide a substantial dataset which would provide detailed insight into Web Science, from what is taught and published within the scope of the discipline.

3.9.3 Rapid evolution

Because of the nature of the Web, Web Science is constantly evolving and changing, requiring a flexible approach to research, which is able to adapt and reflect the rapid changes in the research subject. This involves taking a ‘bottom up’ approach (i.e. evolving the curriculum from looking at the current material available.) This is reflected in this author’s choice to adopt a mainly qualitative, Grounded Theory based research approach, incorporating mixed methods. Traditional quantitative research methods focus on forming a hypothesis, and then setting out to prove it. While this can still be done when studying the Web, it is also useful to look at the bigger picture, and make deductions from observing the Web itself; therefore, drawing conclusions from observing the problem, (qualitative) as opposed to forming a hypothesis and setting out to prove it (quantitative).

4 Further Background Research: Subject Comparison and Classification of Disciplines

4.1 Introduction

In order to understand how Web Science might evolve as an academic discipline, it is helpful to examine how other related academic disciplines have evolved. Following this, there is a review of the similarities between Web Science, and related subjects which demonstrate some overlap in topics taught, including Internet Science and Network Science. Both of these examples are also relatively recent disciplines, and as is also the case with Web Science, neither Internet Science or Network Science yet have examples of formal curriculum guidelines. Therefore, this chapter also examines existing attempts to define curricula for other subjects which may share some element of overlap with Web Science. Resources which are examined include the ACM Computer Science Guidelines and the ACM guidelines for Information Science. These are potentially the most similar examples of structured curriculum guidelines to Web Science, and the Web Science conference is also sponsored by the ACM. In order to discover how much (if any) overlap is demonstrated between Web Science and related subjects, a comparison is made between example modules from Computer Science and Information Science.

In addition to examining how related curricula are defined and how these relate to Web Science, this chapter also examines Biglans's proposals for classifying academic disciplines. Biglan proposes categories for classifying academic disciplines, which will prove useful when examining the nature of Web Science and its component subjects. Therefore an initial overview of Biglan's framework is provided, prior to the classification of Web Science Key and Paper topics according to Biglan's framework, which will be outlined later in chapter 9.

This chapter has the following structure:

- Evolution of disciplines
- Three Subjects: Comparison: Web Science vs Internet Science and Network Science
- Overview of Existing Curriculum Guidelines
    - The ACM Computer Science Curriculum CS2013
    - The ACM Information Science Curriculum
- Biglan's Classification of Academic Disciplines
    - Application for Web Science
- Wikipedia's Outline of Science

4.2 Evolution of Disciplines

The following section examines the process by which subjects evolve into academic disciplines, in order to understand how Web Science might similarly evolve.

(Becher et al. 2001) suggest that the definition of an academic discipline is not always straightforward. They provide the example of the subject 'Statistics' which originates from Mathematics, and suggest that whether or not a University recognises Statistics as a subject in its own right, stems from whether or not a University has a department dedicated to a subject. Therefore, using this criteria, if statistics is a sub-subject provided by a Mathematics department, it would suggest that it is considered by that University as a subsidiary of Mathematics, whereas if a University has a dedicated Statistics department, then it is considered a subject in its own right. Using this same criteria, it could be said that Web Science is considered a dedicated discipline in its own right by, for example, the University of Southampton, as this institution has a dedicated Web and Internet Science research group. This is just one way of determining the status of an academic subject. (Becher et al. 2001) suggest that while this is one method of determining the status of a subject, it is not the only method. (Becher et al. 2001) explain that:

> "*international currency is an important criterion, as is a general though not sharply defined set of notions about academic credibility, intellectual substance, and appropriateness of subject matter*".

Various factors will affect whether a subject is considered a subsidiary subject, or a subject in its own right. These factors include funding, overlap between disciplinary modules, and subject popularity with students and also potentially interdisciplinary collaboration and affiliation at conferences. (Becher et al. 2001) also recognise that academic disciplines are by no means stable entities and are subject to both "historical and geographical" variation. Despite evolution over time, there is generally a consistent pattern to development of a discipline. The same is true for global differentiation, while practices may differ between countries, a discipline will maintain the same fundamental core concepts. Some related disciplines have well defined borders, and others have what (Becher et al. 2001) describe as 'ragged' borders. Overlap between related disciplines can lead to conflict between differing disciplinary schools of thought and practice. However, in other instances, overlapping borders between disciplines can lead to collaboration between related disciplines. This kind of interdisciplinary collaboration is what Web Science aims to promote.

4.2.1 Academic Discipline Creation and Evolution

(Krishnan 2009) examines academic disciplines from a number of different perspectives. These perspectives include, Philosophical, Anthropological, Sociological, historical, management and education. (Krishnan 2009) propose six of the key features which they suggest as defining an academic discipline. They advise that the term 'academic discipline' has become associated with the "*organisation of learning and the systematic production of new knowledge*." It is suggested that an academic discipline may often be associated with a 'taught subject' at a university, however, not all university taught programmes may necessarily represent a distinct academic discipline. (Krishnan 2009) recognise that not all academic disciplines will include the features outlined in Table 5, these are just some of the key criteria by which a discipline may be identified.

(Krishnan 2009) suggests that there are considerable variations between disciplines when subjects' standing in universities is considered. (Krishnan 2009) explain that factors such as the number of students attracted, the number of academics and the amount of research and funding for a discipline can be used as benchmarks to measure to standing or popularity of a discipline. They suggest that larger departments with more expensive equipment and more funding tend to have a greater influence within the wider academic community for that discipline than smaller, lesser funded departments. Also, the number of institutions which have departments dedicated to a specific subject can provide an indication of popularity. (Krishnan 2009) also suggest that discipline differentiation can be attributed to differing research practices, Table 5) citing the examples of Sociology and Anthropology. They explain that one of the key aspects that differs Anthropology from Sociology, is the research practice known as Ethnography, which was first established and practiced by Anthropologists.

Table 5 Principles for Discipline Creation

6 Principles for Discipline Creation (Krishnan 2009) (page 10) http://eprints.ncrm.ac.uk/783/1/what_are_academic_disciplines.pdf
1) disciplines have a particular object of research (e.g. law, society, politics), though the object of research maybe shared with another discipline;
2) disciplines have a body of accumulated specialist knowledge referring to their object of research, which is specific to them and not generally shared with another discipline;
3) disciplines have theories and concepts that can organise the accumulated specialist knowledge effectively;
4) disciplines use specific terminologies or a specific technical language adjusted to their research object;
5) disciplines have developed specific research methods according to their specific research requirements; and maybe most crucially
6), disciplines must have some institutional manifestation in the form of subjects taught at universities or colleges, respective academic departments and professional associations connected to it. Only through institutionalisation are disciplines able to reproduce themselves 'from one generation to the next by means of specific educational preparation.

(Krishnan 2009) suggest that one of the key points at which a subject becomes a recognisable academic discipline, is the point at which an academic department dedicated to that given subject is founded at an academic institution. The number of institutions which then found departments dedicated to that given subject, provides an indication of the subject's prominence within academic circles. It is logical therefore, that the more universities which then found departments dedicated to that subject, the more that the popularity of that subject grows. However, what causes a university to create a department dedicated to a new subject in the first place?

4.2.2 The Development of Computer Science as an Academic Discipline

The following section examines the development of Computer Science as an academic discipline, as it is an example of a technological discipline which is still relatively young, and also features some overlaps with Web Science.

(Denning et al. 1999) suggests that Computer Science originated in the early 1940's with the convergence of algorithmic theory and mathematical logic, and the invention of what he terms the "stored-program electronic Computer". (Denning et al. 1999) then goes on to explain that "By the early 1960's, there was a sufficient body of knowledge to merit the first academic departments and degree programmes."

An article by IBM [15] pinpoints the foundations of Computer Science, attributing it to an earlier meeting between academic Benjamin Wood, a professor at Columbia University, and IBM chief executive Thomas J. Watson in 1928. Wood convinced Watson that IBMs machines of the time could be used to "*measure anything represented by mathematics, including not just business activities but physics, biology, astronomy or any other scientific discipline*". While it must be considered that this article is biased towards IBM, having been written by them, it's implications for Computer Science are indeed notable. It is evidence of a very early collaboration between a company developing computing technology, and an academic institution, which then resulted in IBMs collaboration with academia. The article states that "*the close relationship between IBM and the university was instrumental in the emergence of a new scientific discipline, with Columbia offering one of the first academic-credit courses in computing in 1946."* IBM also began to run their own training programmes in order to instruct clients how to utilise their products.

An article from the University of Birmingham [16], charts the development of Computer Science at that academic institution. The article states that the first lecturer in Computing was appointed in 1959. This position was created within the department of Mathematical Physics. Following this, in 1961 an English Electric KDF9 was purchased by the University. In 1965, this English Electric KDF9 passed tests and entered use within the institution as an 'extra-faculty unit'. This resulted in several academics from 'the department of Mathematical Physics' being transferred to form the newly created unit. Programming courses were initially offered to all academic staff and researchers. The department also offered *"several optional courses on programming and other aspects of computer science that were introduced into the degree programmes in Mathematical Sciences during the 1960s*." In 1969, the Computer Centre at the University of Birmingham was given the status of an academic department, as opposed to merely a service department, for the purpose of offering degree programmes. The teaching of Computer Science taught programmes at Birmingham then began with an MSc programme in Computer Science. The first intake of students took place "*in October 1969*". It is interesting to note, that much like Web Science often has at the institutions in which it is taught, Computer Science began its taught life at the University of Birmingham as an MSc programme.

This paper is an attempt to define Computer Science and break it down into its constituent components. The guidelines set out by (Denning et al. 1989) recognise that Computer Science is a rapidly evolving discipline, and proposes a teaching paradigm for the subject. This highlights the similarity with Web Science in terms of a rapidly evolving discipline. (Denning et al. 1989) underwent a process of dividing the Computer Science discipline into sub-areas. They describe the fact that they decided that Computer Science could be characterised by three key processes, which included 'theory', 'abstraction' and 'design', and that each of the curriculum sub-areas should include these three basic processes. In order to illustrate this, (Denning et al. 1989) developed a matrix structure (Table 6) showing their proposed curriculum model for Computer Science. This is a similar matrix

[15] http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/compsci/

[16] https://www.cs.bham.ac.uk/about/history/

idea to the example proposed by (Heredia & Vinueza 2015) for use within the proposed Web Science Lexicon. (See section 9.2.8 for further information).

Table 6 Definition Matrix for the Computing Discipline (Denning, 1989)

Definition Matrix for the Computing Discipline (Denning, 1989)			
	Theory	Abstraction	Design
Curriculum Model for Computer Science (Denning et al, 1989)			
1 Algorithms and data structures			
2 Programming languages			
3 Architecture			
4 Numerical and symbolic computation			
5 Operating systems			
6 Software methodology and engineering			
7 Databases and information Retrieval			
8 Artificial intelligence and robotics			
9 Human-computer communication			

The curriculum model matrix shown in (Table 6) was then updated with new topics in the 1999 version of the (Denning 1999) paper. Table 7 shows a comparison between the lists of topics from the two versions of the paper. The two versions illustrate the evolution within the subject over the 10-year year period between the two publications. A notable feature in the evolution between the two versions, is the inclusion of Networks within the 'Operating systems' section. This reflects the development of the Web between 1989 and 1999, a consequence of which, means that the 1999 version reflects a greater level of Web connectivity and hence the importance of Networks within the curriculum. The inclusion of Graphics reflects the evolution in technology and the ability to provide enhanced GUIs and present information in a more visual way. The inclusion of the final three entries, 'Computational Science', Organizational Informatics', and 'Bioinformatics', perhaps reflects the increasing application of Computer Science in other disciplines, such as bioinformatics, which is the convergence between Computer Science and Biology.

Table 7 Definition Matrix for Computer Science Comparison

Definition Matrix for Computer Science (Denning et al, 1989)	Definition Matrix for Computer Science (Denning et al, 1999)
1 Algorithms and data structures	1 Algorithms and data structures
2 Programming languages	2 Programming languages
3 Architecture	3 Architecture
4 Numerical and symbolic computation	4 Operating systems and Networks
5 Operating systems	5 Software Engineering
6 Software methodology and engineering	6 Databases and information Retrieval
7 Databases and information Retrieval	7 Artificial intelligence and robotics
8 Artificial intelligence and robotics	8 Graphics
9 Human-computer communication	9 Human-computer communication
	10 Computational Science
	11 Organizational Informatics
	12 Bioinformatics

(Krishnan 2009) suggest that new technological and managerial disciplines are growing in popularity because of their perceived usefulness with regards to future employment. They also suggest that Scientific disciplines have stronger, more well-defined paradigms, and are therefore better connected and more respected. (Bargh, Catherine; Scott, Peter; Smith 1996) suggest that new subjects aimed at student popularity and post education employability are emerging. Computer Science developed because it represented a way of learning how to utilise new developing technology, in order to apply it to solve real-life problems. The same could be said of Web Science. Web Science was proposed "as a way to understand the Web, and as a way to focus its development on key communicational and representational requirements." (Berners-Lee, Weitzner, et al. 2006) The mission statement at the end of (Denning et al. 1999) suggests that "*The fundamental question underlying all of computing is, what can be (efficiently) automated*?" It could be said that the fundamental purpose of Web Science is to engineer "the Web we want", according to the words of the Web's creator, Tim Berners-Lee. It could be said that the scope of Web Science is much larger than that of Computer Science. While Computer Science often focuses designing and engineering the base technological platforms, Web Science looks to not only the technology, but the wider impact upon what has become the technological underpinning of society. Computing is no longer a mere tool to improve the efficiency of processes. Online banking, healthcare services, government services, and online social interactions and more, now represent a fundamental part of society and peoples' lives. Therefore, is Web Science the next generation of Computer Science? (Shneiderman 2007). Yes and no. We still need the fundamental basis of Computer Science, how to solve problems utilising technology, although the technology utilised is becoming increasingly Web based. Web Science is the convergence between the technological and the social, in order to understand and drive the impact of the Web. Because the Web impacts so many areas of society, therefore in order to understand the phenomenon, Web Science must also be inherently multidisciplinary.

4.3 Three Subjects: Comparison: Web Science vs Internet Science and Network Science

In a 2015 paper, (Tiropanis et al. 2015) examine the overlap between Web Science, and what they identify as similar disciplines, which include Internet Science and Network Science. The following section provides an overview of Internet and Network Science, and a simple comparison between these two subjects and Web Science.

In the same way that the information systems discipline was considered a bridge between computer science and the business world, Web Science could be said to expand horizons to include the world and impact of the Web, (Oinas-kukkonen 2010). In 2006 Tim Berners-Lee announced the Web Science research intuitive in 2006 in order to study the Web, in an attempt to "*map the boundaries of the Web"* (Wright 2011) and better understand its impact upon society. Other similar initiatives include the EINS (European Network of 'Excellence in Internet Science'), which aims to promote the study of 'Internet Science'. Network Science is another related field of study which "*aims to understand the evolution of networks regardless of where they emerge: the Internet as a network transforming and forwarding information among people and things, and the Web as a network of creation and collaboration*." (Tiropanis et al. 2015)

4.3.1 Internet Science

The goal of Internet Science is to promote interdisciplinary collaboration, in order to gain a broader understanding of the "*Internet as a societal and technological artefact, whose evolution is increasingly intertwined with that of human societies*" (Hooper et al. 2015). A number of peer-reviewed journals dedicated to Internet Science were published between 2006 and 2014. Its scope included presenting "*empirical findings, methodology, and theory of social and behavioural science concerning the Internet and its implications for individuals, social groups, organizations, and society*." [17] There have currently been two dedicated international Internet Science conferences. Similar to the Web Science Trust[18], the EINS (European Network of 'Excellence in Internet Science') is a body which aims to "*strengthen scientific and technological excellence by developing an integrated and interdisciplinary scientific understanding of Internet networks and their co-evolution with society*"[19]. "*The primary objective of EINS is to foster open and productive dialogues between all disciplines that study Internet systems from any technological or humanistic perspective*." (Guevara & Blackwell 2013).

4.3.2 Network Science

Network Science is a far older and more established discipline than either Web Science or Internet Science. While in its current form, it largely focuses upon the study of the effects of Web-based networks; its roots can be traced back to the 18th century to "*Leonard Euler's seminal work on graph theory*" (Tiropanis & Hall 2015). They explain that Network Science relates to the study of the "*emergence, evolution and outcome*" of Networks and their impact upon the wider world. Before the advent of the Web, such studies focused more upon physical and biological networks. There have been a total of ten dedicated 'NetSci' Network Science conferences, the first of which took

[17] http://www.ijis.net/

[18] http://webscience.org/

[19] http://www.internet-science.eu/network-excellence-internet-science

place in 2006. There have also been a number of international workshops dedicated to the study of network theory.[20] [21]

4.3.3 Subject Comparison

Internet Science can be defined as an emerging science, the aim of which is to study the evolution and development of the Internet as "*a social and technological artefact*" (Cîmpan & Salamatian n.d.). This description is very similar to many given for Web Science. There can be considerable confusion between the terms Web Science and Internet Science; just as the terms 'Web' and 'Internet' are often mistakenly used interchangeably, as are the terms 'Web Science' and 'Internet Science'. In a nutshell, Web Science is the study of the World Wide Web and its impact, and 'Internet Science is the study of the Internet and its impact. In order to understand the differences and relation between the two subjects, one must first understand the differences between the Web and the Internet, and how they also relate to each other.

The Internet is a network of networks and technological infrastructure which connects millions of computes across the globe. The World Wide Web is a means of accessing information using the technology that the internet provides using HTTP (Hypertext Transfer protocol) as a standardised means of communicating information. "*The Web is an application developed to let people share information over the Internet.*" (Easley & Kleinberg 2010)

Table 8. Three Sciences: Subsets of each other?

The Three Sciences: Subsets?
"*The precise delineation between Internet Science and Web Science is not always clear. One might argue that Web Science is a subset of Internet Science (because web technologies are a subset of Internet technologies), while at a 2013 Internet Science / Web Science workshop [28] it was argued that one could view Internet Science as a subset of Web Science*." (Hooper et al. 2015) "*In one sense Web Science is a subset of network science. In another sense, network science is a subset of Web Science*." (Wright 2011)

As explained by (Tiropanis & Hall 2015) there is overlap between Web Science, and other related subjects, Internet Science and Network Science. The following table shows a comparison between the descriptions (Tiropanis & Hall, 2015) give for each of the three sciences. It is immediately obvious from reading these descriptions that there are similarities between the three subjects. It is clear from descriptions that all three subjects share attributes, and each subject is described as potentially being a subset of another. (See Table 8. Three Sciences: Subsets of each other?).

[20] http://sonic.northwestern.edu/news/events/webnetsciworkshop/

[21] https://web.archive.org/web/20171120060603/http://www.ctr.kcl.ac.uk/netscicom14/

Table 9. The three sciences: comparison

The Three Sciences: Comparison	
Network Science	"*The Term Network Science emerged as an interdisciplinary area that draws on disciplines such as physics, mathematics, computer science, biology, economics, and sociology to encompass networks that were not necessarily social*." (Tiropanis & Hall, 2015) "*Network Science is technology agnostic and it overlaps only in part with Internet Science and Web Science, since it explores emergent structural patterns and flows on network structures be they social, biological, the Web or the Internet.*" (Tiropanis & Hall 2015)
Internet Science	*"Internet Science is an emerging science targeting the study of Internet as a societal and technological artifact, whose evolution is increasingly intertwined with that of human societies. Internet. Internet Science is by nature strongly multidisciplinary as several disciplines like network engineering, computation, complexity, security, trust, mathematics, physics, sociology, game theory, economics, political sciences, humanities, law, energy, transport, artistic expression, and any other relevant social and life sciences, study Internet Systems."* (Cîmpan & Salamatian n.d.) "*In Internet Science the evolution and sustainability of the Internet and its services are central objectives*" (Clark et al. 2002) (Tiropanis & Hall 2015)
Web Science	*"Web Science draws on disciplines that include the social sciences, such as anthropology, communication, economics, law, philosophy, political science, psychology and sociology but also computer science and engineering. Studying the Web and its impact requires an interdisciplinary approach that focuses not only on the technological level, but also on the societal, political, and commercial levels.*" (Tiropanis & Hall, 2015) "*In Web Science, the study of the Web in itself is crucial (Hendler, Shadbolt, & Hall, 2008), as is safeguarding the Web and its evolution* (Hall & Tiropanis 2012)."

While it is very evident that similar descriptions may be applied to all three subjects, and the terms are sometimes mistakenly used interchangeably, there are some very important fundamental differences. Web Science arguably incorporates both Internet Science and Network Science, as the study of the technologies underpinning the Web as well as network effects are both impact the Web at a macro level, and therefore have significance and relevance within Web Science. The same also applies form both Network Science and Internet Science, each of the three subjects overlap in some regard. (Wright 2011) explain that both Network Science and Web Science are technically subsets of each other, while in practice; "*Web Science is focused on how we could do things better, while Network Science is more focused on how things work*" (Wright 2011).

4.4 Introduction to Analysis of Existing Curricula

There is not yet a definitive curriculum for Web Science; currently the most comprehensive attempt to outline the scope of Web Science is the Web Science Subject Categorization. Because there are currently no ACM curriculum guidelines for Web Science, it is beneficial to examine existing curriculum guidelines from related displaces. The ACM provide extensive curriculum guidelines for Computer Science and related subjects. (White & Vafopoulos 2012) make comparisons between Web Science and Computer Science. These ACM curricula are an appropriate point of comparison with Web Science, given that the ACM sponsor the Web Science conference, and there is likely to be a fair degree of similarity and shared attributes between the subjects. Therefore, the ACM Computer Science and Information Science guidelines are examined as curricula which may share some common elements with Web Science. A comparison is made between the teaching subjects recommended by (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013), ACM curriculum guidelines for Information systems (Topi et al. 2010) and the University of Southampton Web Science programme. The following section will examine and compare the existing curricula in more detail, identifying similarities with Web Science.

4.4.1 The ACM Computer Science Curriculum CS2013

It is useful to understand the process by which 'traditional' curricula are devised, in order to evaluate how these compare with attempts to create a curriculum for Web Science. The ACM and IEEE-Computer Society have for many years been collaborating, in order to establish curriculum guidelines for computer science. This operates on a roughly 10 year cycle, with a new volume of guidelines being published each decade. The most recent example was published in 2013. (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013). They explain that over recent years, the field of computer science has expanded grown and diversified and that "*there are now curricular volumes for Computer Engineering, Information Systems, Information Technology, and Software Engineering in addition to Computer Science*. The CS2013 report includes guidelines for what they describe as the 'essential' ingredients for a computer science curriculum, as well as exemplar modules from a number of computer science related programmes.

A common characteristic shared by this thesis and the work of the CS2013 report, is the compilation of exemplar modules taught by universities teaching modules in the related field. According to (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) "T*he Body of Knowledge (i.e., the outline of topics that should appear in undergraduate Computer Science curricula) was the most used component of the reports.*" This would imply that institutions found the subject guidelines the most useful component, suggesting that the same should also be beneficial for Web Science.

The CS2013 project involves numerous academics from various institutions, including committees from both the ACM and IEEE. The process also takes years to complete. If the same lengthy process were used for Web Science, due to the rapid evolution of the Web, the results would arguably become outdated by the time the exercise were completed. Therefore, this thesis utilises a 'Bottom up' approach to research, employed over a much shorter time period. The (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) identifies a set of what they term 'knowledge areas'. These are outlined in Table 10. Knowledge Areas for the Computer Science curriculum. It should be noted that knowledge areas are subject related, and are not to be confused with ‘learning outcomes’, similar to those specified in models such as Blooms Taxonomy.

Table 10. Knowledge Areas for the Computer Science curriculum

Knowledge Areas According to (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013)
“The CS2013 Body of Knowledge is organized into a set of 18 Knowledge Areas (KAs), corresponding to topical areas of study in computing. The Knowledge Areas are: *● AL - Algorithms and Complexity* *● AR - Architecture and Organization* *● CN - Computational Science* *● DS - Discrete Structures* *● GV - Graphics and Visualization* *● HCI - Human-Computer Interaction* *● IAS - Information Assurance and Security* *● IM - Information Management* *● IS - Intelligent Systems* *● NC - Networking and Communications* *● OS - Operating Systems* *● PBD - Platform-based Development* *● PD - Parallel and Distributed Computing* *● PL - Programming Languages* *● SDF - Software Development Fundamentals* *● SE - Software Engineering* *● SF - Systems Fundamentals* *● SP - Social Issues and Professional Practice”*

The (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) guidelines explain that ‘knowledge areas’ should not necessarily be associated with a specific module, and that they in fact discourage this practice. They explain that “*We view the hierarchical structure of the Body of Knowledge as a useful way to group related information, not as a stricture for organizing material into courses*.” Even though this may be the case with some modules, and there will in many cases be some courses containing material from only one knowledge area, as a general rule, it is expected that modules should include material covering multiple knowledge areas, and that this may especially be the case in introductory modules. As explained previously, knowledge areas subject related and are therefore different to ‘learning outcomes’ described by Blooms Taxonomy. (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) explain that while they explored and utilised curriculum approaches, such as Bloom’s Taxonomy, they have not directly applied models such as Blooms within their report “*in part because several of them are driven by pedagogic context, which would introduce too much plurality in a document of this kind.”* They have instead derived their own ‘Levels of Mastery’ scale, outlined in Table 11.

Table 11. 'Levels of Mastery' according to the ACM Computer Science Curriculum.

'Levels of Mastery' from (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013)
- *"**Familiarity**: The student understands what a concept is or what it means. This level of mastery concerns a basic awareness of a concept as opposed to expecting real facility with its application. It provides an answer to the question "What do you know about this?"* - ***Usage**: The student is able to use or apply a concept in a concrete way. Using a concept may include, for example, appropriately using a specific concept in a program, using a particular proof technique, or performing a particular analysis. It provides an answer to the question "What do you know how to do?"* - ***Assessment**: The student is able to consider a concept from multiple viewpoints and/or justify the selection of a particular approach to solve a problem. This level of mastery implies more than using a concept; it involves the ability to select an appropriate approach from understood alternatives. It provides an answer to the question "Why would you do that?""*

The 'levels of mastery' shown above are the (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) equivalent of a leaning model used to assess learning objectives.

4.4.2 The ACM Information Science Curriculum

The (Topi et al. 2010) report explains that there is a close relationship between Information Systems and other Computing subjects. However, while there are many similarities shared between Information systems and for example, Computer Science, there are also some key differences. (Topi et al. 2010) explain that "*The context for Information Systems is an organization and its systems. In contrast, the context for Computer Science would typically include algorithmic processes for information processing and associated technical issues*." (Oinas-Kukkonen 2010) explain that Information Systems "*also resembles web science in its multidisciplinary nature and strong utilization of reference disciplines."*

> "*Web science may learn substantially from the methods and approaches in information systems. In a manner similar to how many think of web science currently, information systems was years ago considered only as a bridge between the business world and computer science. This provides a model that could also apply in the development of web science from a multidisciplinary field as it is now into a sound discipline."* (Oinas-Kukkonen 2010)

4.5 Curricula Comparison

(Chin et al. 2007) explain that due to rapid advances and changes in technological development, students are now confronted with many differencing options for study when choosing a technology-based degree course, and that the differentiation between programmes is not always clear. They also explain that universities often struggle to ensure that their programmes continue to *"stay aligned*" with advances in technology. They therefore propose an ontology for differentiation between programmes, which they term "*a high-level IT (information technology) pedagogical knowledge base that will make it possible to clearly distinguish between computer-related pro-*

grams of study." (Oinas-Kukkonen 2010) and (The Joint Task Force for Computing Curricula 2005) explain that the ACM has "*defined five sub disciplines of computing*". These are shown in the diagram displayed in Figure 3.

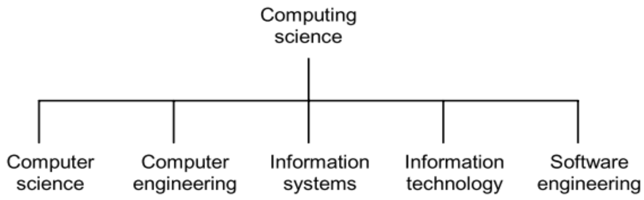


Figure 3. Sub-disciplines in computer science

(Tiropanis & Hall 2015) suggest that there are similarities shared by the subjects Web Science, Internet Science and Network Science. Both Internet Science and Network Science are also emerging subjects. While Network Science has existed for many years in the form of mathematical graph theory, with applications in subjects such as Chemistry, its application to online networks is still relatively recent. Internet Science is also a young discipline, with the first international conference in Internet Science being convened in 2013, with a second in 2015[22]. Despite some obvious overlaps with Computer Science, neither of these two subjects are represented by the ACM family of computing disciplines shown in Figure 3.

4.5.1 ACM Curriculum Guidelines

Because Web Science, Internet Science and Network Science are relatively new disciplines, there is little material and no examples of ACM curriculum definitions available. Therefore, in order to examine the process of curriculum creation, the ACM Computer Science (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013) and Information Science curriculum guidelines (Topi et al. 2010) for BSc were used as a reference. Although different to Web science, these fields do share some attributes with Web Science.

A comparison between the subjects was made, in order to provide an initial overview of the amount of shared modules the subjects may have in common. The large number of choices referred to by (Chin et al. 2007) are reflected in both the ACM curriculum guidelines and example course syllabi investigated. The ACM Computer Science curriculum includes a very comprehensive list of suggested modules, however, there are too many to alternatives to facilitate a direct comparison. Therefore, the University of Southampton Computer Science BSc [23] was used as an example, as it follows the ACM guidelines. The large number of optional modules offered, still consequently make it difficult to formulate a concise comparison between the subjects. Therefore, only core and compulsory modules were included.

Table 12 shows the comparison of key modules for Computer Science, Information Science and Web Science. The Information Science column is drawn from the list of 'recommended modules' provided by the ACM Information Science curriculum guidelines for BSc (Topi et al. 2010). (The variations in colour depict an overlap in subjects between programmes.) Because there is no example of a Web

[22] http://internetscienceconference.eu/insci2015.html

[23] https://web.archive.org/web/20160222141213/http://www.ecs.soton.ac.uk/programmes/bsc-computer-science#modules

Science curriculum definition comparable with that of the ACM Computer Science and ACM Information Science curriculum guidelines, the University of Southampton Web Science BSc programme modules have been used as an example. The University of Southampton offers two variations of Web Science BSC; Social Science[24], and Computer Science[25]. Both biases have been included for contrast. This initial comparison superficially suggests that there is little direct overlap between Web Science, Computer Science and Information Science; although some of the same broad topics are covered. However, a more detailed understanding and investigation of module content may yield a more accurate comparison. The initial lack of overlap in the curricula is evidence in favour of the need for Web Science as a separate taught subject.

24 https://web.archive.org/web/20161018150409/http://www.southampton.ac.uk/webscience/undergraduate/courses/i200_bsc_web_science_social_science.page

25 https://web.archive.org/web/20161005230228/http://www.southampton.ac.uk/webscience/undergraduate/courses/i201_bsc_web_science_computer_science.page

Table 12. Comparison of Knowledge Areas for Computer Science, Information Science and Web Science

Comparison of Key Modules for Computer Science, Information Science and Web Science BSc			
ACM Computer Science	**ACM Information Science**	**Uni of Southampton: BSc Web Science (Social Science)**	**Uni of Southampton: BSc Web Science (Computer Science)**
Core: COMP1202 Programming I	**CORE: IS 2010.1 Foundations of Information Systems**	**PAIR1002 Political Systems**	**Core: COMP1202 Programming 1**
Core:COMP1203 Computer Systems I	**CORE: IS 2010.2 Data and Information Management**	**WEBS1001 Information, Technology and Social Change**	**Core: WEBS1001 Information, Technology and Social Change**
Core: COMP1205 Professional Development	**CORE: IS 2010.3 Enterprise Architecture**	**COMP1056 Web Design**	**Compulsory: COMP1056 Web Design**
Core: COMP1215 Foundations of Computer Science	**CORE: IS 2010.4 IS Project Management**	**SOCI1002 Transformations of The Modern World**	**Compulsory: COMP1206 Programming 2**
Core: COMP1201 Algorithmics	**CORE:IS 2010.5 IT Infrastructure**	**STAT1003 Introduction to Quantitative Methods**	**Compulsory: SOCI1014 Foundations in Social and Anthropological Theory: Traditions of Thought and Argument**
Core:COMP1204 Data Management	**CORE: IS 2010.6 Systems Analysis and Design**	**SOCI2031 Social Theory**	**Compulsory: STAT1003 Introduction to Quantitative Methods**
Core:COMP1206 Programming 2	**CORE: IS 2010.7 IS Strategy, Management and Acquisition**	**WEBS2001 Cybernetics, Societies and the Web**	**Core: WEBS2001 Cybernetics, Societies and the Web**
Core:COMP1216 Software Modelling and Design	**Application Development**	**WEBS2002 Interdisciplinary Group Project**	**Core: WEBS2002 Interdisciplinary Group Project**
Compulsory: COMP2207 Distributed Systems and Networks	**Business Process Management**	**COMP2213 Interaction Design**	**Compulsory: COMP2213 Interaction Design**
Compulsory: COMP2208 Intelligent Systems	**Enterprise Systems**	**SOCI2020 Research Skills**	**Compulsory: SOCI2020 Research Skills**
Compulsory: COMP2209 Programming III	**Introduction to Human-Computer Interaction**	**UOSM2008 Living and Working on the Web**	**Compulsory: UOSM2012 Online Social Networks**
Compulsory: COMP2210 Theory of Computing	**IT Audit and Controls**	**COMP3016 Hypertext and Web Technologies**	**CORE: WEBS3001 Dissertation**
Compulsory: Core:COMP2211 Software Engineering Group Project	**IS Innovation and New Technologies**	**WEBS3001 Dissertation**	**Compulsory: UOSM2008 Living and Working on the Web**
Compulsory: Core:COMP2212 Programming Language Concepts	**IT Security and Risk Management**	**UOSM2012 Interdisciplinary: Online Social Networks**	
Compulsory: COMP2213 Interaction Design			
Compulsory: COMP3200 Part III Individual Project			

4.6 Biglan's Classification of Academic Disciplines

The key proposed outcome of thesis is to better understand the scope of Web Science by examining its 'component subjects,' determining which constituent disciplines are taught and studied as part of Web Science. A notable example of existing work into classifying disciplines, is that of (Anthony Biglan 1973b). Biglan's work proposes the cataloguing of academic subjects according to their attributes. Biglan classified a range of academic disciplines, deriving a matrix framework which groups academic disciplines according to characteristics shown in Table 13. Although the study by (Anthony Biglan 1973b) was conducted for a range of academic disciplines, as opposed to examining the details of a single discipline, Web Science is by nature multidisciplinary. Therefore, the work conducted by (Anthony Biglan 1973a) and (Anthony Biglan 1973b) is relevant and interesting, and could provide a framework for classifying the sub disciplines or component subjects of Web Science.

Biglan suggest that the academics in a given area are the best source of knowledge about the characteristics of the given academic area. Biglan performed a multidimensional scaling analysis of disciplines, utilising the collaboration of academic staff from the University of Illinois, and a small unnamed art college located in the state of Washington. Biglan's study consisted of a manual survey, which involved academics from all of the disciplines involved in the study, manually sorting piles of cards, grouping together cards representing subjects that they considered to be related. Where the survey was conducted remotely, academics were asked to staple together slips of card representing topics which they considered to be related. An example of the results from the (Anthony Biglan 1973a) classification of disciplines is shown in Table 13. This is not a fully exhaustive range of academic disciplines, it includes only the examples of subjects taught in the institutions surveyed by (Anthony Biglan 1973a) and (Anthony Biglan 1973b) at the time of their study, however the study does provide an indication of the characteristics of most major sciences of the time. A notable omission is that of Computer Science, which would undoubtedly have been included were the study conducted during the present day, under the 'hard/non-life/applied' field.

Table 13 Biglan's Model of differences among academic disciplines (Anthony Biglan 1973a)

	Non-Life		Life	
	Pure	**Applied**	**Pure**	**Applied**
Hard	Physics Chemistry Mathematics Engineering	Engineering (e.g., Chemical, Mechanical, Civil, Nuclear)	Biology Zoology Bacteriology Philosophy Life	Medicine Agriculture Physical Education
Soft	(e.g., Chemical, Mechanical, Civil, Nuclear)	Library Science Finance Operations Research Foreign Languages Law Architecture	Psychology Sociology Anthropology Literature	Education Organizational Behaviour Social Work Performing Arts

As seen in Table 13, (Anthony Biglan 1973a) and (Anthony Biglan 1973b) propose three types of categories for academic disciplines, these include hard vs soft, pure vs applied and life vs non-life, (Shin & Arimoto 2014) (Anthony Biglan 1973a) (Anthony Biglan 1973b). These three categories allow for eight possible types of classifications for academic disciplines, (Alise 2008) these include:

- hard/life/pure,
- hard/non-life/pure,
- soft/life/pure,
- soft/non-life/pure,
- hard/life/applied,
- hard/non-life/applied,
- soft/life/applied,
- soft/non-life/applied

The following Table 14 explains the characteristics of each of these types of academic disciplines.

Table 14 Description of classifications used by (Anthony Biglan 1973a)

Description of classifications used by Biglan
Pure: *Pure sciences are "primarily theoretical" such as for example, 'Mathematics'.* (Krishnan 2009) *"a science depending on deductions from demonstrated truths, such as mathematics or logic, or studied without regard to practical applications." (Google)*
Applied*: "Applied science consists in a system of concrete interpretations of scientific propositions directed to some end useful for human life." "Applied science, in which are included all applications of the experimental pure sciences. These are concerned with the improvement of human means and ends and with nothing else."* (Feibleman 1961) *"Applied science is a discipline of science that applies existing scientific knowledge to develop more practical applications, like technology or inventions." (Wikipedia)*
Pure vs Applied: *"pure science has as its aim the understanding of nature; it seeks application. Applied science has as its aim the control of nature; it has the task of employing the findings of pure science to get practical things done"* (Feibleman 1961) (McGrath 1978)
Hard vs Soft: *Every discipline has its paradigm and may be characterized as being well-developed, or in a state or pre-development. The social sciences or humanities may be pre-paradigmatic, whereas physics and engineering have well-developed paradigms. According to Biglan, a high paradigm or well-developed discipline would be "hard" and a low paradigm discipline would be "soft".* (McGrath 1978) ***"****Hard science and soft science are colloquial terms used to compare scientific fields on the basis of perceived methodological rigor, exactitude, and objectivity. Roughly speaking, the natural sciences are considered "hard", whereas the social sciences are usually described as "soft"." (Wikipedia)*
Life vs Non-life: *"The distinctiveness of life disciplines is that their subject matter refers to any type of living thing; therefore botany and zoology as well as anthropology are life sciences. Non-life fields are those whose subject matter deals with anything non-organic, for example geology."* (Alise 2008)

The classifications outlined by (Anthony Biglan 1973a) and (Anthony Biglan 1973b) are logical and largely self-explanatory. Pure disciplines are those that involve purely theoretical deductions such as mathematics. Applied disciplines represent those which have a practical application, for example engineering, which has a physical usable output. Hard science and soft science refer to methodological exactitude. So-called hard sciences employ rigorous methods with definitive results. Soft sciences have less clear and definitive methodologies and results, for example, there may be no clear 'right' answer to a given research question, instead there may be multiple answers. Natural sciences are general perceived as "hard" disciplines, whereas social sciences are usually classified as "soft"." Life sciences refer to disciplines which study living things, whereas non-life sciences involve the study of anything that is not living, such as geology.

4.6.1 Application for Web Science?

(Anthony Biglan 1973b) also suggests that academics of each given subjects are the best source of knowledge about each discipline. To this end, Web Science academics and students will be a key source of information for this study into the nature of Web Science. Web Science academics and students will be surveyed and asked to rate the relevance of constituent subjects of Web Science. Whereas at the time of their study, (Anthony Biglan 1973b) conducted a manual paper study, this study of Web Science will be achieved with an online survey instrument. Similar to the study conducted by (White et al. 2011), academics and students will be asked to rate the relevance of the subjects of the Web Science subject categorization in order to determine subjects' perceived relevance according to the expertise of participants.

When outlining their proposals for Web Science, (Berners-Lee, Weitzner, et al. 2006) state that Web Science is socio-technical. (Halford et al. 2010) also suggest the concept of heterogeneous actors, in which the Web is formed by social and technical elements, incorporating both society and technology. Therefore, Web Science is by nature a combination of both hard and soft subjects. Following data gathering and the identification of constituent subjects of Web Science, it should be possible to examine the key component subjects of Web Science. It will then be beneficial to group the subjects according to the (Anthony Biglan 1973b) classifications, in order to represent and categorise the spread and types of sub-disciplines of Web Science.

4.7 Outline of Science Wikipedia

The Outline of Science was highlighted as a useful resource, and an example of a classification matrix for subjects by Rafael Melgarejo Heredia (Heredia & Vinueza 2015). Rafael utilised the topics of the Outline of Science in his example of a proposed Matrix structure for the Web Science Lexicon. See section 9.2 for further details. The Outline of Science[26] is a resource hosted on Wikipedia which is an example of a classification of scientific disciplines.

4.7.1 The Six Main Categories of the Outline of Science

There are six categories listed within Wikipedia's Outline of Science. These include:

- Natural Science
- Formal Science
- Social Science
- Applied Science
- Interdisciplinary
- Philosophy and History

The final category is a combination of two sub categories, which includes the philosophy of science and the history of science. These do not refer to philosophy or history in general, these only represent the philosophy and history of the study of science.

[26] https://en.wikipedia.org/wiki/Outline_of_science

Table 15 Wikipedia's Outline of Science Categories

Outline of Science Category Descriptions
Natural Science **Google:** "Branch of knowledge which deals with the study of the physical world." **Wikipedia**: "Natural science is a branch of science concerned with the description, prediction, and understanding of natural phenomena, based on observational and empirical evidence."
Formal Science **Google:** "A "formal science" is an area of study that uses formal systems to generate knowledge such as in Mathematics and Computer Science. Formal sciences are important subjects because all of quantitative science depends on them." **Wikipedia**: "Formal sciences are language disciplines concerned with formal systems"
Social Science **Google: "**the scientific study of human society and social relationships." **Wikipedia**: "Social science is a major category of academic disciplines, concerned with society and the relationships among individuals within a society. It in turn has many branches, each of which is considered a "social science"."
Applied Science **Wikipedia:** "Applied science is a discipline of science that applies existing scientific knowledge to develop more practical applications, like technology or inventions."
Interdisciplinary **Wikipedia:** "Interdisciplinarity involves the combining of two or more academic disciplines into one activity (e.g., a research project). It is about creating something new by thinking across boundaries."
Philosophy and History **Wikipedia:** "Philosophy of science is a sub-field of philosophy concerned with the foundations, methods, and implications of science." **Wikipedia:** "The history of science is the study of the development of science and scientific knowledge, including both the natural sciences and social sciences."

4.7.2 Outline of Science vs Biglan's Classification of Academic Diplines

The Outline of Science differs from Biglan's Classification, in the sense what while Biglan's model classifies all academic disciplines, the Outline of Science classifies only subjects which are considered sciences. While not directly matching Biglan's classifications, the Outline of Science categories do at least loosely correspond with Biglan's categories. The following list shows the overlap between Biglan and the Outline of Science where an overlap is evident:

- Natural Science = Life + Prue
- Formal Science = Hard + Non-Life
- Social Science = Soft + Life
- Applied Science = Hard + Applied

It is problematic to apply all categories of Biglan to the Outline of Science categories. For example, Social Science is a broad term which can include both pure and applied social sciences. For example, it is possible to have both pure and applied sociology. Pure Sociology focuses primarily on the acquisition of knowledge about a given subject in an attempt to better understand it. In contrast, applied sociology focuses on the investigation into social issues in an attempt to provide a practical solution to solve them [27]. Wikipedia describe the fact that there are three major branches of science, natural, formal and social. They explain that these constitute the 'fundamental sciences' and that

[27] http://www.differencebetween.com/difference-between-pure-and-vs-applied-sociology/

these then "*form the basis of interdisciplinary and applied sciences such as engineering and medicine.*" [28]

4.8 Conclusions

This chapter has featured background research, exploring the evolution of disciplines, as well as the overlap between Web Science and related disciplines, Internet Science and Network Science. Examples of formal curriculum guidelines for similar subjects Computer Science and Information Science were evaluated. This included a comparison between Web Science modules, and Computer Science and Information Science modules, in order to identify and overlaps between the subjects. The comparison showed limited overlap between the subjects, reinforcing the argument that Web Science is and should be recognised as a subject in its own right, separate from that of traditional Computer Science (Hendler et al. 2008).

The classification of subjects according to (Anthony Biglan 1973a) and (Anthony Biglan 1973b) and the Outline of Science[29] were also examined, as these are existing resources which feature the classification of academic subjects. These two resources will provide a valuable reference when producing the structure of the Lexicon during chapter 9.

[28] https://en.wikipedia.org/wiki/Branches_of_science
[29] https://en.wikipedia.org/wiki/Outline_of_science

5 Research Questions

5.1 Introduction

This chapter outlines the research questions that this thesis will address, as well as providing a brief overview of the research methods by which the author intends to address them. A detailed overview of the research approach utilised is provided in Chapter 6, and the Data sources utilised are outlined in greater detail in Chapter 7.

The Research Questions are as follows:

- What is the extent of the Web Science field of study?
    - Where is Web Science Taught?
    - What is taught as Web Science?
    - What subjects are studied by Web Scientists in practice?

The ultimate aim is to identify a clearer definition for Web Science, employing a 'bottom up' approach, in order to provide academics and students wishing to teach or study the subject with a resource outlining the scope of the subject. The following section looks at each of the research questions in more detail. Table 16 outlines each research question, as well as a brief overview of the methods proposed for addressing each question. The following paragraphs then explain each question in more detail.

5.2 The Hierarchy of Research Questions: Explained

The key objective of this study is to answer the question 'What is the extent of Web Science?' This will be achieved by answering the questions 'Where is Web Science taught?', 'What is taught as Web Science? and 'What subjects are studied by Web Scientists in practice?'. The answers to the three 'sub-questions' will provide an answer to the higher-level question: 'What is the extent of the Web Science subject?' The following table and section includes a breakdown of each research question, explaining each in more detail.

Table 16. Breakdown of Research Questions

Research questions	Research instrument\data source	Data type\ Outcome	Notes
1. What is the extent of the Web Science field of study?			
Determine an outline of the current scope of the Web Science subject, identifying its key constituent topics. Present a representation of the current scope of Web Science as an ontology of key constituent topics. (This will be named the Web Science Lexicon) The purpose of this ontology would be to communicate the scope of Web Science for student and academics wishing to study or teach Web Science.	Desk survey, analysis of Web Science conference papers, Background reading\research, survey, interviews	Most frequently occurring subjects taught and published as part of Web Science, contributing to a Taxonomy of topics representing the scope of Web Science	Web Science is NOT static, a fully comprehensive and definitive representation is not realistic. However, a representation of the subject in its current state is possible, and would be beneficial to people wishing to gain an understanding of the scope of the subject.
2. Where is Web Science taught?			
Which institutions across the world teach Web Science, and where are they located?	desk survey, survey,	Survey results\data – will require processing using a packages such as NVivo	Initial desk survey will identify Web Science teaching institutions worldwide. Will information on which institutions to target for a more detailed modules survey.
3. *What* is taught as Web Science?			
What is the taught definition of Web Science, and Is there a difference between existing subject definitions, (such as the Web Science Subject Categorisation) and what is taught as Web Science in practice?	Background reading\ research, improved version of the teaching institution desk survey, survey, interviews	Identification of what is taught, and comparison between this and the subject definition.	Will need to perform a more detailed desk survey of Web Science teaching institutions, in order to facilitate the analysis of module content.
4. What subjects are studied by Web Scientists in practice?			
• Investigate which interdisciplinary subjects are researched and studied by Web Scientists in practice. • What is being published by Web Scientists as part of Web Science Research? • How do Web Scientists perceive and describe the Web Science subject?	Background reading\ research, survey, interviews\ study of existing Web Science conference literature	Opinions of data subjects	What subjects are Web Scientists actually studying? What do Web Scientists themselves say about the scope of the subject? What is the extent of students' understanding of Web Science having completed a Web Science programme?

Table 16 shows a breakdown of the research questions. The highest level question is 'What is the extent of the Web Science subject'. All three other 'sub questions' relate to addressing this key question, each exploring it from different aspects.

5.2.1 What is the extent of the Web Science Subject?

This highest-level question, relates to gaining an overview of what constitutes Web Science as a subject, with the aim of better representing and communicating its scope and benefits to both academics, students, as well as the interested intelligent layperson. This question and the following question are interrelated. Gaining an understanding of what is taught as part of Web Science will provide an indication of the answer to 'what is the extent of the Web Science subject?' While determining what is taught as part of Web Science programmes is the main benchmark that will be used to determine the current nature of Web Science, there are additional indicators of the subject content which will be used as a point of comparison. For example, topics covered by Web Science related events such as the Web Science conference will provide an indication of the scope of the wider subject.

5.2.2 Where is Web Science taught?

In order to identify what is being taught as Web Science around the world, it is necessary to first identify *where* it is being taught. In order to address this question, a desk survey will be conducted in order to identify institutions worldwide that teach Web Science. A survey of Web Science students and academics will also seek to identify additional institutions that offer Web Science taught programmes.

5.2.3 What is taught as Web Science?

This question focuses upon examining what is taught in practice as Web Science across institutions worldwide; with a view to identifying the Web Science taught topics common to most programmes. The main data source for achieving this will feature a more detailed version of the desk survey of Web Science teaching institutions. This second desk survey will focus upon the identification of specific modules which are taught as part of Web Science programmes. This will provide module data which can then be analysed, in order to determine which topics are most frequently taught as part of Web Science programmes.

5.3 What subjects are studied by Web Scientists in practice?

As suggested by (Anthony Biglan 1973b), academics in any given subject are the best source of knowledge about the subject. While it is possible to gain a good understanding of the subjects within a Web Science taught programme by examining curriculum specifications, this does not provide insight into the subjects which are studied as part of Web Science in practice. This research question will be addressed with a study of Web Science literature published at the Web Science conference series, which has run since 2009. This will help to determine what research is being conducted by Web Scientists in practice. A series of interviews will also be conducted, targeting Southampton Web Science researchers; the purpose of which will involve gaining insight into what Web Science PhD students are researching as part of Web Science in practice. It will also be beneficial to ascertain how Web Science researchers describe the subject, 'Web Science', as they are at the forefront of the discipline.

5.4 Conclusions

The research questions outlined in this chapter stem from the need to provide a clearer definition of the emerging subject, Web Science. A hierarchy of four research questions have been outlined, with the three lower level questions: 'What is taught as Web Science', 'Where is Web Science taught?' and 'What subjects are studied by Web Scientists in practice?' address the answer to the higher level question: 'What is the extent of the Web Science Subject?'. The research methodology proposed for achieving this outcome is explained in Chapter 6. The results of the answer to the main question; 'What is the extent of the Web Science Subject?' will be presented in the form of a structured taxonomy of topics representing the scope of Web Science. This will be termed the 'Web Science Lexicon', and will be the major contribution of this project.

6 Research Approach and Tools Employed

6.1 Introduction

The following chapter outlines the research methodology adopted for use during this thesis. This includes a review of the research approach adopted, details of the specific research methodology employed, as well as the practical research tools utilised in order to conduct research.

The structure of this chapter is as follows:

- Qualitative vs Quantitative
- Justification for the Research Approaches Used
- Adopting a Qualitative, Mixed Method Approach to Web Science
- Utilising a Grounded Theory Approach Incorporating Mixed Methods
    - Grounded Theory
    - Thematic Analysis
    - Keyword Analysis: Bibliometric Mapping
    - Ethnographic Study
- Data Analysis Tools
    - Mendeley
    - Nvivo
    - Evernote

6.2 Web Science: A Bottom up Approach

A key problem when defining curriculum guidelines for Web Science, is the fact that Web Science is a rapidly evolving field. Use of traditional top down methods of formulating curriculum guidelines, such as those used by the ACM for Computer Science, involve an onerous and time consuming process of development. A panel of experts will typically be convened and produce a set of guidelines over a period of time, this will then be put forward for consultation by the wider community, before being reviewed and formalised. The guidelines will then be published and will become a set of standards, which are used as a benchmark in the development of taught curricula.

Any similar attempt for Web Science using the top down method would be outdated by the time it were completed. To this end, as outlined in Chapter 3, Web Science is suited to a bottom up research approach, in which the content of existing Web Science curricula and literature will be used to develop a taxonomy of topics representing Web Science, in a reverse engineering fashion.

When using a traditional quantitative research approach, the researcher will typically begin with preconceived ideas or theories, and then set out to prove them. However, when utilizing a qualitative approach, the outcome is influenced by the findings of the research, and the direction of the project is influenced by findings, which is conducive to the proposed bottom up approach. The following section will look at the differences between quantitative and qualitative approaches, and will discuss the application for Web Science.

6.3 Background: Qualitative vs Quantitative

> *"Qualitative methodology implies and emphasises on discovery and description, and the objectives are generally focused upon the meaning of experience.*
>
> *These objectives are contrasted with those of quantitative research, where the testing of hypotheses to establish facts and to designate and distinguish relationships between variables is usually the intent.*
>
> *The intent of qualitative research is to examine a social situation or interaction by allowing the researcher to enter the world of others and attempt to achieve a holistic rather than a reductionist understanding".*

(Bloomberg & Volpe 2012)

(Strauss & Corbin 1998) state that the term 'qualitative research' covers any type of research that is not conducted using statistical methods "*or other means of quantification*". This type of research focuses upon discovering details of peoples' life experiences, feelings and insights about a given situation (Flick 2009). As explained by (Bryman 2001) the qualitative researcher maintains an open approach to their research, enabling them to identify and pursue unforeseen avenues of research that may not have been foreseen when they initially planned the study, which would not be possible within a structured quantitative approach. It is arguable, however, that the quantitative method allows the researcher to remain unbiased, whereas by utilising the qualitative method, the researcher and the subject are far more interactive, and therefore the opinions of the researcher may influence the outcome to an extent.

6.3.1 Qualitative vs Quantitative: The Application for Web Science

The quantitative approach enables the researcher to reliably and rigorously gather large amounts of data which is suited then to the application of statistical analysis (Berg 2006). "*Quantitative studies emphasize the measurement and analysis of casual relationships between variables, not processes*" (Denzin & Lincoln 2011). Traditional methods of studying technology involve theorizing how a given technology is used in certain situation. (Orlikowski 2008) propose the alternative method of observation of the practice of how technology is actually used by stakeholders, using observation of practice to inform studies. This method of using a 'practice lens' can be applied to the study of Web Science, by examining the reality of what is actually taught as part of real Web Science taught curricula. Traditional quantitative methods for defining a curriculum are time-consuming, large data samples are often required for accuracy, and due to the rapid evolution of Web Science, the results are likely to become outdated by the time such efforts are completed. Therefore, (White et al, 2011) propose a bottom up approach to defining Web Science, which is more suited to a flexible qualitative approach, as opposed to a traditional heavily structured quantitative example. As suggested by (Tinati et al. 2012), this will be augmented by incorporating some mixed methods, employing some quantitative analysis techniques as a means of comparison. The following section will outline the qualitative research approaches used during this work.

According to (Strauss & Corbin 1998), a reason for choosing a qualitative research approach, is determined by the nature of the research problem. "*For example, research that attempts to understand the nature of experience of persons*" – in this case people's experience of the Web. The Qualitative research approach is particularly applicable to Web Science, because the data available potentially requires both positivist and interpretative methods of data gathering, in order to successfully interpret it. For example, raw data about Web usage - such as stats showing the number of people to visit a specific Website, or the number of people enrolled on a Web Science degree programme could be ascertained using positivist methods, and can provide valuable data. However, in order to also learn what people actually thought about the Website or the course that they were enrolled on, this would then require an interpretative approach, in order to gain user feedback. Depending upon this feedback, the direction of the research may alter; for example, if users said that the Website was not useful, but that there was another similar Website that they visited which was more useful, this could be ascertained using an interpretive approach, while the positivist approach would simply determine how many people used the original website - perhaps at the most including a rating scale of its effectiveness rather than peoples' actual opinions of it. If we are to judge what users think of the Web and how they actually use it; then it is arguable that interpretive methods play an important part, providing valuable insight which could not be gained from stats alone.

(Orlikowski 2008) propose a "*Practice Lens for Studying Technology", by this they* refer to gaining an understanding the use of technology in organisations by practical observation of how the technology is actually used in practice, as opposed to forming theories about its use. "*Actors shape the world at the same time that it shapes them*" (Chomic 1988). This can be applied to the Web Science. By using a bottom up approach to observe what is taught and published as part of Web in practice, and comparing this with what is taught as part of Web Science literature, it should be possible to build a picture of what constitutes Web Science in practice. This will arguably provide a more detailed, first-hand insight into the subject, than the theoretical top-down method employed by traditional approaches to defining subjects.

6.4 Utilising a Grounded Theory Approach Incorporating Mixed Methods

The following section outlines the research approaches adopted throughout this project. Qualitative research includes a mixture of both positivist and interpretivist methods; which is necessary in order to build a full picture of a subject. In their paper, (Lin 1998) argue that an approach featuring entirely one or other method will produce flawed results, and that a combination of both methods is the most sensible option. In their paper on 'Mixing Methods and Theory to Explore Web Activity' (Tinati et al. 2012) suggest the benefits of adopting as mixed methods approach for Web Science as a social technical subject, explaining that *"both approaches examine the same phenomena from a different perspective, and by combining them we can grasp the Web at micro and macro levels".* They explain that a mixed methods approach is ideal for understanding both the technological and the sociological perspectives of the Web. Therefore, the author proposes to adopt qualitative Grounded Theory based approach which incorporates mixed methods, in order to build a full, well-rounded picture of Web Science. The different aspects of this qualitative, mixed methods based approach are outlined in greater detail in the following sections.

6.4.1 Grounded Theory

A Grounded theory is one of the most widely recognised approaches to conducting qualitative research. (Bryant, A, & Charmaz 2007) explain that the Grounded Theory approach is a "*a systematic, inductive, and comparative approach for conducting enquiry"* (Bryant, A, & Charmaz 2007) and constructing theories, and "focuses on building theory from data". (Braun & Clarke 2013).

This approach is particularly pertinent for Web Science, as the research is informed by the constant evolution of the Web. It is also suggested by (Bryant, A, & Charmaz 2007) that memo writing is a key method of the researcher gaining an understanding their subject, and that the process of memo writing evolves throughout the project and enables the researcher to see emergent patterns in the data, enabling them to form conclusions, reflecting the 'bottom up' approach as explained by (White et al., 2011). According to (Bryant, A, & Charmaz 2007) and (Glaser & Strauss 1967), although Grounded Theory is largely described as a qualitative approach to research, (Glaser & Strauss 1967) state that it may also be applied to quantitative research. This means that it is an ideal mixed method approach to use in order to evaluate both qualitative and qualitative data, something which will be valuable during this project.

The proposed Bottom Up approach is suited to qualitative Grounded Theory, incorporating mixed methods, as outlined by Strauss and Corbin. According to (Strauss & Corbin 1998), theory derived from data is far more likely to resemble reality than a theory drawn only from a series of concepts or speculation. Because Grounded Theory involves deriving theories from the data, it is more likely to offer accurate insight, (Bryant, A, & Charmaz 2007). The Grounded theory approach suggested by (Strauss & Corbin 1998) is not purely qualitative, it has the benefit of integrating mixed methods. While this study includes largely qualitative data it also includes some qualitative analysis of the data, including keyword analysis, and also quantitative presentation of qualitative data, facilitated by Nvivo. (Strauss & Corbin 1998) suggest that there is no standard set of methods equally useful for each research step; therefore researchers will typically use a combination of methods, in order to "foster the development of theory". They also propose that there can be interplay between the qualitative and quantitative in an iterative cycle, in which each method contributes and provides insight in ways only it can.

6.4.2 The Straussian Method of Coding (Thematic Analysis)

Thematic analysis involves "identifying and analysing patterns in qualitative data" (Clarke & Braun 2013). Thematic analysis is a term for coding data. (Strauss & Corbin 1998) outline a set of procedures for coding, which are designed to help the researcher foster the development of theory. (Strauss & Corbin 1998) explain that these guidelines need not be adhered to rigidly, they are there as a framework to aid the researcher, and should assist and not hinder the analytic process. Their suggested procedures for coding are as follows in Table 17.

Table 17. Straussian Method of Grounded Theory Coding

Straussian Method of Grounded Theory Coding according to (Strauss & Corbin 1998)
Open Coding The analytic process through which concepts are identified and their properties and dimensions are discovered in the data.
Axial Coding The process of relating categories to their subcategories, termed 'axial' because coding occurs around the axis of a category, linking categories at the level of properties and dimensions.
Selective Coding The process of integrating and refining the theory.

The (Strauss & Corbin 1998) coding methodology as outlined in Table 17 translate to the following proposed steps for practical use within this thesis. As suggested by (Strauss & Corbin 1998), these stages can (and will be) an iterative process.

6.4.2.1 Open Coding

The process of identifying a series of 'key topics'; topics which are part of the Web Science curriculum. This will be achieved by analysing the contents of a survey of Web Science taught modules, in order to identify what concepts or 'key topics' are most commonly taught as part of Web Science curricula.

6.4.2.2 Axial Coding

The process of relating higher level Web Science 'key topics' (categories) to lower level key topics. For example, when a higher level 'key topic' might be 'Web Society'; this is a broad category under which other lower level key topics (subcategories) need to be grouped for the purpose of creating a structured Web Science Taxonomy. Any key topics which may be too closely related will need to be identified and potentially combined.

6.4.2.3 Selective Coding

This process is the final stage of refining and grouping the 'key topics' (categories), in order to create the final taxonomy of Web Science key topics.

6.4.3 Conducting Thematic Analysis with a Software Package

It is possible to conduct thematic analysis manually, or using a software package. When the process of thematic analysis is referred to in this thesis, it essentially means that the data will be coded using the procedures outlined above by (Strauss & Corbin 1998).

The qualitative research software package Nvivo is also capable of facilitating coding or thematic analysis, utilising the principles for Thematic Analysis as outlined by (Strauss & Corbin 1998) and (Clarke & Braun 2013). The following Table 18 shows the stages of thematic analysis as outlined by (QSR 2010), a manual showing the processes involved when conducing thematic analysing utilising the qualitative data analysis package, Nvivo. These stages of analysis as described by (QSR 2010) are loosely mapped to the coding stages as described by (Strauss & Corbin 1998). This is an attempt to illustrate the stages of coding which loosely correspond to the stages of thematic analysis conducted using NVivo. However, the process of coding is iterative, so each stage may be completed more than once.

Table 18. Stages of Thematic Analyses as Defined by (QSR 2010) *and* (Strauss & Corbin 1998)

Stages of Thematic Analysis as Defined by (QSR 2010) and (Strauss & Corbin 1998)
Open Coding
• *"Phase 1: Data familiarisation,*
• *Phase 2: Initial coding,*
Axial Coding
• *Phase 3: Theme identification,*
• *Phase 4: Reviewing themes,*
Selective Coding
• *Phase 5: Defining and naming themes,*
• *Phase 6: Report writing"*

The practical application of the Nvivo coding processes will be outlined more detail in chapter 8, when the practical application of data analysis is described.

6.4.4 What is Keyword Analysis/Bibliometric Mapping?

(Hooper et al. 2015) explain that the theory bibliometric mapping suggests that a given subject may be defined by a list of important keywords.

"*In the quantitative study of science, which is also known as scientometrics or (if focused on literature data) biblio-metrics, 'maps of science' provide overviews of structures existing within science fields*." (Buter et al. 2004). While the process of bibliometric mapping is a mainly quantitative approach to research; (Buter et al. 2004) go on to explain that the identification of trends within the data identified using bibliometric mapping require a qualitative approach. This suggests that a Grounded Theory based thematic analysis approach in order to identify data trends may complement the quantitative process of bibliometric mapping, reinforcing the positive benefits of adopting a mixed method approach to research.

The (Hooper, Bordea, et al. 2013) and (Hooper et al. 2015) studies used the bibliometric mapping approach to compile a list of keywords in order to analyse the spread of topics publishes within Web Science literature. A similar approach and comparison with keywords gathered from Web Science taught curricula may provide interesting insights and a direct comparison into how what is being taught as part of Web Science curricula in practice compares with Web Science published material.

6.4.5 Thematic Analysis and Keyword Analysis: Application

A process of keywords analysis or bibliometric mapping will be utilised in combination with thematic analysis, in order to identify key themes in Web Science taught programme content identified from the desk survey of Web Science teaching institutions. The combined process of keyword analysis and thematic analysis will also be utilised to analyse the Web Science conference publications, in order to identify keywords and topics most commonly published for Web Science.

6.5 Software Tools Utilised

The following section of this chapter outlines the software tools utilised within this work in order to support the research and also to facilitate the hosting of the proposed Web Science Lexicon.

6.5.1 Data Analysis Tools

The following data analysis tools will be utilised to support a variety of tasks through the project, facilitating the implementation of the thematic analysis using a grounded theory approach.

6.5.1.1 Mendeley

Mendely is a reference management tool which is invaluable for organising resources and creating bibliographies. It has a Word plugin which facilitates the creation of reference lists. It also allows resources related to the project to be capture and stored in one accessible location. There is also a social aspect to the tool; it is possible to create groups within Mendeley which allow resources to be shared between researchers. It is also possible to export resources from Mendeley to other research packages, such as Nvivo.

6.5.1.2 Nvivo

Nvivo is a qualitative data analysis software package. It has been designed for qualitative researchers working with very rich text-based and/or multimedia information, where deep levels of analysis on small or large volumes of data are required. There are a considerable number of tools offered by Nvivo, which enable the user to perform a thematic analysis of the data. Nvivo is an invaluable tool used in order to organise project resources. It also allows keyword analyses to be run easily, a feature which will be very valuable for collating key terms relating to Web Science. It has been used in combination with Mendeley to gather resources related to this project. While NVivo's main focus is on qualitative analysis, it does provide support for researchers who work with mixed methods. For example, it is possible to analyse the open-ended questions in a survey. This is an aspect which will be invaluable and will be explore further in order to analyse the results of the questionnaire based survey once reasonable number of results have been gathered. As well as simply being a useful means of gathering resources, this tool enables what is termed as 'coding', which enables users to assign tags to sections of text and images, which then enables queries for tagged topics. This is an invaluable feature which allows comparison of different opinions on a topic in several publications.

NVivo realises this ability to 'code' data according to categories, and perform thematic analysis, identifying themes and patterns within the data. Nvivo is also capable of performing word frequency and keyword analyses required for simple bibliometric mapping. This feature will be utilised in order to identify keywords and themes in the desk survey of Web Science teaching intuitions, the Survey of Web Science students and academics and the material published in the Web Science conference series.

6.5.1.3 Evernote

Evernote is an online notes tool, which is used for recording memos, thoughts and observations about project content, as well as minutes from meetings. A significant benefit to using this tool, is the fact that it is online, and therefore can be accessed from any device, including mobile, ipad and multiple desktop PCs. Another key feature of this software, is the ability to tag notes with related keywords, as well as a web capture tool available on the desktop version, which can be used to capture entire webpages. This has proved useful in capturing Web Science related web content into a format which can potentially be imported into other packages such as NVivo for keyword analysis. Evernote is the main tool which was used to record memos according the Grounded Theory methodology.

6.5.2 Platforms considered for creating the Web Science Lexicon

The following section details the investigation of software platforms which could facilitate the creation of the proposed Web Science Lexicon.

6.5.2.1 Investigation of Wiki's as a Platform for the Web Science Lexicon

A series of open source Wiki platforms were investigated as a platform for the proposed Web Science Lexicon. The reason that wikis were considered as an option is because wikis allow the creation of a series of linked pages or articles which facilitate editing and use by multiple users. This would facilitate the ongoing contribution of content from the wider Web Science community. Investigation revealed that there are two general varieties of wikis, which include: personal wikis, which are aimed at smaller scale usage by individual users, and large-scale wikis which require hosting on a server. The wikis considered for the Lexicon are included in Table 19.

6.5.2.2 MediaWiki

Media Wiki is a powerful and well established large scale free open source wiki platform. It is the platform which was utilised to create Wikipedia, and includes the facility to create wiki pages in this well recognised format. It is a highly customisable and flexible platform, and would offer the option to create a series of referenceable pages for the Web Science Lexicon, which would also be editable by multiple users. It is also customisable for users without extensive knowledge of HTML. Although free, Media Wiki requires a large-scale server in order to host.

6.5.2.3 DokuWiki

Another example of an open source wiki which is suitable for use within an enterprise context. It is described as 'highly versatile', including additional features beyond that of a standard wiki. Doku Wiki allows the content to be 'categorized in namespaces' and browsed via an 'automatic index', which would be ideal for allowing navigation of the Web Science Lexicon entries. DokuWiki has the benefit of not requiring a dedicated database, however, similar to Media Wiki, it still requires a server in order to host.

6.5.2.4 TiddleyWiki

TiddlyWiki is a free open source wiki platform, which describes itself as providing a unique '*non-linear notebook for capturing, organising and sharing complex information*'. The platform is designed to be tailored according to users' needs, and can facilitate the production of personal databases, collections and to-do lists and more. It also includes additional plugins which allow the production of mind maps and visualisations. The TiddlyWiki platform has the benefit of being a well-established application, it's initial release dates back to 2004 and is one of the more well-established personal wikis available. However, this platform is not ideal for editing by multiple users, which could be problematic when encouraging participation by the wider Web Science community. TiddlyWiki does not offer a hosting service, however, a benefit is the fact that it is smaller scale than some other wiki platforms, and does not require a dedicated server for installation. Users can create content locally and then export for Web use, therefore content created using TiddlyWiki could potentially be hosted on a small scale low-cost web hosting service.

Table 19 Wikis Considered

Free Open Source Wiki Platforms	
Media Wiki	**https://www.mediawiki.org/wiki/Manual:What_is_MediaWiki%3F**
Doku Wiki	**https://www.dokuwiki.org/dokuwiki**
Tiddly Wiki	**https://tiddlywiki.com/**

6.5.2.5 Conclusions Regarding the use of a Wiki

The majority of personal usage wikis are most commonly personal organiser or notepad type applications. Larger scale free open source wikis generally do not offer hosting services, and are designed to be hosted on the user's own server. Prolonged use of a server for hosting is not something which is available within the scope of this project. None of the wikis considered are ideal for the creation of the Web Science Lexicon, as none provide inclusive hosting, or in the case of TiddlyWiki, are not readily accessible for multiuser contributions. Therefore alternatives will need to be considered.

6.5.2.6 Drupal Platform – a Brief Overview

During the WAIS Fest event brainstorming group (see section 9.2 for a description and account of the event) it was proposed by Chris Gutteridge, member of the University of Southampton's IT Innovation team, that the Web Science Lexicon could be hosted as part of the University of Southampton's Website, which could increase the profile of the Lexicon to a wider academic audience.

Because of the limitations of the free wikis investigated, it was decided to employ the Drupal platform utilised by the University of Southampton web team. The Drupal platform is the content management platform which allows non-web developers to author University of Southampton web content. The Drupal content management platform allows users' to author multiple pages organised under higher level headings, which will be utilised to create the Lexicon structure. The Drupal platform requires a user login in order to make contributions. However, participation is possible via a login, and the necessary use of a login would prevent spam content from being added to the Lexicon. Despite the Drupal platform having some limitations, it is a more effective solution than a free wiki platform, especially given that hosting the wiki would be an issue. A crucial benefit of hosting the Lexicon as part of an official University of Southampton Website, is that this provides an academic platform with higher search engine rankings, as it associated with the web address of an academic institution. It could also make the Lexicon more visible to potential users within the University who may benefit from the content, or indeed people who may be able to contribute content.

6.6 Conclusions

As was explained during Chapter 3, the field of Web Science is rapidly evolving. One of the solutions to studying Web Science is therefore to use a 'bottom up' approach, looking at the material available in order to derive a curriculum from existing content. This is particularly suited to a qualitative Grounded Theory based research methodology utilising some mixed methods, as proposed by (Strauss & Corbin 1998), in which a theory is derived directly from the data. A process of Thematic Analysis will be utilised in order to identify themes and patterns in data. This will involve the identification of 'key topics', which represent themes coded from Web Science taught modules, and 'paper topics', which represent themes coded in the analysis of Web Science conference publications. This will be achieved by utilising the qualitative data analysis package Nvivo, as a utility for coding the data and identifying themes.

While the process of analysing sources is conducted qualitatively, Nvivo allows results of qualitative coding to be quantified for ease of presentation. Therefore, the resulting themes identified from the data during the qualitative analysis, will be quantified and presented in the form the key and paper topics most frequently coded. The frequency of each of key and paper topics will provide an indication of how prominent each theme or topic is within Web Science, and this information will be used to judge which final key and paper topics should be included within the Web Science Lexicon.

7 Data Sources

The following sections outline the data sources used to gather information relating to Web Science throughout this project. This will involve triangulation; utilising more than one data source in order to compare and contrast findings (Bryman 2011).

This chapter on data collection has the following structure:

- Data Collection Methods: The Desk Survey of Web Science Teaching Institutions
- Questionnaire Survey of Web Science Academics and Students
- .ac.uk Web Crawler
- Interviews of Web Science Students
- Analysis of Web Science literature

These are the main data sources utilised during this work. The following section outlines each of the methods listed above and explains why each is used. A full analysis of the findings of each these methods will be provided in chapter 8.

7.1 Data Collection Methods: The Desk Survey of Web Science Teaching Institutions

There are many features which can introduce bias into taught programmes, which can be considered. The first, is that new taught programmes can be subject to financial constraints, so they may be constructed by a combination of as many modules from existing programmes as possible, as opposed to beginning entirely new subject specific modules. This makes the cost of introducing a new programme more viable. In the case of an emerging subject such as Web Science, it is also inevitable that the founding academics do no originate from the subject itself, but from alternative disciplinary backgrounds. The disciplinary background from which academics originate will undoubtedly introduce subject related bias into the teaching of Web Science. For example, Tim Berners-lee, the creator of the Web itself and co-founder of the Web Science research agenda is himself from an Engineering and Computer Science background. The same was indeed true of the Computer Science discipline, which was founded by many with an original background in Mathematics. The research agenda set out for Web Science by (Berners-Lee, Weitzner, et al. 2006) encourages input from multiple disciplines, citing this interdisciplinarity as a key part of Web Science. These are some of the many biases which may impact the subject matter of Web Science programmes, however, the programmes remain a key source of data in an environment where little information exists in terms of a definition for Web Science.

A desk survey of Web Science teaching institutions was conducted in order to ascertain which institutions around the world currently teach Web Science. This was then followed by a more detailed desk study of modules. The initial desk survey of Web Science teaching institutions involved the compilation of a list of intuitions that teach Web Science programmes worldwide. This exercise followed the model of a previous study conducted by (White et al., 2011) which outlined a proposal for gathering information about the Web Science curriculum, and also conducted a brief study of Web Science educational institutions. As this study demonstrated, traditional surveys often yield low response rates from participants. Therefore for this study, it was decided to begin with a desk survey of university websites. This involved manual web searches in order to identify a list of institutions which teach an active Web Science syllabus, and also included institutions which teach a module or other content relating to Web Science. The full results of this desk survey may be found in Appendix 16.1. When searching for Web Science taught programmes, the search criteria for the identification of Web Science or related programmes, was that each programme identified should have a broad affiliation to the Web Science agenda. For example, programmes identified should be named as 'Web Science' programmes by title, or should match the same Socio-technical pattern of a 'typical' Web Science programme. So, for example, taking a typical Web Science programme as a benchmark, such as those offered by the Universities of Southampton or Koblenz which both feature a fairly even sociotechnical balance of modules.

7.1.1.1 Search Terms: Searching for Web Science Teaching Institutions

When searching institutions which teach Web Science, the list of Web Science intuitions found on the Web Science Trust Website was used as a starting point. Standard Google searches for the Term 'Web Science' were another means of locating several unknown programmes.

A frequent problem is the inability of Google to distinguish between the terms 'Web Science' and 'Web of Science'; two very different results which contain almost identical keywords. In an attempt to decrease the number of irrelevant 'Web of Science' results, a method investigated involved Boolean searching using the following terms in the standard Google search bar:

"Web Science" AND NOT "Web of Science"

as suggested by (Barker n.d.) and (Branscomb 2000). In reality, this produced the opposite result to the one desired – instead of 'Web Science' most of the results using this search were 'Web of Science', suggesting that this method of entering traditional Boolean term searches is ineffective when using the current version of Google search.

Instead, the advanced Google search features include a form of the above Boolean options, as there is an option to include 'none of these words'. Therefore the following combination was tested:

"Web Science" -Web -of –Science'

However, when this combination is used, however, Google cannot differentiate between the differing contexts for the use of the words 'web' and 'science' and simply excludes these words altogether. A far more effective search followed the exclusion of the word 'of 'only, resulting in the following search: '***"Web Science" –of'.***

This search produced a number of results indulging several Web Science research groups, including 'African Web Science' and '*Web Science - Informatik 5 - RWTH Aachen University*. This result was an outdated webpage for an old module run in 2008, but following some of the links on the page redirected the browser to a new site, and yielded a result for an up-to-date Web Science module.

Another method of searching which proved to be fruitful was the use of country or region names in Web Science searches. For example, searching for ***"Web Science Australia"*** yielded a result relating to a research group dedicated to Web Science based in Australia.

7.1.2 Initial Data Gathered from the Google Search

The initial desk survey of Web Science teaching institutions located a total of 22 universities worldwide which teach a Web Science or related programme. These are shown in Table 20.

Table 20. Web Science Teaching Institutions

Web Science Teaching Institutions
1. Aristotle University of Thessaloniki - Thessaloniki, Greece
2. "British university in Egypt - EL SHEROUK CITY, Cairo, Egypt"
3. Cologne University, Germany
4. Eindhoven University of Technology: Netherlands
5. Georgia Institute of Technology
6. Goldsmiths, University of London
7. Johannes Kepler University Linz
8. Korea Advanced Institute of Science and Technology (KAIST)
9. MIT - Massachusetts Institute of Technology
10. Northwestern University School of Communication
11. Oxford Internet Institute
12. Rensselaer Polytechnic Institute
13. RWTH Aachen University
14. Saint-Joseph University of Beirut
15. The University of Edinburgh: School of Social and Political Science
16. UAH MediaLab, University of Alcalá (Spain)
17. University College London
18. University of Erlangen-Nürnberg
19. University of Koblenz-Landau, Institute for Web Science and Technologies,
20. University of Liverpool
21. University of Southampton, UK
22. VU University Amsterdam; the Network Institute

Table 20 shows a list of the 22 institutions which teach Web Science or related subjects. These institutions originate in diverse world locations, however, the majority found originate in Europe. A map illustrating the location of these institutions is shown in appendix section 16.4.

7.1.2.1 Correlation Utilising Dedicated Masters Search Engines

An additional means of locating Web Science programmes, involves the utilisation of dedicated degree search engines which exist to help students search for degree programmes. The following Table 21 shows the top 15 results from a 'Findamasters.com' search. The search terms utilised when searching were 'Web + Science'.

Table 21 'Find a Masters' search results[30]

Course Title	University
Master in Web Science (M.Sc.)	University of Koblenz-Landu
Web Science – MSc	University of Southampton
Web Science and Big Data Analytics (MSc)	University College London (UCL)
Web Science and Big Data Analytics (MRes)	University College London (UCL)
Computer Science – MSc	University of Southampton
Web Technologies	University of the Highlands and Islands
Computer Science: Informatique MSc	Swansea University and University Grenoble Alpes
Soil Science MSc or MRes	Aberdeen University
Advanced Computer Science MSc	Swansea University
Computer Science MSc	Swansea University
Environmental Science MSc or MRes	Aberdeen University
Data Science – MSc	Goldsmiths London
Digital Marketing	University of Southampton
Data Science	Swansea University
Advanced Web Technologies	University of Manchester

The results returned by 'Find a Masters' provided limited usefulness when compared with manual Google searches for Web Science. Google searches can return any results and are not confined to University programmes, whereas the 'find a masters' results are purely degree programmes, when taking this into consideration, the results returned by the master search engine are a little disappointing for Web Science. Within the first page of search results (consisting of 15 entries), the top three results consisted of dedicated Web Science programmes. These three programmes offered by Koblenz, Southampton and UCL were previously identified by the manual search. The remaining twelve results were not specific Web Science programmes. Given that a total of 15 specifically MSc level Web Science programmes were identified in the manual study, this is somewhat less effective than anticipated.

Within the remaining twelve programmes which were not specifically Web Science; over half were either Computer Science, Web Technologies or Data Science programmes, which are all subjects which demonstrate at least some overlap with Web Science, but are not specifically Web Science. However, the search also returned some results which seem completely unrelated to Web Science, such as 'Soil Science' and 'Environmental Science'. Given the lack of relevance to Web Science demonstrated by these results, the overall accuracy of the search algorithm is somewhat questionable, especially given that these feature in the first page of results, and are therefore supposedly the most 'relevant'. Some of the programmes may demonstrate some overlap with Web Science, such the digital marketing MSc at the University of Southampton, whereas others appear to have no relevance to Web Science whatsoever, such as the 'Soil Science MSc' offered by Aberdeen University. The second page of results yielded no specifically 'Web Science' related programmes.

[30] https://www.findamasters.com/masters-degrees/?Keywords=web+science

Roughly half of the results on the second page were data science programmes. This may be due to the fact that these programmes refer to 'Web' and 'Science' frequently, but not in conjunction.

An alternative dedicated Masters search engine, is 'Mastersportal.com'. The following table shows the first page of results returned by the search engine 'Masters Portal' for the search terms 'Web + Science'.

Table 22 Search Results from 'Masters Portal'

Course Title	University
Web Science	Koblenz – Germany
Web Science	**Southampton – England**
Computer Science – Web Technology	**Kuala Lumper – Malaysia**
Web Science	**Koln – Germany**
Web Science	**Cairo, Egypt**
Advanced Computer Science-Advanced Web Technologies	**Manchester, England**
Computer Science – Internet and Web Technology	**Amsterdam, Netherlands**
Mobile and Web Computing	**Thessaloniki, Greece**
Web Technology	**Southampton, England**
Computer Science	**Southampton, England**

The results returned by Masters Portal were more directly relevant than those offered by 'Find a Masters'. Masters portal returned more specifically 'Web Science' related programmes. This search revealed the programme offered by Cairo, which seems less well publicised than some other programmes, therefore it would seem that this particular search engine is relatively effective. However, despite locating the Cairo, British University in Egypt programme, this even this search engine still failed to locate the total of 15 MSc programmes located by the manual search.

The programmes identified by Masters Portal which did not include the title 'Web Science', were alternatives such as Computer Science and Web Technologies. These are related subjects, but not specifically Web Science. There were no examples of completely unrelated programmes, such as 'Soil Science', which also suggests that this search engine is more effective than the 'Find a Masters' alternative. Despite 'Masters Portal' being the more effective of the two Masters search engines utilised, neither search engine returned over half of 15 the manual results found utilising manual Google searches. One possibility for a lack of results, is that such MSc search engines are more effective for well-established programmes such as Mathematics, than they are for emerging subjects such as Web Science, which are likely to be less frequently searched for. However, the search engines do serve to provide a benchmark for correlation against the manual search results.

7.1.3 Web Science Taught Modules Desk Survey

Having successfully compiled a list of institutions, the study to locate Web Science teaching institutions was then expanded to include details of individual modules relating to Web Science. This survey was recorded within a spreadsheet, and included the following fields:

Table 23: Key Fields Used in the Modules Desk Survey

Intuition Location	Resource Title	Topic(s) Taught	Teaching\ Assessment format
Contributors(s)	Materials Used (urls)	Level (e.g. Masters)	Module Dates\Duration

An abridged version of the full modules worksheet may be found in Appendix 16.1. The process involved time-consuming navigation of web pages for each of the institutions identified, in order to manually gather the information relating to the headings shown in 'Table 23: Key Fields Used in the Modules Desk Survey'. The data was then recorded in an Excel spreadsheet with the above cell headings. This process was repeated for each of the institutions identified in the previous stage of the study. It was only possible to gather data from institutions which provide public information relating to modules. An additional difficulty faced during the exercise, related to the fact that information is often formatted differently by each institution. For example, some universities provide detailed dates for module teaching times, whereas others only provide basic information such as semester 1 or semester 2, whilst others completely omit such information. Appendix 16.5 shows a Word frequency table showing top keyword results for modules. This is unfiltered in any way, and simply returns the top results for keywords taken from the 'topics' field of the modules worksheet. This gives an initial indication of which keywords occur the most throughout the taught modules. The modules desk survey has located **a total of 148 modules**. The full modules dataset can be found in Appendix section 16.1. The content of these modules will be analysed in Chapter 8.

7.1.4 Use of the Desk Survey Data

The data gathered by both desk surveys is the subject of further analysis. Chapter 6 explains the process of thematic analysis used to analyse the data gathered during the modules desk survey, in order to identify the key topics which are taught as part of Web Science programmes. A refined list of these 'key topics' will ultimately be used to aid the creation of the proposed taxonomy of Web Science key topics, as outlined in Chapter 2.2.1.

7.2 Data Collection Methods: The Web Crawler

This section examines the creation of an automated method of identifying additional Web Science teaching institutions, and Web Science related links. This option was explored In order to supplement the manual process of identifying Web Science teaching institutions.

The manual process of identifying Web Science teaching institutions is time-consuming, and while thorough, may not be exhaustive. Therefore, in accordance with the (Strauss & Corbin 1998) method of Grounded theory, which advocates the inclusion of mixed methods, a draft web crawler was coded. The purpose of the crawler, was to identify links relating to Web Science, and also to identify any Web Science teaching intuitions missed by the manual analysis.

7.2.1 Creating the Crawler

The author worked in collaboration with Dr Ash Smith, (iSolutions, University of Southampton), to create a simple web crawler using Python (Coskun et al. 2015), initially outputting a file in JSON format. Ash Smith coded the crawler, which searches the .ac.uk domain for a combination of the keywords 'web' and 'science', in an attempt to identify institutions which teach web science related content. The key information that is ultimately sought includes: Web Science Institutions, taught syllabi, resources and associated people.

The crawler currently returns the results in the format of a list of URLS, which can be traced to teaching institution. The intention was to generate a list of pages on the websites of educational institutions that contained the words 'Web Science' in some combination in order to discount irrelevant pages. The .ac.uk domain is enormous, and it is therefore impossible to carry out a complete crawl without a powerful server farm, but it was decided this would provide a reasonable starting point, with the intention of adding more intelligent search parameters later. It is already possible to rule out the sites not belonging to learning providers thanks to existing datasets. The system was designed to have a reasonably common yet stable back-end data store for two reasons. Firstly, it is possible to have multiple systems crawling in the foreground whilst pushing their results to a single central database, allowing parallel processing, but avoiding duplicated work. Secondly, it is possible can continue development on other, more efficient crawlers without having to stop the existing examples. A simple MySQL database was used for the back-end. For the second phase, more targeted searches were investigated. As well as simply crawling every page until a relevant result is discovered, it was decided to use specific searches from well-known search providers. Google no longer has a search API, however, its rival Bing does include that option, and allows queries to be restricted to a certain domain. The second crawler takes each known learning provider domain in turn and conducts a Bing search for "Web Science", adding each hit to the crawl queue. The free version of the Bing API allows 5,000 queries per month, which is ample as the crawler only searches once per domain.

7.2.2 Crawler Results

The crawler results mainly included Southampton and also identified the University of Liverpool as another institution which teaches web science. The Top 3 crawler results ranked in order of the number of URL results found were: Southampton, Liverpool and UCL.

Table 24 shows an exert of the top 25 results. The 'number of pages' field indicates how many pages the crawler has looked at for each URL. The 'number of mentions' equal the number of positive keyword matches it has found. Amongst the lower end results containing lesser number of positive mentions for the key terms, there are quite a few entries which are just variations of various .soton.ac.uk addresses. Perhaps unsurprisingly, a significant proportion of the total number of results pointed back to Southampton, which, from this search alone, would appear to be the main institution in the UK associated with Web Science. As the number of positive matches for the key terms decrease, the results become less useful, with results such as 'www.jobs.ac.uk'which simply leading a page for feedback on a PhD studentship application. Some of the results are displayed as numerous septate links; for example: ecs.soton.ac.uk is listed several times, even though it should potentially be included under the same URL heading.

Table 24 Exert of top 25 Crawler Results

URLs	No. of Mentions	No. of Pages
dtc.webscience.ecs.soton.ac.uk	405 mentions	22 pages
www.southampton.ac.uk	399 mentions	69 pages
wstweb1.ecs.soton.ac.uk	272 mentions	19 pages
acrg.soton.ac.uk	202 mentions	134 pages
users.ecs.soton.ac.uk	168 mentions	32 pages
moocs.southampton.ac.uk	123 mentions	34 pages
www.liv.ac.uk	116 mentions	103 pages
www.ecs.soton.ac.uk	108 mentions	47 pages
virtualopenday.southampton.ac.uk	101 mentions	40 pages
digitaleconomy.soton.ac.uk	90 mentions	23 pages
www.wais.ecs.soton.ac.uk	62 mentions	19 pages
id.southampton.ac.uk	51 mentions	105 pages
eprints.ecs.soton.ac.uk	43 mentions	15 pages
www.iam.ecs.soton.ac.uk	36 mentions	10 pages
digitalhumanities.soton.ac.uk	29 mentions	11 pages
www.edshare.soton.ac.uk	18 mentions	10 pages
blogs.exeter.ac.uk	16 mentions	21 pages
www.cs.ucl.ac.uk	16 mentions	16 pages
www.icc.ecs.soton.ac.uk	13 mentions	2 pages
sdp.oii.ox.ac.uk	12 mentions	12 pages
www.wun.ac.uk	12 mentions	6 pages
oro.open.ac.uk	11 mentions	22 pages
eprints.ulster.ac.uk	10 mentions	269 pages
www.psychometrics.cam.ac.uk	10 mentions	49 pages
admissions.ecs.soton.ac.uk	10 mentions	12 pages

One of the most successful and fruitful results produced by the crawler, (Table 24) was the University of Liverpool's connection to Web science. This was an institution missed completely by the previous manual study, and did not appear in any of the Google searches conducted at the time of the manual desk survey, despite running a Masters program labelled explicitly as 'Web Science'. This highlights the benefits of using mixed methods in order to detect the most possible results.

7.2.3 Conclusions: Crawler Limitations and Future Development Potential

While a useful supplement to the manual qualitative method of identifying Web Science teaching institutions, the crawler has some serious limitations. It is limited to the .ac.uk domain, consequently it can only detect universities within the United Kingdom. The crawler perhaps unsurprisingly failed to detect results from Goldsmiths and Edinburgh, which is likely to be due to the fact that these courses are not explicitly labelled as 'Web Science' although they relate to Web Science. This suggests that to improve a potential future a crawler, it would be desirable to create a more intelligent algorithm, which could identify derivatives and terms relating to of the term 'Web Science'. Another desirable further development would be to ultimately expand the scope of the crawler beyond the .ac.uk domain, to include institutions worldwide; however, this would require considerable computational resources which are beyond the scope of this work.

7.3 Data Collection Methods: Questionnaire Survey of Web Science Academics and Students

This section has the following structure:

- Survey Structure
  - Survey Section One
  - Survey Section Two
  - Survey Section Three
- Promoting the Survey: Difficulties Faced
- Survey Questions vs Research Questions
- Initial Results: Background of Participants

While it is possible to gain a great deal of information about a subject by analysing its curriculum specifications; deeper insight and understanding may be gained by speaking to people who actually studied the programme in practice. To this end, an online questionnaire based survey was formulated in order to gain insight into the skills and experience that Web Scientists gain from their curricula in practice. The ultimate aim is to compare people's views of their Web Science curricula, and what they perceive Web Science to be, with what is taught in practice. The information gathered in this survey should also provide a flavour of the backgrounds from which Web Scientists originate.

7.3.1 Survey Structure

For the full survey, please refer to Appendix 16.9 Questionnaire. The documentation for the survey ethics application can also be found in appendix section 16.8. The Survey is divided into three sections, and has the following structure:

1. About the Respondent
2. About the Respondent's Academic Institution
3. About Web Science: Rating WSSC Subjects

7.3.1.1 Survey Section One

The first section relates to discovering some basic information about the respondent. This includes basic contact details, such as their name and email if they are happy to provide it, and also questions which aim to identify their background in Web Science. For example are they a student, academic or other. (If they select other, they are asked to clarify, this option is included in case the respondent is neither an academic nor a student, but is instead someone in industry with an interest in Web Science.) The respondent is then asked to state the level at which they teach or study Web Science, and as well as number of other open questions which provides them with the opportunity to explain their background in the subject.

7.3.1.2 Survey Section Two

The second section relates to the research question 'what is taught as web science', and relates to the gathering of information about the academic institution or organisation that the respondent is affiliated with, and is based upon the fields shown in Table 23: Key Fields Used in the Modules Desk Survey'.

7.3.1.3 Survey Section Three
The third and final section attempts to collate information relating to the respondents opinion of how specific subjects are relevant to Web Science. This section uses the WSSC as a template for rating the importance of subjects within Web Science, and was based on the layout used by the survey outlined by (White, et al., 2011).Table 25. shows a list of the key survey questions, and the corresponding research question that they relate to. (***Please note; survey questions are repeated if they relate to more than one research question.)***

7.3.2 Promoting the Survey: Difficulties Faced
One of the key difficulties faced with this form of survey is a lack of participants. As explained by (Ackland 2013), a low response rate is the most common difficulty faced with online surveys. (Ackland 2013) suggests that graphics may be used to improve the appeal of a survey, although, there is no statistical evidence to prove that it actually improves response rates. A promotional flyer was created with the aim of advertising the survey. These flyers were used to promote the survey at the Web Science conference.

The survey was also promoted using a number of Web Science dedicated Facebook groups. This yielded the highest levels of participation of promotion methods so far. This may be due to the fact that people are usually at their computers when they check the groups and then fill out the survey immediately, instead of being handed a flyer on the move, and then forgetting to follow it up. The fact that there was a spike in participants in a number of hours after each post confirms this. The spike in the number of participants inclines for roughly twenty-four hours after each post, before dwindling. Another frequent observation, includes the fact that out of the number of people attempting the survey, only roughly a third of people then go on to actually complete it.

A higher number of respondents have been achieved from promoting the survey on Facebook groups than using the flyers or promoting participation at the Web Science conference. Therefore the conclusion is that Social media is a more efficient way of promoting the survey than traditional methods. This could be because people are conveniently placed to respond immediately if they see a survey when they are already on the Web.

7.3.3 Survey Questions vs Research Questions
Table 25 outlines the survey questions that feature in the questionnaire, classified according to the overall project research questions. Survey questions appear in the category of each research question to which they relate.

Table 25. Survey Questions in relation to Research Questions

Research Question	**1. What is the extent of the Web Science Subject?**
Related Survey Question(s)	• Section 1. Question 7. Are you aware of the term "Web Science"? • Section1. Question 8. If you have a background in Web Science, please explain any significant contributions you have made to the discipline: • Section 2. Question 5. Please provide a summary of the key topics covered by your institution’s Web Science (or related) programme(s): (If this is a full degree programme, please provide module titles if possible) • Section 3. Question 1. Please rate the following categories relating to: WEB TECHNOLOGIES • Section 3.Question 2. Please rate the following categories relating to: WEB ANALYSIS • Section 3.Question 3. Please rate the following categories relating to: WEB SOCIETY • Section 3. Question 4.Do you feel that the above framework of the WSSC omits any key subjects which relate to Web Science or your individual curriculum? If yes, please list subjects and why you feel they are relevant: • Section 3. Question 5. Are you aware of any differences between the terms ‘Web Science’ and ‘Internet Science’? If yes, please elaborate.
Research Question	***2. Where* is Web Science taught?**
Related Survey Question(s)	• Section 1. Question 4: At what level do you teach or study? • Section 2. Question 1. What is the name of your academic institution or company? • Section 2. Question 2. In which part of the world is your institution located? • Section 2. Question 4. What level(s) of Web Science (or related) qualification are taught by your academic institution? (Please tick all that apply) • Section 2. Question 6. Please list the language(s) in which your programme is taught: • Section 2. Question 7. Please provide an estimate of the number of teaching hours within your Web Science (or related) taught programme: • Section 2.Question 8. What is the duration of the programme? (Please give dates if possible) • Section 2. Question 9. What date was your programme first created? (Optional) Section 2. Question 10. In what way(s) is your programme taught? (Lectures, online material, etc.)
Research Question	**3. *What* is taught as Web Science**
Related Survey Question(s)	• Section 2. Question 3. Please give the name(s) of the Web Science (or related) taught programmes offered by your institution Section 2. Question 5. Please provide a summary of the key topics covered by your institution’s Web Science (or related) programme(s): (If this is a full degree programme, please provide module titles if possible)
Research Question	**4. What subjects are studied by Web Scientists in practice?**
Related Survey Question(s)	• Section1 Question 5.Which field(s) best represent your educational or work background? Please tick all that apply: • Section 1. Question 6.Please specify your degree title, research field or job title • Section1. Question 8. If you have a background in Web Science, please explain any significant contributions you have made to the discipline:

This table shows the research question each survey question addresses. The survey was conducted over a number of months during 2015/2016; the primary participants targeted were the attendees of the 2015 Web Science conference, held in Oxford. The survey was also promoted to Web Scientists worldwide via social media, including on a number of Web Science Facebook groups. The full results of the survey will be presented and analysed in Chapter 8.

7.4 WAIS Fest Events

WAIS Fest (WAIS standing for Web and Internet Science) is an annual research event held by the Web and Internet Science research group at the University of Southampton. This annual research event provides a short period of time (often three days) in which researchers have the opportunity to work on short projects which differ from their chosen research subjects. Researchers requiring outside input into their own research also have the opportunity to put forward ideas for WAIS Fest project topics, in order to enlist the help of others within the research group. Therefore, the author made use of this opportunity on two occasions. The first WAIS Fest theme relating to this work took place in 2016, and is described below. A description of the second event is provided in section 9.2, as it directly related to structuring the proposed Web Science Lexicon.

7.4.1.1 Web Science PhD Researcher Interviews

The final data source constituted of a series of interviews with Web Science PhD students from the University of Southampton Web and Internet Science research group. This was proposed in order to supplement content provided by the survey of Web Science students and academics. This will facilitate first-hand insight into the subjects studied by Web Science researchers in practice, as well as providing insight into which subjects within Web Science are being chosen by students as research topics. The 2016 WAIS Fest event was utilised as an opportunity to interview Web Scientists within the research group, in order to gain a deeper insight into the aspects of Web Science being studied by Web Science researchers in practice as well as their individual views on the subject. Prior to WAIS Fest, ethics approval was first sought in order to conduct a study involving human participants. The ethics application can be found in appendix section appendix section 17. A total of 16 participants agreed to be interviewed, the majority within the WAIS Fest period, with several additional participants also agreeing to be interviewed at a later date. The list of questions asked during the interviews can be found in section 8.7.1 and the analysis of the results can be found in section 8.7.6.

7.5 The Web Science Conference Literature

Another key data source utilised during this thesis is the Web Science conference series. The papers published during this conference provide a valuable insight into what subjects are being researched and published by the Web Science community. The foundation of the Web Science conference dates back to 2009, with the first conference held in Athens. There have been a total of 9 Web Science conferences to date. The full papers published in these conferences will be analysed utilising the thematic analysis approach described in chapter 6, in order to determine what subjects are most commonly published as part of the Web Science community. This will help to determine the scope of Web Science according to what is published within the Web Science research community, and should help to identify the topics being researched by Web Scientists in practice.

7.6 Conclusions/Future Work

This chapter has outlined the data sources which will be utilised during this thesis. These include a desk survey of Web Science teaching institutions and modules, a simple Web Crawler to supplement the findings of the desk survey, a questionnaire survey of Web Science students and academics, a series of interviews with Southampton Web Science PhD researchers, and a review of Web Science conference literature. The full analysis and results produced from the data sources described in this chapter will be outlined in chapter 8.

8 Analysis and Results

8.1 Introduction

This chapter presents the analysis and results of the data sources outlined in chapter 7. The process of analysing the data sources outlined in this chapter corresponds with the Grounded theory process of 'Open Coding' as described by (Strauss & Corbin 1998). This involves the initial identification of themes in the data, which in the case of this thesis will involve the identification of 'key topics' (originating from the analysis of Web Science taught modules) and 'paper topics' (originating from the analysis of Web Science conference publications). The differentiation between the two sets of topics is maintained in order to facilitate comparison between what is taught as Web Science and what is published. The data from the analysis of Web Science taught modules and the analysis pf Web Science conference publications provided the most detailed data, and are therefore the from key basis for deriving the Lexicon subjects. However, these analyses will also be supplemented with the findings of the survey and interviews of Web Science academics and students. While providing less detailed results than the prior two analyses, the contribution of input from Web Scientists provided by the survey and interviews will provide valuable input for verification and comparison with the results of the two prior analyses.

Much of the data analysis described in this chapter is conducted qualitatively utilising manual coding in a Word document and later in Nvivo. However, the results of coding the key topics and paper topics identified are quantified and displayed in terms of coding frequency for ease of presentation.

The final outcome of this study is to present a taxonomy of Web Science subjects, to be named the Web Science Lexicon, which will represent Web Science as it is taught and published. The process of structuring the Lexicon will be presented in the following chapter 9.

This chapter has the following structure:

- Analysing the Desk Survey Data
- Analysis of Web Science conference data
- Presentation of Survey Results
- Review of Interview results

8.2 Desk Survey Data: Analysing Web Science Taught Modules

As outlined in Chapter 7, a desk survey was conducted, in order to identify Web Science taught modules. This Web Science taught module data was collated and stored in a 'modules' spreadsheet, and stored under the fields shown in Table 26. A total of 146 modules were identified across 16 institutions which offered programmes featuring taught modules.

Table 26. Fields used in the modules desk survey

Intuition Location	Resource Title	Topic(s) Taught	Teaching\ Assessment format
Contributors(s)	Materials Used (urls)	Level (e.g. Masters)	Module Dates\Duration

Following the process of data collection, the spreadsheet data was then imported into Nvivo for analysis. In order to facilitate ongoing identification of additional institutions, modules from each institution were sorted into separate worksheets, one for each institution, before being imported into Nvivo. The insertion of modules from each institution in a separate worksheet also facilitated the automatic generation of an Nvivo case node for each institution, in order to facilitating easier querying of the data in Nvivo following the coding process.

8.2.1 Manual Thematic Analysis of Web Science Taught Programmes

Once imported into Nvivo, the 'module topics' field for each for each Web Science taught programme was manually examined and coded. As previously described in chapter 6, a thematic analysis was conducted following guidelines suggested by (Strauss & Corbin 1998). The initial stages of Opening Coding involved manual coding of recurring themes in the data, labelling each of these themes identified as a 'key topic'. Further iterations of coding saw recurring occurrences of these key topics coded throughout all the modules. The titles of the modules were used a guide as to the content of each module. Module titles are generally effective summaries and concise deceptions of content, therefore if a keyword appeared in the module title, it was frequently coded as a key topic. Topics not in titles, but which were found to be recurring, were also coded as key topics. Following the initial Open Coding iteration, additional iterations of coding were used to provide thorough coding coverage. Topics which were identified later in the coding process where then identified in some of the earlier coded modules, and therefore added to coding. The initial Open coding phase saw higher-level topics being coded, for example, 'Web Technologies' or 'Web languages'. During following iterations of coding, frequently occurring lower level key topics were identified and coded, for example, 'Data Mining'. Although a lower level, more specific topic, it was identified as an example of a frequently recurring topic, it was deemed worth coding it as a topic in its own right.

The identification of lower level key topics can be cited as a form of Axial Coding. However, the final structuring of the key topics and identification of higher level topics and lower level topics will be conducted once the analysis of all data is complete. The structured key topics will form the basis of the proposed Web Science Lexicon. Additional data sources such as the Web Science conference papers may yield more information about which topics occur more frequently, and are therefore potentially more important, higher-level key topics. Also, additional new topics and sub topics will be identified from other data sources, and these will also contribute to the structure of the Lexicon.

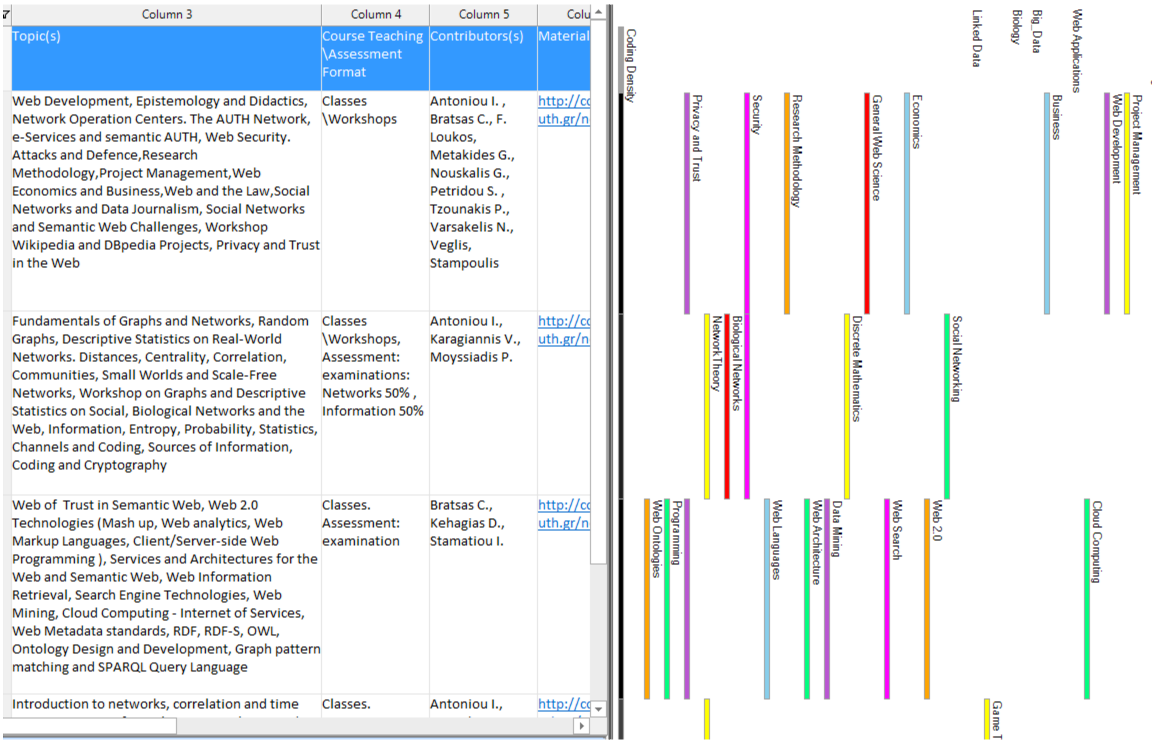


Figure 4. Coding Key Topics in Nvivo

Figure 4 shows the Nvivo ‘coding stripes’ which indicate coded data. The process of coding the data required several iterations of coding in order to fully identify and saturate the coding of the data.

8.2.2 Key Topics

Following several iterations of coding in Nvivo, a list of key topics were identified. In order to present and the data and illustrate the most commonly identified key topics or themes, the qualitative data was quantified and presented in a table. Nvivo shows the frequency at which nodes are coded, and this information was presented in table format. Table 27 shows the key topics identified. The darker colours indicate more frequently occurring key topics. These key topics reflect the most commonly taught subjects within the Web Science programmes identified. The ‘sources’ column indicates the number of programmes in which each given topic appears. So for example, the key topic ‘Business’ appears in a total of 7 taught programmes. The ‘refs’ column shows the frequency at which each key topic occurs overall throughout all programmes and within each programme. For example, the key topic ‘Business’ appears 11 times; the ‘sources’ column indicates that it is present in a total of 7 sources or programmes, so therefore the ‘refs’ column shows that within the 7 sources that the key topic appears, it was identified in 11 instances.

Table 27 Web Science taught modules: Key Topics Identified

Name	Sources	Refs	Name	Sources	Refs
Accessibility	3	3	Network Theory	9	16
Artificial_Intelligence	3	3	Operating Systems	1	1
Big_Data	7	9	Performance Management	1	1
Biological Networks	2	3	Politics	5	6
Biology	3	3	Privacy and Trust	8	9
Business	7	11	Programming	4	6
Cloud Computing	5	6	Project Management	3	3
Computer Graphics	2	2	Psychology	3	6
Computer Vision	2	4	Quality Management	1	1
Criminology	1	1	Research Methodology	8	16
Data Mining	7	8	Risk Management	2	2
Digital Copyright	4	5	Security	5	9
Digital Ethnography	1	1	Semantic Web	10	21
Digital Sociology	4	6	Social Networking	11	14
Discrete Mathematics	1	1	Sociology	6	17
Economics	7	11	Software Agents	1	1
Education	1	1	Software Engineering	4	8
e-Learning	1	1	Statistical Analysis	4	9
Entrepreneurship	2	2	Systems Design	4	6
Ethics	2	3	Virtual Environments	1	1
Game Theory	3	3	Visualisation	6	6
General Web Science	5	7	Web 2.0	6	6
Hardware	3	5	Web Analytics	5	7
HCI Human Computer Interaction	3	3	Web Applications	4	4
Hypertext	2	3	Web Architecture	8	12
ICT Information Communications Technology	5	9	Web Design	3	4
Information Systems	2	2	Web Development	5	5
Information Theory	2	2	Web Governance	4	5
Internet of Things	3	5	Web Graph	4	7
Internet Protocols	1	1	Web History	3	3
Law	6	10	Web Languages	10	15
Linked Data	5	7	Web Modelling	5	8
Management	4	6	Web Ontologies	5	6
Marketing	1	2	Web Search	5	7
Media	5	7	Web Society	9	23
Mobile Web	3	3	Web Technologies	7	8
Networks	9	24			

The entire range of key topics identified indicate slightly more of a bias towards technical subjects, although the key topics most frequently coded suggest a fairly even spread between socio-technical subjects. The higher level (e.g.'Web') topics have a tenancy to be more frequently coded, but this is likely to be due to the fact that the higher level topics naturally cover a wide range of sub topics, whereas any lower level topics identified only occur in a small number of instances. Descriptions of the most frequently occurring topics will also be included in the final proposals for the structure of the Web Science Lexicon, shown in section 9.

8.3 Analysis of the Web Science Conference Series

Since 2009, there has been an annual conference dedicated to Web Science. The publications produced as part of this conference represent a valuable data source for understanding what is published as part of Web Science in practice. To this end, an analysis was conducted of the papers published in the Web Science conference series, from 2009 to 2017. A total of 357 papers were included in the analysis of Web Science conference proceedings, see Table 28.

Table 28. Web Science Conference Papers included in the Analysis

Web Science Papers Included in the Analysis	
Conference Year	**No of Papers Analysed**
WebSci 2009 Athens, Greece 18th–20th March, 2009	**27**
WebSci10. Raleigh, North Carolina, 26 & 27 April 2010	**33**
WebSci11, Koblenz, Germany, June 14-17, 2011	**30**
WebSci '12, Evanston, IL, USA - June 22 - 24, 2012	**47**
WebSci 2013 Paris, France, 02-04 May 2013,	**75**
WebSci 2014, Bloomington, USA, June 23 to June 26, 2014	**33**
WebSci 2015, Oxford, United Kingdom, June 28 - July 1, 2015	**32**
WebSci 2016, Hannover, Germany, May 22 to May 25, 2016	**36**
WebSci 2017, Troy NY, USA, 25-28 June, 2017	**44**

8.3.1 Data Recorded

All full papers in each of the Web Science conferences from 2009 to 2017 (See Table 30) were analysed utilising this same process. For each paper, the following information was recorded for each paper:

- Title,
- authors,
- key topics,
- paper key topics,
- author keywords.

This same information was recorded for each full-length paper, and stored in a Word document in the following format:

Title: Closed-Loop Opinion Formation **Authors:** Larissa Spinelli, Mark Crovella **Key Topics:** Psychology, Web Society, Web Analytics, Marketing, Web Modelling, **Paper Key Topics:** Content Analysis, Online Engagement, Recommender Systems, User Behaviour **Author Keywords**: Filter Bubble, User Experience, Metrics, User Behaviour, Recommender Systems

The two separate lists of key topics and paper topics were recorded and maintained. This was in order to facilitate a comparison between key topics identified form Web Science taught programmes, and paper topics published within Web Science conferences. The complete list of coded papers can be found in appendix 19.

8.3.2 Coding the Web Science Conference Papers

As seen in section 8.2.2, an initial list of 'key topics' representing Web Science as it is taught was identified, following the analysis and coding of the Web Science taught programmes. (See Table 27) This list of key topics was utilised as a starting point for coding key topics within the Web Science conference literature. However, many of the papers analysed included new topics not previously identified within the analysis of taught programmes. Table 29 outlines the criteria which were used to decide the topics which were coded within the analysis of conference literature.

Table 29 Criteria for Coding of Web Science Conference Literature

Criteria for Web Science conference paper coding
• Existing key topics are coded when located within a paper • High level subject areas such as 'social science' which are not already present in the list of key topics are coded • Frequently discussed topics, present in more than one or two papers are coded, e.g. 'open data' • Where a topic is the main focus of a paper, if it is not already represented by an existing key or paper topic, it is coded as a new paper topic. For example the topic 'nutrition' is not frequently discussed, but it is the main focus of one paper, and is not represented by any existing topic. It is coded, despite being an example of the long tail. • If a topic is included within the author keywords in several papers, it is considered for inclusion if it is deemed to not be well represented by existing key or paper topics coded.

The process of analysing the conference papers was as follows; each paper was manually read. During the initial read through and first iteration of coding, the list of the key topics identified from coding the Web Science taught modules was consulted, and any key topics present in a paper were manually identified and recorded for each conference paper. Where the paper included a notable topic not represented by the list of key topics, this new topic was recorded, contributing to a new list of paper specific 'paper topics'.

Early in the coding process, only new paper topics deemed as high level, e.g. 'Social Science' were coded. However, as the analysis progressed, additional lower level paper topics, such as for example 'Open data', were identified as commonly discussed topics within papers, and on some occasions, such topics also demonstrated common occurrences in the lists of author keywords. While for example, 'open data' is similar to some existing key topics such as 'linked data', it is not identical, and it is something which is mentioned frequently in its own right, and is even the main topic of some publications. Therefore, such topics were deemed to be relevant and worthy of coding within the list of paper topics. Some paper topics such as for example, 'algorithms', are topics with a broad context, which in this case has the potential to refer to many types of algorithm. However, it was mentioned numerous times in the author keywords sections of papers, frequently enough to be of note. It is also a frequent tool utilised to address research questions outlined in Web Science papers, and is therefore a notable component of Web Science research, worth recording and including in the list of paper topics.

Upon completion of the manual word document based analysis for each of the Web Science conferences, the key topics and paper topics identified and coded where were then coded using Nvivo. This second iteration of coding enabled further saturation of coding, and identification of additional key topics and paper topics in the papers which had not previously been identified on the first iteration of coding.

The process of reading and coding the papers could arguably have been conducted automatically utilising an algorithm. While this is true, an automated analysis would have been unable to interpret the context of the papers. For example, a paper might heavily feature the key topic ‘data mining’. However, it may not specifically mention the keyword ‘data mining’. It may explain in depth, the process of collecting data, without once mentioning the keyword ‘data mining’. An automated analysis would potentially miss this, whereas a human is able to interpret the meaning and context of the paper content and identify a topic, even though a specific keyword is not mentioned. While manual analysis is far from flawless, it does allow for human interpretation of subjects, and also has the benefit of the researcher conducting the analysis gaining knowledge of the subject.

Author keywords were not coded in Nvivo, because, while descriptive of the content, they did not conform to consistent naming conventions. For example, one paper might have a keyword ‘social networks’, while another might have the keyword ‘social media’. Therefore, the paper topics attempt to encapsulate and represent the author keywords. This will facilitate consistent naming allowing a related theme to be fully coded, instead of some papers being coded with e.g. ‘social media’ and others with ‘social networking’, whereas in reality they are both discussing the same topic.

Where a paper topic is only coded against one or two nodes, it generally indicates that the given topic only appears as focus of one or two papers, but where it does occur, it is the key focus of the given paper, and it therefore deemed worthy of coding with a dedicated paper topic. For example, in the case of (Arbor et al. 2012), ‘Recipe recommendation using ingredient networks’, this paper was coded with the paper topic, ‘nutrition’, as this was the key focus of the paper. Nutrition is only coded on this one occasion, as it is not a topic that is frequently discussed, however its inclusion does represent the diversity of topics which can feature as part of Web Science. During the structuring of the final Web Science Lexicon, the decision will need to be made as to whether to include such topics in the taxonomy, as while they represent the diversity of Web Science; alone they represent examples of the ‘long tail’ of Web Science, and a very small portion of Web Science research.

Table 30 Web Science Paper Topics identified

Name	Freq.	Name	Freq.
Academic publishing	11	MOOCs	2
Algorithms	31	Narrative	14
Archaeology	1	Natural Language Processing	4
Arts and Culture	9	Network Science	47
Bibliometrics	2	Nutrition	1
Cognitive Science	8	Online Advertising	6
Communication Science	23	Online Engagement	41
Computer Science	15	Online Identity	14
Content Analysis	121	Online Offline Community	10
Crowdsourcing	23	Open Access	4
Cyber Bullying	5	Open Data	18
Cybercrime	11	Philosophy	7
Democracy	7	Provenance	2
Demographics	15	Quality Control	5
Digital literacy	4	RDF	10
Digital Native	2	Real-time data	3
Disaster Response	5	Recommender Systems	8
eCommerce	9	Religion	3
Environmental Science	1	Scientific Method	5
Folksonomy	4	Sentiment Analysis	15
Gamification	4	Social Machines	12
Geographic Information Systems GIS	3	Social Science	79
Geography	28	Socialbot	2
Geo-tagging	12	Synthesis Ranking	1
Government	16	Theology	1
Health	15	Topic Modelling	19
Information Theory	1	User Behaviour	145
Journalism	4	Virtual Community	42
Knowledge patterns	10	Web 3.0	2
Machine Learning	2	Web Archiving	25
Microblogging	77	Web Publishing	21

Table 30 shows the quantified results of the qualitative analysis, illustrating the frequency at which the paper topics where coded. The following Table 31 shows the range and frequency of key topics which were coded within the paper analysis. Some topics have a value of '0' in the frequency column; this indicates that while these key topics were defined and coded in the prior modules analysis, these key topics were not identified during the paper analysis.

Table 31 Number of sources coded for key topics within Web Science conference papers

Name	Freq.	Name	Freq.
Accessibility	8	NetworkTheory	28
Artificial_Intelligence	5	Operating Systems	0
Big_Data	51	Performance Management	0
Biological Networks	1	Politics	28
Biology	1	Privacy and Trust	45
Business	4	Programming	0
Cloud Computing	6	Project Management	3
Computer Graphics	3	Psychology	81
Computer Vision	0	Quality Management	22
Criminology	3	Research Methodology	38
Data Mining	98	Risk Management	4
Digital Copyright	2	Security	8
Digital Ethnography	9	Semantic Web	40
Digital Sociology	87	Social Networking	182
Discrete Mathematics	0	Sociology	25
Economics	9	Software Agents	0
Education	13	Software Engineering	0
e-Learning	6	Statistical Analysis	21
Entrepreneurship	1	Systems Design	2
Ethics	9	Virtual Environments	1
Game Theory	1	Visualisation	7
General Web Science	38	Web 2.0	26
Hardware	2	Web Analytics	213
HCI Human Computer Interaction	23	Web Applications	15
Hypertext	12	Web Architecture	21
Information and Communications Technology	7	Web Design	4
Information Systems	5	Web Development	8
Information Theory	0	Web Governance	4
Internet of Things	4	Web Graph	29
Internet Protocols	3	Web History	11
Law	8	Web Languages	5
Linked Data	55	Web Modelling	36
Management	1	Web Ontologies	22
Marketing	13	Web Search	23
Media	45	Web Society	223
Mobile Web	5	Web Technologies	27
Networks	34		

While the key topics shown in Table 31 were already identified during the analysis of Web Science taught modules, and therefore classified within the key topics list, these key topics may be just as relevant to the Web Science conference publications. To this end, Table 31 shows the frequency of coding for key topics within the analysis of Web Science conference publications. The paper topics include a wider range of specific activities relating to the Web, than those within the key topics identified during the analysis of Web Science taught modules. Such examples from the paper analysis include 'sentiment analysis' and 'Topic modelling'. The greater diversity of Web based activities within papers compared with those identified during the modules analysis, represents the wider range of research covered in the conferences, compared to the limitations of specific curricula. This is why a combination of both data sources is useful for determining the overall scope of Web Science

8.4 Utilising Nvivo automated text search to saturate coding of topics

Once initial coding was complete, it was suspected that some topics identified during later stages of analysis had be omitted from earlier stages of coding, at which time they may not yet have been classified as a keyword. Where coding of a key topic or paper topic was suspected not to be fully saturated, Nvivo automated text search was then used to search for any topic which was suspected to be 'under coded'. Nvivo automated text search and coding is potentially beneficial when utilised as a supplementary tool, however it needed to be employed with caution. For example, even if the key topic 'computer science' is mentioned, it may only be included once as a brief reference, and may not be the focus of the paper. It would therefore be inappropriate to code it as a keyword.

Nvivo facilitates the search for stemmed words and synonyms, so it is possible to search for, for example, plurals of a word. However, a significant drawback of Nvivo's text search function, is that it seemingly only works effectively for a single word topic such as 'algorithms'. Where a phrase topic needs to be searched for, such as for example, 'cloud computing', Nvivo will return results which include 'computing' in isolation, without the word 'cloud'. This is unhelpful, as many of the key topics and paper topics identified consist of more than one word. Mendeley's text search feature proved to be a somewhat more useful tool, as when a key topic of paper topic is searched for, Mendeley usefully highlights the part of the document in which it appears. It is also more effective than Nvivo at returning results for topics consisting of more than one word, however it also returns results where only one of the words is present.

8.4.1 Nvivo Automated Keyword Analysis

The Nvivo automated 'Word Query' function was utilised as an experiment with the aim of providing a qualitative comparison to the manual qualitative analysis and coding of Web Science taught modules and Web Science conference publications. For the purpose of this analysis, results were limited to the 100 most frequently occurring of the top keywords for both the Web Science taught modules and the Web Science conference publications. A significant limitation of this method of automated keyword analysis, is the fact that Nvivo is only able to detect keywords, and not 'key phrases'. For example, Nvivo can detect the number of occurrences of the isolated words 'Web' and 'Science', but not a combination of the two.

8.4.1.1 Conference keyword and module keyword comparison

A comparison was made between the top 100 conference keywords and the top 100 keywords identified from the modules desk survey data. Specifically, those from the 'module topics' field, which describe what subjects are being covered as part of Web Science taught modules. Investigations suggested that Nvivo is not capable of performing this type of comparison; to this end, MS Excel was used. Research into Excel tools revealed that utilisation of the 'conditional formatting' feature facilitated easy comparison of the contents of two columns. The process for this was carried out using the following method: The two columns were selected, followed by `'conditional formatting' > new rule > select 'format only unique or duplicate values'` and then choose one of the two options, in addition to formatting the display colour using the 'format' button. The results of this produced the following colour coded cells, **Blue** for matching cells, and **Green** for unique values.

Table 32 shows the most frequently occurring results of the analysis. Initially it was thought that there was an error in the function, as two words in the left-hand column were highlighted as duplicates, and yet they did not appear in the right-hand column. It was then noticed that in two cases, the left-hand column contained two instances of the same word. Because of Nvivo's lack of ability to detect keywords comprising of more than one word, it is difficult to make a direct comparison between the keywords found in this analysis and those identified by the manual analysis. In isolation, many of the keywords identified in the Nvivo analysis also lack context.

8.4.2 Flaws of Nvivo Keyword Analysis Feature

Whilst the automated detection of keywords and the subsequent comparison was intended to provide a point for comparison with the 'key topics' identified during the manual analysis, the automated keywords lack the level of useful detail provided by the manual analysis. A key issue is Nvivo's lack of ability to detect key phrases', as many of the terms within Web Science, such as 'social networking' and 'semantic web' consist of more than one word. Words such as 'web' mean little in insolation, as context is unknown, and the words may occur in multiple combinations, for example, Web Science, or 'Web of Data', or 'Semantic Web'. The NVivo automated analysis also lacks the ability to filter out erroneous content, for example, 'com' refers to 'dot com' from a web address, which provides little useful insight. The keyword Nvivo keyword detection also lacks context; for example, the keyword '2011' means nothing in isolation, presumably, it refers to publication dates, but it could also be that a large number of the references included in papers have the date 2011. In isolation, 2011 has little context; it is also another example of a 'keyword' which fails to directly describe Web Science as a subject.

Table 32. Conference and module keyword comparison

Conference Keywords	Module Keywords	Conference Keywords	Module Keywords	Conference Keywords	Module Keywords
social	web	tweets	language	process	quality
data	social	different	management	value	time
web	data	features	computer	see	application
users	networks	however	engineering	digital	content
information	course	table	languages	public	performance
user	information	http	services	semantic	technical
number	science	example	society	graph	used
one	network	com	world	section	using
also	research	well	approaches	2009	well
network	design	many	retrieval	words	approach
time	systems	search	security	trust	communication
research	analysis	system	tools	org	future
twitter	media	approach	user	large	mining
based	new	acm	online	world	professional
figure	understanding	given	process	conference	theory
online	technologies	paper	search	first	topics
work	models	within	concepts	like	areas
used	applications	2011	virtual	different	critical
analysis	computing	related	also	page	impact
may	students	high	big	open	intelligence
media	issues	2010	understand	first	marketing
new	techniques	pages	architectures	degree	participation
two	development	nodes	business	event	processes
using	methods	knowledge	context	distribution	professionals
science	semantic	university	principles	post	protocols
people	technology	systems	algorithms	important	related
networks	digital	topic	architecture	level	structures
content	internet	facebook	environments	make	study
set	module	human	including	similar	analyse
results	software	terms	innovation	wikipedia	behaviour
model	introduction	order	interaction	values	concept
community	based	group	mobile	news	develop
study	knowledge	Internet	programming	part	human

While a more detailed automated analysis capable of detecting key phrases would potentially provide a useful comparison with the manual analysis, the analysis features offered by Nvivo are somewhere limited. Future work could include the development of an algorithm facilitating more advanced textual analysis, however this is beyond scope of this thesis and the current skills of skills of the author.

8.5 Analysis and Results of the Questionnaire Survey of Web Science Students and Academics

The section shows the responses to the questionnaire survey of Web Science students and provides some thoughts and analysis of the implications of the responses. The questions are numbered according to the three sections included in the online iSurvey. These three sections include the following areas, as previously explained in greater detail in section 7.3.1:

- Survey Section 1: Background of Participants
- Survey Section 2: Your Curriculum Details
- Survey Section 3: Curriculum Responses: Rating the WSSC responses

The questions are listed in order, however, not all questions are included. Reasons for this include the need to protect the identity and confidentiality of participants. This section also only includes questions which provide useful insight which contributes to the study. Where sections yielded limited useful information, they are omitted.

8.5.1 Survey Section 1: Background of Participants

The survey of Web Science students and academics was completed by a total of 42 participants. The first few questions relate to the identities of participants, and are therefore omitted from the presentation of the results in order to maintain the confidentiality of participants.

8.5.1.1 Question 1.4: At what level do you teach or study?

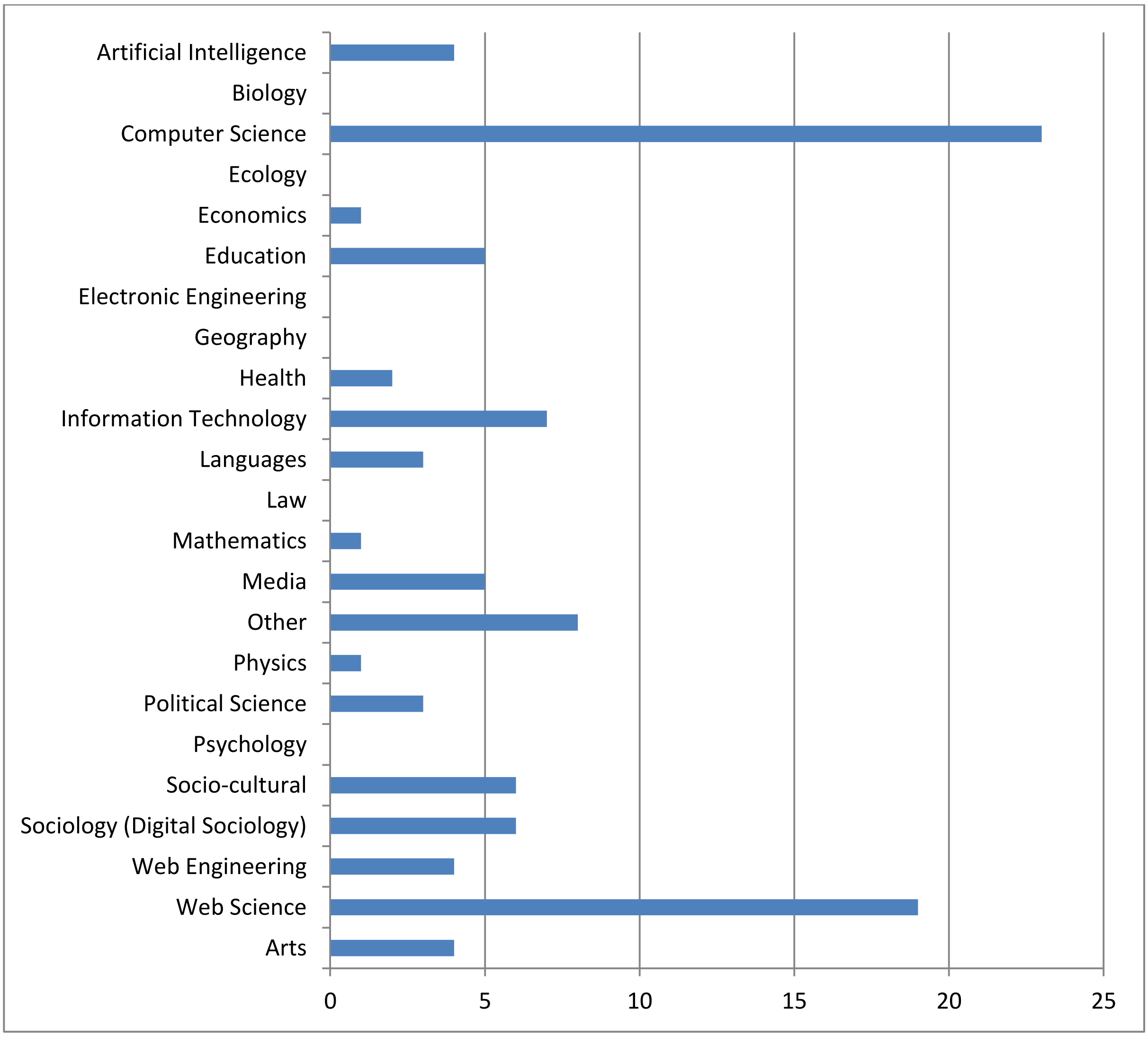
Of the 42 participants respondents, a total of 28 identified as students, 8 identified themselves as academics, while 6 choose the 'other' category. Of the respondents who chose the 'other' category, 4 choose to clarify. They identified as the following:

1. Postdoc researcher (formerly WebSci student)
2. Applied researcher
3. Completed PhD in CS at Soton
4. Information security consultant

8.5.1.2 *Question 1.5: Which field(s) best represent your educational or work background? Please choose all that apply:*

Participants were asked to select a range of options representing their disciplinary background. The topics included in the list of options originated from the main headings of the Web Science Cluster Diagram, with a few extra options, such as health and education. The following list in Table 33 shows the collection of results. Unsurprisingly the most frequent choices were Computer Science and Web Science. Areas which seem to be underrepresented include Law, which would correspond with the (Hooper et. al 2013) findings, which suggest that Law is an area under represented within publications at the Web Science conference series.

Table 33 Subject Areas of Survey Participants



Question 1.5a: If other, please elaborate:

Of the few participants who selected the 'other' option and chose to elaborate, the following responses were given:

- *Music*
- *Book-Retail Marketing*
- *Communications*
- *Criminology*
- *Commerce (practical experience - not Economic theory Information/Knowledge management*
- *Marketing Design*
- *Philosophy*
- *Business*

The diversity of these responses confirm the board interdisciplinary origins of Web Science students and academics. The majority of these additional subjects recommended by respondents were identified in either the list of key topics or the list of paper topics. Book-retail marketing is a very specific subject, and is not directly covered by either key topics or paper topics, however general 'marketing' represents it in a broad context. Music is the other example which is not specifically covered, however 'arts and culture' and 'media' may adequately cover it.

8.5.1.3 Question 1.7 Are you aware of the term "Web Science"?

This question asked respondents if they were aware of the term 'Web Science', to which all answered conclusively, 'Yes'.

8.5.1.4 Question 1.7a: If yes, please provide a brief summary of what you believe the term 'Web Science' means:

Respondents where then asked to describe the nature of Web Science. The following Table 34 shows the responses to this question.

Table 34. Survey respondents' descriptions of Web Science

No.	How do respondents describe Web Science?
1	The social scientific study of the World Wide Web
2	An interdisciplinary approach to studying the reciprocal effects of the Web on society.
3	Web Science is the scientific investigation and shaping of the socio-technological system that is the Web.
4	The theory and practice of social machines; an inherently interdisciplinary effort to both understand and influence the web's evolution.
5	interdisciplinary approach to understand the web |
6	I have absolutely no idea.
7	Web Science deals with Information Science Computing and the Internet from a sociotechnical perspective.
8	The study of the Internet's technical and social elements particularly as they relate to the traditions of the University of Southampton's Comp Sci programme.
9	It is the study of systems like the world wide web and the interaction between this technology and the users
10	Interdisciplinary Study of the Web as a Social and Technical Phenomenon
11	The study of the relationship between people and technology
12	Web Science is the study of systems composed of people and technology arising from technology such as the web. It speaks both to technology and society.
13	Science concerning the relationship of mutual influence between the web and society
14	A field dedicated to understanding more about the technical and social implications of web technologies
15	Web Science is the interdisciplinary study of the relationship between people (society) and the Web - The

No.	How do respondents describe Web Science?
	effect society has had on the development and nature of the Web and the effect the Web has had on the nature of society.
16	Is the study of the Web and to explore topics related to the impact that the Web has on the society and technology
17	'The way the Web influences the world and the world influences the Web'(to quote someone or other!)
18	The multidisciplinary study of the Web as a socio-technical phenomenon.
19	Web Science is a multi-disciplinary practice rooted in computer science which attempts to examine the socio-technical complexities associated with the Web.
20	In my general terms the study of the Web in all aspects and from all facets. How it is engineered how it affects people societies countries. How it is used. It is the study of and engineering of the Web. ||Formally I go along with what's on the web (wikipedia)||Web science is the study of large-scale socio-technical systems such as the World Wide Web. It considers the relationship between people and technology the ways that society and technology co-constitute one another and the impact of this co-constitution on broader society. Web Science is inherently interdisciplinary[1][2] and combines research from disciplines as diverse as sociology computer science economics and mathematics.[3]||[1] Berners-Lee T.; Hall W.; Hendler J.; Shadbolt N.; Weitzner D. (2006). ';Computer Science: Enhanced: Creating a Science of the Web; . Science 313 (5788): 769 771. doi:10.1126/science.1126902 . PMID 16902115 .|[2] Hendler Jim; Shadbolt Nigel; Hall Wendy; Berners-Lee Tim; Weitzner Daniel (2008). 'Web science: an interdisciplinary approach to understanding the web' (PDF). Communications of the ACM 51 (7). doi:10.1145/1364782.1364798 .|[3] Why do we need web science research http://www.slideshare.net/webscikorea/why-do-we-need-web-science-research
21	Understanding how the web impacts society and how society impacts the web
22	The study of the Web as a socio-technical phenomenon from a multidisciplinary and interdisciplinary perspective.
23	Web Science explores large-scale sociotechnical constructs and behaviours attending on uses of the World Wide Web.
24	Science of Web. Well Web is a huge net drawn and it covers and affects various disciplines. When we study the sociological and political changes we do use lot of computing algorithmic and mathematical tools at our disposal to study not just the history but also predict the evolution of the web.
25	Understanding the social and technical underpinnings of the web in order to limit the potential problems that might occur and further its use in a positive way
26	A study of the World Wide Web (and Internet based interaction) as a socio-techinal system.
27	The area of study of the Web and the impact it produces on society.
28	The study of both technical and social aspects of the web.
29	Studying the web as a culture and sociological entity
30	the study of how the Web intersects with society and vice versa how each impacts the other from an interdisciplinary standpoint.
31	interdisciplinary study of the web
32	It is an interdisciplinary research area in which research questions span fields of law education computer science psychology .... And these research questions should be in some way related to the web.
33	An interdisciplinary study of the Web and its impacts on society.
34	study the web
35	The study of the technosocial processes and interactions between society and the web; the study of how society changes the web and equally how the web changes society
36	Interdisciplinary study of the relationship between people and the web
37	a dynamic set of disciplines in order to understand networked people
38	Web Science is what I'm looking at when I say 'Web Science';
39	studying the relationship of the technical infrastructure of the Web to the social interactions it permits and particularly trying to develop an understanding that can inform engineering of the Web infrastructure in a way that provides affordances for the social interactions|
40	A interdisciplinary study of the Web.
41	Web Science is about understanding how the Web changes the society and how society changes the Web

Virtually all the respondents provide an answer which implies the socio-technical nature of Web Science. Many include the term 'social-technical' or technology and society and the mutual shaping of one on the other. The fact that this is the case, suggests that the majority of participants demonstrate at least a basic understanding of the fundamental concept of Web Science. An exception to this may include respondent 14, who suggests that Web Science is merely an extension of Web technologies. They describe Web Science as "A field dedicated to understanding more about the technical and social implications of Web technologies". While this statement is not untrue about Web Science, the subject is more than the mere study of the impact of the technology, involving the study of social machines, and the mutual shaping of society on technology and vice versa. While the statement itself is not incorrect, given in isolation as the sole description for Web Science, it fails to convey the mutual shaping of technology and society.

A number of respondents, e.g. 31, 34 and 40 fail to provide any level of detail. While other responses do include the core social-technical concept of Web Science, most are fairly brief descriptions, lacking any great level of detail. This is perhaps not unsurprising given the amount of time participants may be prepared to dedicate to a survey of this kind. It may also reflect a lack of any greater level of knowledge about the nature of Web Science as a subject. Therefore, a more in-depth study in interview format would be beneficial in order to assess Web Scientists knowledge of the Web Science field in greater detail. Details of this can be found in section 8.7.

8.5.1.5 Question 1.8 If you have a background in Web Science please explain any significant contributions you have made to the discipline:

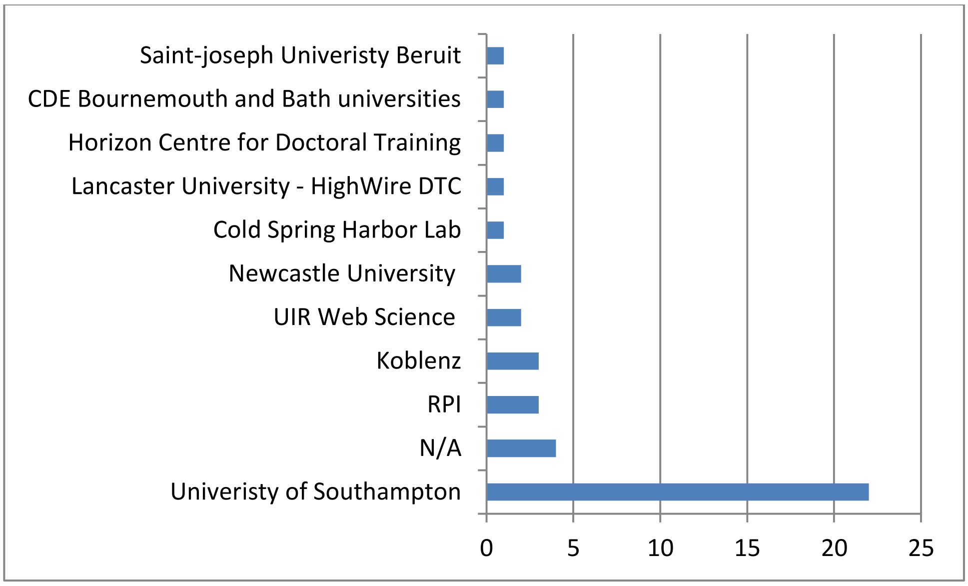
Responses to this question did not reveal anything useful contributing to the direction of research. A summary of responses mainly include the fact that respondents had either published a few papers relating to Web Science, (and in these cases, they failed to provide any further information), or that respondents were Web Science students. The only response that yielded any useful information was that of a respondent who claimed to be the former deputy director of a Web Science Research centre, but their identity and contributions cannot be revealed in order to protect their privacy.

8.5.2 Survey Section 2: Your Curriculum Details

The following section relates to gathering information about participants' academic intuitions, with the aim of identifying additional Web Science institutions omitted from the desk survey of Web Science teaching institutions.

8.5.2.1 *Question 2.1: What is the name of your academic institution or company?*

The following table shows the respondents' affiliated academic institutions.



One of the purposes of this question is to provide an answer to the research question 'Where is Web Science taught?' The purpose of discovering where Web Science is taught, is to facilitate the collection of additional curriculum data for Web Science, in order to answer the research question, 'What is taught as Web Science? Therefore, the key purpose of this question was to identify any Web Science teaching institutions which may have been missed by the desk survey of Web Science teaching institutions.

Even when survey participants provide information on which university they originate from, this is not necessarily an indication of an intuition which teaches Web Science. The survey was promoted within various Web Science Facebook groups, and was also promoted during the 2015 Web Science conference in Oxford. While the attendees of this conference included many academics and students who would class themselves as Web Scientists, conversations with participants during the conference indicated that it also included a good number of participants from other disciplines, representing the interdisciplinary nature of Web Science. Therefore, some of the survey respondents do not originate from intuitions which teach Web Science programmes.

The CDE Bournemouth and Bath Universities[31] response does not indicate the presence of a dedicated Web Science programme. The response instead indicates a collaboration between two universities via a centre for doctoral training. This does not feature Web Science specific research, and instead focuses upon a 'Centre for Digital Entertainment', hence the initials 'CDE'. While this doctoral training centre is not dedicated to Web Science, its presence in the survey results suggests that the group conducts research which has some overlap with Web Science, as the participant must have either been involved within a Web Science Facebook group or attending the Web Science conference to have received an invitation to participate in the survey.

[31] http://www.digital-entertainment.org/applicants/

Similar to the Bournemouth and Bath example, the Horizon Centre for doctoral training[32] is an example of another centre for doctoral training which conducts research which may demonstrate some overlap with Web Science. The Horizon Centre for doctoral training is described as focusing on digital economy research. This is a subject which features in Web Science, as indicated by survey respondent 7 in section 8.5.3.1, Table 38. This is an example of another research group with research dedicated to one of the 'component topics' of Web Science, but does not have a dedicated Web Science research programme. The Lancaster University Highwire CDT[33] is another example of a centre for doctoral training demonstrating some overlaps with Web Science. The Website for Highwire describes a research focus of 'digital innovation through technologies'. They also describe their programme as demonstrating the overlap between the three disciplines: 'Computing, management, and design'. This is an example of interdisciplinary research, similar and related to Web Science, but not strictly Web Science. Cold Spring Harbour Lab[34] is a private research facility which focuses on medical research. It is not an institution which would immediately be associated with Web Science, and it does not offer any form of taught programme relating to Web Science. However, its presence in the survey results illustrates the diversity of Web Science, and the fact that the Web impacts many disciplines, including medical research.

8.5.2.1.1 Conclusions from Results

While not providing additional information on specific Web Science programmes, these findings do provide information on some of the interdisciplinary areas, which are related to Web Science. This provides avenues for further research into the overlap between and Web Science related disciplines.

8.5.2.2 Question 2.2: In which part of the world is your institution located?

Figure 5 shows the world locations of survey participants. Perhaps unsurprisingly, the largest number of participants originate from Europe, with a small number from America and a smattering from the Middle East. This directly corresponds to the current results from the Desk survey of Web Science teaching institutions, which suggests that the majority of Web Science courses originate in European Universities, with a small number in America, and the University of Beirut in the Middle East. The rest of the world is seemingly under represented; however, this may be at least partly due to the language difference, as it is likely that searches missed Universities which have websites in languages other than English.

[32] https://www.horizon.ac.uk/about-the-cdt/

[33] http://highwire.lancaster.ac.uk/about/training-at-highwire/

[34] http://www.cshl.edu/

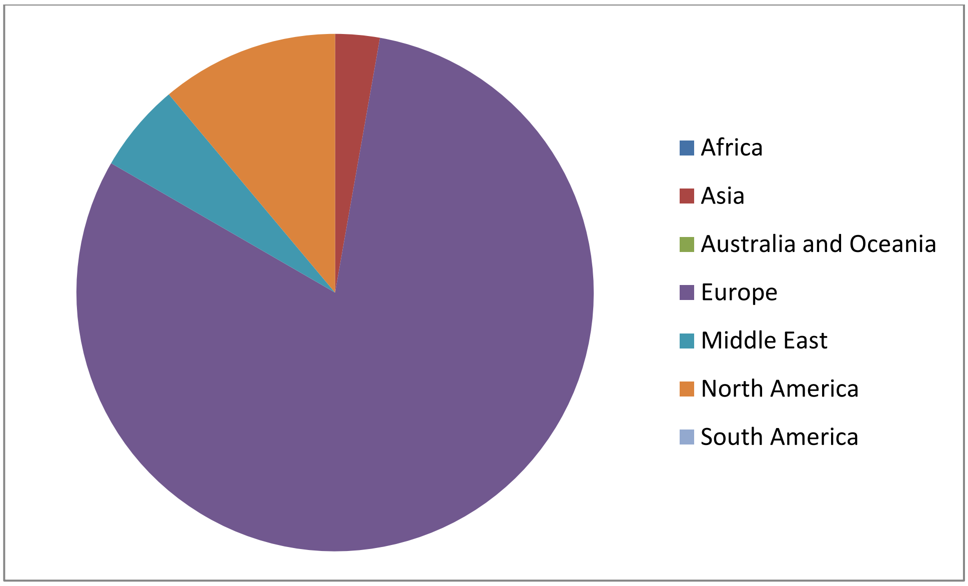


Figure 5. World locations of Participants' Academic Institutions

8.5.2.3 Question 2.3: Please give the name(s) of the Web Science (or related) taught programmes offered by your institution:

It was hoped that answers this question might reveal additional Web Science taught programmes missed by the desk survey of Web Science of Web Science teaching intuitions. However, answers to this question failed to reveal any detailed information about programmes other than those already identified. Nearly a third of respondents simply put 'Web Science', not providing any detail at all, whilst the remaining did not provide much further detail, and others listed programmes already identified. The fact that some respondents failed to provide additional detail may relate to the fact that as seen previously, 22 of the respondents were from the University of Southampton, therefore knowing that the survey was run from the same institution, they did not feel the need to provide additional detail.

8.5.2.4 Question 2.4: What level(s) of Web Science (or related) qualification are taught by your academic institution? (Please tick all that apply)

This question asks participants to select all levels of Web Science taught programme is offered by their academic institution. Responses superficially seem to indicate a fairly even split between all three levels. Superficially this is would seem to contradict the findings of the prior desk survey of Web Science teaching intuitions, as according to findings of that study, there are a far greater number of MSc programmes in existence than undergraduate examples. According to the desk survey, the number of PhD programmes are also less than the number of MSc programmes. However, when it is taken into consideration that a large number of the respondents originate from the University of Southampton, which offers all three levels of programme, this then explains the apparent imbalance in the results.

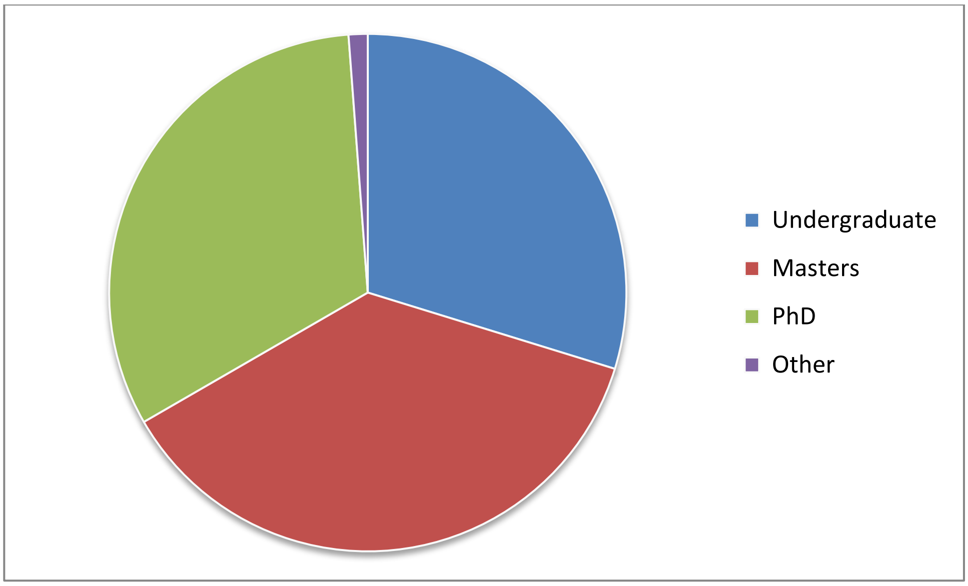


Figure 6. Levels of Web Science Programme Offered

8.5.2.5 Question 2.5: Please provide a summary of the key topics covered by your institution Web Science (or related) programme(s): (If this is a full degree programme please provide module titles if possible)

Due to a large number of respondents being based at the University of Southampton or other institutions already identified, this section yielded little new insight as to the topics covered by Web Science programmes. The following table shows the respondents answers in the combination with their answers to question 2.1, in order to show respondents' answers in the context of their institution of origin. Answers were filtered to exclude University of Southampton results, as the subjects for Southampton are already well known. Blank results were also filtered.

Table 35 Respondents' Descriptions of Web Science Modules (Excluding Southampton)

Uni of Origin	Respondents' Descriptions of Web Science Modules (Excluding Southampton)
Koblenz	Master's programme in Web Science; module names: • Introduction to Web Science, • Network Theory and Dynamic Systems, • Semantic Web, • Information Retrieval, • Web Engineering, e-Participation, • Online consumer behaviour, • Research methods, • Computational social science • Entrepreneurship, • plus further electives
UIR Web Science	• Cyberwar • web and education • web in the middle east
Lancaster University - HighWire DTC	• Qualitative Methods in Social Sciences. • Applied Innovation. • Design Driven Innovation. • Analysis and Design. • Special Topics (Lit Review on topic of our choice.
Horizon Centre for Doctoral Training	• Innovation and Technology (semi-technical over-view of how the internet works, profs' research, current concerns in digital technology, framed largely from an HCI perspective) • Emerging Technologies (near-future technologies currently in development, and use of current technologies in various institutions like the local museums; assessed via shooting short utopic and dystopic films) • Innovation and Society (basic sociology as applied to technology) • Module on business considerations • Research methods • Recommended optional modules: Computer Programming, Human Performance (basic qualitative and quantitative research), Cognitive Ergonomics
Rensselaer Polytechnic Institute	• Topics include: • Semantic Web technologies • Data Science • X-informatics
Rensselaer Polytechnic Institute	We have course at different levels ranging from web development to web science per se. The latter is fairly similar in scope to the material in http://www.nowpublishers.com/article/Details/WEB-017

Having filtered out the Southampton results, respondents still largely originated from institutions which were already included in the Web Science modules desk survey. The only exceptions include the results from 'Lancaster University - HighWire DTC' and 'Horizon Centre for Doctoral Training'. As previously described, neither of these institutions offer Web Science taught programmes. Despite not being Web Science programmes, the content of both demonstrate interdisciplinary. The Horizon example especially, bears a significant resemblance to attributes which might be found in a Web Science programme. The programme includes technological development, business, research methods and the application of technology in a societal context. The reason for the presence of the same basic structure could be explained by the fact that the programme has some foundations in Human Computer Interaction, which is an interdisciplinary subject which shares some similarities with Web Science, specifically, the socio-technical aspect of investigating the application of technology and its effects on society, (Hooper & Dix 2012).

8.5.2.6 Question 2.6: Please list the language(s) in which your programme is taught:

Responses to this question are almost universally English, only two responses differed; these included the following answers: 'English and French, and 'English French and Arabic'.

8.5.2.7 Questions 2.7 to 2.10

These questions were deigned to identify additional information relating to Web Science teaching institutions missed by the desk survey. The answers to these questions yielded a lack of useful results, especially given that most respondents originated form Southampton, and therefore the answers provided were already known.

8.5.2.8 Question 2.11: Please provide a link to the programme website: (Optional)

The following table includes links to the websites of respondents' affiliated Web Science teaching institutions. The majority of responses show Web Science intuitions already identified by the desk survey of Web Science teaching intuitions. A few responses showed non Web Science intuitions, such as a Computer Science programme offered by the University of Bath. This shows some affiliation between Web Science and Computer Science students.

Responses 3 and 20 identified Highwire DTC based at Lancaster University, the research of tis intuition is not directly identified as Web Science, it focuses upon research the Digital Economy. The identification of this intuition was not directly useful to the desk survey of modules as it does not have a structured taught programme, however, it does illustrate ties between Web Science and research into the Digital Economy, which could conceivably be a sub topic of Web Science.

Table 36 URLs of Respondents' Web Science Taught Programmes

No.	URLS of Respondents' Web Science Taught Programmes
1	http://west.uni-koblenz.de/en/studying/master-web-science
2	http://webscience.blogs.usj.edu.lb/
3	http://highwire.lancaster.ac.uk/About
4	http://www.horizon.ac.uk/
5	http://www.bath.ac.uk/study/ug/prospectus/subject/computer-science/
6	https://west.uni-koblenz.de/de/mws/aims-contents
7	http://dtc.webscience.ecs.soton.ac.uk/
8	http://www.ecs.soton.ac.uk/programmes/msc-web-science and http://www.southampton.ac.uk/webscience/undergraduate/courses/i200_bsc_web_science_social_science.page http://www.southampton.ac.uk/webscience/undergraduate/courses/i201_bsc_web_science_computer_science.pa ge
9	you have this
10	https://www.rpi.edu/dept/IT/facts_figures.html
11	I know this is being reported by others at my institution. Please refer to their responses.
12	n/a
13	http://naa.url
14	http://www.ecs.soton.ac.uk/programmes/msc-web-science#modules (this is the current So'ton course - different curriculum).
15	www.soton.ac.uk
16	http://www.rpi.edu/dept/IT/
17	http://dtc.webscience.ecs.soton.ac.uk/
18	http://west.uni-koblenz.de/en/studying/master-web-science
19	http://webscience.blogs.usj.edu.lb/
20	http://highwire.lancaster.ac.uk/About
21	http://www.horizon.ac.uk/
22	http://www.bath.ac.uk/study/ug/prospectus/subject/computer-science/
23	https://west.uni-koblenz.de/de/mws/aims-contents
24	http://dtc.webscience.ecs.soton.ac.uk/

8.5.3 Section 3: Curriculum Responses: Rating the WSSC responses

During this section, respondents were asked to rate the categories of the Web Science Subject Categorization. The numbers in indicate the number of respondents who answered in each category.

Table 37 Survey Respondents' Ratings of WSSC Categories

WSSC Category	not at all relevant	not very relevant	unsure	relevant	very relevant
Please rate the following categories relating to: WEB TECHNOLOGIES					
20.General Web Technologies	0	0	2	13	13
21.Web Milieux: Document technologies; Hypertext technologies; Internet technologies; Mobile Web technologies; Grid and Cloud computing technologies	0	2	4	17	4
22.Basic Web Architecture: HTTP and related technologies; URIs; HTML; XML; CSS and related technologies; Interfaces and Browsers; Servers Web Services	0	2	3	12	11
23.Web 2.0 technologies	0	2	2	14	9
24.Semantic Web/Linked Data: Metadata; Knowledge Representation; Ontology Languages; Linked Data; Natural Language Processing; Provenance systems in the Web; Other in Semantic Web/Linked Data	0	2	3	13	9
25.Internet/Web of Things	0	2	2	17	7
Please rate the following categories relating to: WEB ANALYSIS					
26.General Web Analysis	0	2	3	11	10
27.Mathematical Methods of Web analysis: Web data sampling and analytics; Logic and Inference in the Web; Statistical Inference in the Web; Statistical Analysis of the Web; Web as a Complex System; Graphs; Networks; Mathematical methods for describing Web services; Crawling; Indexing and Searching; Data Mining; Information Retrieval and Machine Learning; Other Algorithms for the Web	0	4	2	12	8
Please rate the following categories relating to: WEB SOCIETY					
28.Economics and Business: Goods in the Web; The Web economy; Antitrust Issues and Policies in the Web; Intellectual property and digital rights management; Web-based economic development	0	4	4	8	13
29.Economics and Business (Business): E-commerce Business models in the Web; Advertising in the Web; sponsored search	0	2	4	12	10
30.Social Engagement and Social Science: Social networks; Mass phenomena; Collective intelligence; Peer production; Globalization; Systems; Social structures and processes; Virtual communities	0	2	1	11	14
31.Personal Engagement and Psychology: System Psychology and Behaviour; Child and adolescent psychiatry; Tele-working	2	7	4	9	7
3**2.Philosophy:** Philosophy of information; Objects; Reference and Cognition in the Web; Ethics in the Web	0	6	6	8	8
33.Law:Intellectual Property in the Web; Digital Rights Management; Digital crime; Laws for Web access; Antitrust Law	1	2	4	14	7
34.Politics and Governance: Political science**;** E-Government; E-Politics; E-Democracy; Policy and Regulation; Web Governance; Privacy; Trust; Security ;Network neutrality; E-Inclusion	0	2	4	14	9

The results of section 3 of the survey suggest that the majority of respondents rate the topics of the WSSC as 'relevant', with a slightly lesser number rating the topics as 'very relevant'.

Given the results of this section, findings of the survey would largely seem to indicate that Web Science academics and students involved in the survey, largely agree with the topics included in the WSSC. However, responses to the following question (3.4) also suggest that it fails to cover all areas of Web Science. Therefore, while it serves as a useful reference for topics which are related to Web Science, there are topics which is does not cover, indicating that improvement and expansion is needed.

Only two categories, 'Personal Engagement and Psychology' and 'Law: Intellectual Property in the Web' received rating which indicate that participants strongly disagree with the category being relevant, and this was only in the case of a very small number of respondents. In these instances, it would be interesting to have a justification for why respondents answered as they did, although time limitations of such a survey restrict the number of questions which can be asked. One of the answers to the following question (3.4) suggests that one respondent answered as 'unsure' due to the fact that their Web Science taught programme did not include certain social science based subjects, so this could potentially be a reason for the low ratings.

It is slightly surprising that the Personal Engagement and Psychology categories were the lowest rated examples, given that Psychology was one of the more highly coded subjects identified from the analysis of Web Science conference publications. Although not in the top 15, the subject 'psychology' was identified in a total of 84 publications, while 'user behaviour' was the top most identified subject out of the paper topics identified, and was coded in a total of 145 papers. A good number of papers in the review of Web Science conference publications related to the behaviour and motivations of Web users, and therefore 'Psychology'. Therefore it seems strange that this would be the worst rated WSSC topic, with two 'not at all relevant' ratings and less higher level ratings. The low survey rating for the WSSC topic may be attributed to the more specific nature of the WSSC sub topics: 'System Psychology and Behaviour; Child and adolescent psychiatry; Tele-working'. These may be too specific and therefore not reflect general Web user behaviour and motivation.

8.5.3.1 *Question 3.4: Do you feel that the above framework of the WSSC omits any key subjects which relate to Web Science or your individual curriculum? If yes please list subjects and why you feel they are relevant:*

Respondents were asked to list any subjects that they felt were omitted by the WSSC. Results suggest that this should have potentially been a compulsory question, as the results of this question are particularly interesting to this thesis, identifying Web Science 'component subjects'; yet not all participants choose to answer.

Table 38 Subjects omitted by the WSSC

No.	Subjects omitted by the WSSC
1	Education
2	Web Entrepreneuership - companies like Facebook Twitter etc. shape the Web||
3	No
4	no
5	I have no idea what webscience is. I have only heard of it from people who are involved with it at Southampton University but I have never got a description of what it entails and don't really feel I can answer most of the questions about it here.
6	Not sure
7	Digital anthropology (different from digital sociology!); ethnographic methods; ethnomethodology; and (broadly critical social theory and philosophy; though I think a lot of that has to do with the tendency of programmes like these to originate in the concerns of the sciences more than the humanities / social sciences. There's a lot that web science (or digital economy or Internet science etc...) COULD be. I'd like to see more of an underlying incorporation of other critical traditions (rather than just the selective smattering I more often see) though the next best thing is to be as upfront and clear about curriculums'; underlying concerns traditions and epistemologies as possible.
8	No
9	seems pretty comprehensive
10	Interdisciplinary - because the web can hardly be researched from one discipline alone|Computer Science Basics - all the technology mentioned above doesn't mean anything for me really. Basics are required to bring scholars from various disciplines at least to a stage where they can discuss technical aspects of the web.
11	Anthropology
12	Yes - Many sociological issues seem to be missing. For example the issue of the digital divide and power structures on the Web. Associated applications (such as language-related issues on the Web and the dominance of English) are also missing.
13	No
14	NO
15	No
16	I think it would be useful to include research and training on how to apply web science as part of a business intelligence function.
17	No
18	Education!
19	WSSC is cool but you guys stopped following the butterfly diagram. Anyways in the above web society its all optional and we don't have much offered ..so I put up not sure for some
20	Information management curation of knowledge. There is an unsupported assumption that finding and organising things is a natural product of more/better computing power/code. The human element in the triage and curation of information is overlooked in favour of more process. A thematic problem with teaching is that it was situated in an in-discipline way with little elucidation as to *why* the [topic] is relevant to WebSci. My course has no data science at all. There was little attempt to connect the humanities approach with the engineering; rather it was posited as a way to "not let the engineers take charge". The downside effect is few of my cohort appeared to move beyond their 'parent' discipline. Coming to study after a career in Forces/business I was surprised how much time was wasted on issues of process and discipline as opposed to achieving actual outcome.
21	No
22	Entrepreneurship Marketing Innovation Maths Education
23	No

Respondent 2 suggests that entrepreneurship in the context of the Web is a key topic which is omitted from the WSSC. They state that companies such as Twitter are central in the development of the Web, and that consequently, their development should be studied.

Respondent 5 admits that they possess little understanding of the nature of Web Science as a subject. This raises questions over the target audience of the survey, and why this participant decided to contribute to the survey. The survey was promoted within dedicated Web Science Facebook groups, and also during the 2015 Web Science conference, which was assumed to be the one of largest gatherings of Web Scientists from across the globe, and therefore one of the best opportunities to promote a survey such as this. However, the conference would have included some participants from disciplines outside Web Science, as indicated from the earlier results of question 9. The response of this participant may simply highlight the need for greater promotion of Web Science as a subject, in order to raise awareness of an important field of study.

Respondent 7 highlights a number of topics which they feel should be better represented by the WSSC. Two of the examples they cite, ‘ethnography’ and ‘philosophy’, accord with results of the analyses of Web Science taught modules Table 27 and Web Science conference proceedings Table 30, as these are two topics which were identified as part of these analyses. However, respondent 7 also highlights two additional subjects, Digital Anthropology, and Ethnomethodology, which they cite as relevant to Web Science. They justify their recommendations for these subjects by explaining that they feel that Web Science is biased towards the opinions and methods of the sciences, as opposed to the humanities and social sciences. They explain that they would like to see greater incorporation of interdisciplinary methods and perspectives from social sciences; as opposed to what they consider to be the current, very selective sample of methods from the social sciences included in their Web Science taught programme.

Respondent 10 is another example of a participant who, from their answer, would seem to possess limited knowledge of Web Science as a subject. They suggest an ‘interdisciplinary perspective’, and then appear to slightly contradict themselves, by stressing the importance of having all Web Scientists able to discuss the Web with at least a basic knowledge of Web technology. While this is an aspect of Web Science, it is not the only focus and lacks any suggestion of dual importance of both the socio and the technical.

Respondent 11 simply states the subject ‘Anthropology’ with no justification for their recommendation. Respondent 7 also suggested ‘Anthropology’, in the form of ‘Digital Anthropology’. Anthropology is the study of human society, culture and development. Given the prominence of the Web in modern society, it is very conceivable that Web Science and Anthropology would demonstrate some overlap.

Respondent 12 highlights the lack of representation of sociological issues such as the ‘digital divide’, suggesting that there is little investigation of the various languages spoken on the Web, and the fact that English is the dominating language. During the analysis of Web Science conference literature, there were only a few examples of research into the digital divide and languages spoken on the Web. Given that this is likely to be a huge issue; this is somewhere surprising. Respondent 16 appears to be another example of a participant who has answered in the context of their Web

Science taught programme, as opposed to Web Science as a subject in general. They suggest that they would like to see more teaching on Web Science in relation to 'business intelligence function'. While this refers to specific training opportunities desired by the respondent, it may also indicate the suggest that the topic of 'business' should be a category which is covered by the proposed Web Science Lexicon, and this is also supported by the results of the analysis of Web Science taught modules. Respondent 19 also appears to have answered with the assumption that they are rating only areas which are featured in their own taught curriculum, as opposed to which subject should represent the wider Web Science subject. Consequently, they only rated Web Society topics as 'unsure'.

Respondent 20 is another example of a participant answering the question in the context of their own Web Science curriculum. However, in contrast to respondent 19, they provide some useful insight into what they feel are some flaws of Web Science teaching. Respondent 20 explains that within the modules of their Web Science programme, there is a lack of 'the bigger picture' of Web Science provided. They explain that individual subjects are taught in isolation, with little reference to the wider context of Web Science. This corresponds with the findings of the analysis of Web Science conference publications, which seems to indicate that during latter years of the conference, there is a lack of reference to the wider Web Science subject within research published under the Web Science conference.

Respondent 20 suggests that 'knowledge management' is something that was lacking from their Web Science programme. They describe the fact that within their programme, the management of knowledge was assumed to originate from proficiency at coding, neglecting the need for the human element of knowledge management. They then explain that there was very little link between the social and the technical modules. Whilst a key feature of Web Science should be the interplay between the two, this was lacking from the teaching of the modules, with very little communication and links provided between the social and the technical modules. This is a likely consequence of the fact that economic constraints mean that Web Science programmes often consist of modules drawn from other programmes. For example, Web Science students will join Computer Science and Social Science cohorts for lectures, taught from each of these perspectives. While this is likely to be necessary for the provision of Web Science programmes financially, the response of participant 20 suggests that more needs to be done to provide context between interdisciplinary subjects within Web Science teaching. This also echoes the sentiments of respondent 7, who also suggested that there was a lack of interdisciplinary perspectives within Web Science teaching.

Respondent 22 recommends a number of subjects which they feel should be included in the WSSC, these include: 'Entrepreneurship Marketing Innovation Maths Education'. No punctuation was included in the answer; however, it can be assumed that the respondent intends to describe 5 separate subjects which they feel are omitted from the WSSC. This is not accompanied by any detailed justification, however, entrepreneurship and education correspond with the responses of other participants. Marketing also corresponds with the results of the analysis of Web Science taught modules (Table 27).

8.5.3.1.1 Conclusions drawn from the results of Question 3.4

To conclude, overall themes which emerge from the analysis of the results for this question include the following. There is a general consensus between all participants, that WSSC is missing social science subjects, and that it is more biased towards the technical aspect of Web Science.

Many of the respondents answered with the apparent preconception that they were being asked which subjects were not represented by their own Web Science taught programme, as opposed to which subjects are underrepresented in the WSSC topics and Web Science as a subject in general. This suggest the possibility that wording of the question may have not been clear, although the subsequent answers still provide valuable insight into participants' view of what they feel should be included in Web Science taught programmes. It is also natural that many participant's views of Web Science are heavily influenced by their Web Science programme, as in many cases this will be participants' only experience of Web Science as a subject.

It is suggested according to the experiences of two respondents, that Web Science teaching lacks interdisciplinary context between modules. Web Science subject recommendations made by participants will be considered along with the results of interviews with Web Science students and compared with the results of the analysis of Web Science taught curricula and Web Science conference material, in order to devise the content and structure of the proposed Web Science Lexicon.

8.5.3.2 Question 3.5: Are you aware of any differences between the terms Web Science and Internet Science? If yes please elaborate:

The following table illustrates respondents' understanding of the differences between Web Science and Internet Science. (Please note that some of the responses to this question were incomplete, but they have still been included. Respondent 15 appears to be 'trolling' and cannot be considered seriously.)

Table 39 Differences between Web Science and Internet Science

No.	Difference Between Web Science and Internet Science
1	Similar but Internet science slightly more focused on the architectural underpinnings of the web along with network science.
2	I would define "internet science" to be mostly restricted to the level of networking and questions that arise for solving networking issues. Some of these issues will have to do with Applications but these applications will not be constitutive for internet science per se.|Web science rather takes networking as a given and asks questions about the ecology of applications living on top of the internet how people use them which patterns emerge - the set of applications constitute the Web.
3	Yes arguments made in terms of the IP stack differences in the make-up of the two communities (see the EINS 2015 paper on this by Hooper Neves Bordea). Also look out for Thanassis's CACM article out next month.
4	Internet Science studies Internet Internet Networks and society and Web Science studies larger sociotechnical systems in the web environments.
5	No
6	not really...
7	No
8	Focus on social versus technical (internet)
9	I believe there is a difference but I do not know what that difference is.
10	I thought that Internet Science focuses more on the technical infrastructure which underlies the Web.
11	Internet Science is more inclusive of international politics and governance hardware and system designs.
12	I'm not familiar with the term "Internet Science."
13	Internet science is more about the technical functionalities. web science is about the impacts.
14	Web Science: studies how people interact in the Web and how it affects different domains in society. |Internet Science: study of the technologies that support Internet and the Web and how people deal with such technologies.
15	ofcourse dont ask the difference between a connected network and world wide web to a web science student :P|Congrats and all the best for your research its quite a good questionnaire . I will read your PHD thesis :)
16	No
17	Yes Internet Science is studying the underlying technologies used to provide the web such as IPv6 multicast DNS routing etc.
18	I always think of Internet science to refer to more topics pertaining more to the hardware or to a more computer science perspective on web science. internet seems to be used colloquially to mean Web so it is hard to know what people really mean
19	-
20	No
21	As far as I can tell they mean exactly the same thing. The
22	Web is logical level Internet is physical connections.
23	In the US 'internet Science'; is virtually unheard of and Web Science itself is pretty rare.
24	Yes but I'm not sure there is a large difference. It seems more like Internet Science is closer to computer science and looks at the internet layer more than just the web.

Responses appear to be split evenly between those who understand the nature of the differences between the two subjects, and those who do not. Some of the respondents who are not aware of the differences simply answer 'no'. However, some provide an incorrect answer along the lines of 'they are the same thing'; for example, respondent 21. Respondent 15 appears to be answering with a level of sarcasm, therefore their response cannot be considered seriously.

Those who do provide a ‘correct’ justification for their answers according to the description of differences provided by (Tiropanis et al. 2015), generally demonstrate consensus along the lines of Internet Science involving the study of the technologies which underpin the Web, whereas Web Science focuses of the socio-technical impact of the Web. Respondent 3 actually refers to the (Hooper et al. 2015) and (Tiropanis et al. 2015) papers on Internet Science, and the differences between Web Science, Internet Science and Network Science.

Respondent 2 describes Internet Science as involving the study of Networks. While this is true to an extent; as seen previously in chapter 4, Network Science is a subject which is dedicated to the study of Networks. They also describe Web Science as the study of the applications which run on the Web and their impact on users. While this is also true, the description is still somewhat limiting the scope of Web Science to the mere study of Web Applications.

Respondent 14 makes the differentiation between Web Science involving the study of the impact of the Web in Society, and Internet Science involving the study of Web Technologies. While this differentiation is partly supported by (Tiropanis et al. 2015), it is important not to overlook the socio-technical nature of Web Science when describing the subject; as Web Science involves both the study of both technology and society, and the interaction between the two.

8.6 Main Conclusions from the Survey

The survey failed to highlight any Web Science taught programmes which had not already been identified during the desk survey of Web Science teaching institutions. However, the survey did provide insight into how participants describe Web Science. While the majority of respondents did demonstrate at least a superficial understanding of the interdisciplinary and socio-technical nature of Web Science, very few of them were able to provide detailed explanations. This may be partly due to the fact that participants are prepared to devote limited time to a survey such as this, and therefore this will be addressed with the proposed interviews of Web Science PhD students. Another reason for the lack of detail could be attributed to participants’ lack of knowledge regarding the scope of Web Science, and if so this reinforces the need for greater awareness about the scope and nature of the subject, something which the proposed Web Science Lexicon aims to address.

Participants were asked to rate the relevance of the WSSC topics, and the results of this section of the survey will prove useful when determining structure of the final Web Science Lexicon. The outcome of the ratings suggested that despite the perception that the WSSC is outdated, the majority of participants still considered the topics relevant. However, it was suggested that the WSSC was missing a number of subjects, and several suggestions were made, which can be considered when structuring the Lexicon.

8.7 Web Science PhD Researcher Interviews

The purpose of these interviews was to build on the findings of the previous survey of Web Science academics and students, providing more detailed insight into the specific subjects being researched by Web Scientists in practice. The data was gathered in the form of one to one interviews, conducted in person, where the interviewer used an online iSurvey instrument as a basis for questions, and also to record the information provided by the participant. There was one instance where an interview was conducted remotely via Skype.

All participants interviewed were PhD students at the University of Southampton. A common element shared by all participants, was the fact that each had completed the Web Science MSc taught at the University of Southampton. All participants had also progressed from this to study Web Science related PhDs at the university of Southampton. Participants were at varying stages of their PhDs, some having only just completed the Web Science MSc, while others were more advanced and had been studying PhD's for 3 years or more.

The first dozen Interviews were conducted as part of the University of Southampton's department of Web and Internet Science annual research week, known as 'WAIS Fest', with several additional interviews being arranged shortly afterwards. The project was promoted to participants via a pitch at the start of the event, and also by e-mail, with participants being sent a doodle pool of slots from which they could choose availability for meeting slots.

A total of 16 people agreed to interviews during and following the period of the WAIS Fest research week. While this is a relatively small sample of Web Scientists, it still offers some insight into what subjects are researched by Web Scientists, as well as their research methods utilised. The amount of information gleaned from the interviews is not enough on which to base a detailed analysis of Web Science constituent subjects, such as that conducted during the analysis of Web Science taught modules and Web Science conference publications. However, it is a useful sample of how Web Science is studied in practice, and will provide enough information to serve as a comparison with the two afore mentioned analyses.

The data recorded from the interviews will be analysed in Nvivo, in order to identify the key topics and paper topics most frequently corresponding with the research topics detailed by the interviewees. This analysis will also seek to identify any significant Web Science component subjects not previously identified by the analysis of Web Science taught module and Web Science conference material.

The first three survey questions relate to the identities of participants, and are therefore not included, in order to protect the identify of participants.

8.7.1 What is your Research Project Title?

The following table outlines the research project titles of the participants.

Table 40. Interviewee Research Project Titles

No.	Research Project Titles
1	Measuring online/offline community
2	Demand and innovation for open data
3	Longitudinal maintenance of hypertext
4	use of the internet in decision making processes in the green party
5	Value values exchange process on the web
6	The public health analogy in web security
7	The impact of community and sociability
8	Evaluating the mobile web accessibility of electronic text for print impaired user in higher education
9	Socio-technical construction of MOOCS and educator and learning designer roles and practices in higher education
10	What do teachers talk about on social media
11	Looking at ways to improve the information spread and warnings natural disasters using social media
12	Perceptions and use of social media and support of non-formal learning by secondary school pupils
13	Whether there's a technological solution to adults with basic literacy skills
14	Archiving social media
15	Exploring emergency responders use of social media
16	Looking at the evolution of web technology

8.7.2 What keywords would you use to describe your research project?

If you were identifying your academic subject area or discipline/field of study to someone what words would you use? E.g. criminology, education, law

Table 41. Interviewee Project Keywords

No.	Research Project Keywords
1	Online/offline community social networks community studies interdisciplinary studies |||research model - called Spence interdisciplinary construct ||
2	Open data demand innovation digital ethnography open source community collaboration platforms tools
3	Relevance accessibility collaboration anonymous access hypertext
4	Decision making online participation democracy political parties
5	Human action worth ethical values values exchange space alienation ethics economics psychology epistemic issues sociology social construction
6	Regulation security economics simulation
7	Computer supported collaborative work online communities crowdsourcing HCI
8	Metadata user agents accessibility requirements tuples etexts
9	MOOCs socio-technical interaction network roles and practices learning designers
10	Topic modelling data mining blog sphere ethnography social networks network graphs
11	Social media reliability accessibility disaster management
12	Social media non-formal learning constructivist technology TEL - technology enhanced learning digital literacy social capital
13	Inclusion digital equality digital, inclusion exclusion accessibility adult literacy
14	Archiving social media legal issues
15	Emergency management emergency responders social media UK floods twitter
16	interdisciplinary empirical theatrically grounded web technology evolution

8.7.2.1.1 Mapping Interviewee keywords to key and paper topics

Most of the project keywords listed by participants can be easily summarised with one or more predefined key topics or paper topics. In the case of some examples, the wording was changed slightly, but in most cases, the prior identified key and paper topics provided a good match for the participants own words used to describe their projects.

Translating participant keywords to predefined key topics and paper topics:

1. offline community and social networking
2. Open Data, Digital Ethnography
3. Accessibility, Hypertext
4. Online engagement, Politics
5. Ethics, economics, psychology, sociology,
6. Law, Web Governance, Security, economics
7. Virtual Communities, Crowdsourcing, HCI
8. Accessibility,
9. MOOCs, Networks, e-learning
10. Topic Modelling, Data Mining, Digital Ethnography, Social Networking, Networks, Web Graph,
11. Social Networking, Accessibility, Disaster Management,
12. Social Networking, e-learning, accessibility, Digital Literacy,
13. Accessibility, Digital literacy,
14. Social Networking, Law,
15. Disaster Management, Social Networking, Microblogging,
16. Research Methodology, Web Technologies,

The following topics studied by participants represented topics which were not identified during the prior analyses of Web Science modules and Web Science conference publications. However, they represent topics which are being studied by Web Scientists in practice, and therefore warrant further investigation, in order to determine if they should be represented within the Web Science Lexicon

Technology enhanced learning - although not identified as a topic in its own right, this could conceivably be represented by the related topics: e-learning, accessibility and HCI.

Citizen Science – this can be represented by the higher level topic, Social Science. However, this is a very generalised representation.

Digital equality and digital literacy – these topics were not frequently coded during the paper analysis, yet these are both significant issues which is surely frequently researched as part of Web Science. They are potentially related to pre-existing topics, Ethics, and Accessibility.

Entrepreneurship – was identified as a key topic taught from the Web Science programmes, but was only coded once in the analysis of Web Science conference publications. However, a survey participant cited that Web Entrepreneurship is an interesting topic, which should feature more prominently in Web Science taught material, due to the entrepreneurial nature of Web enterprises such as Google and Facebook.

8.7.3 Please briefly describe the subject area of your research:

This question asked participants to provide more detail about the subject area of their research. The following data, along with the keyword and research question data was coded in an Nvivo analysis in order to identify the Web Science constituent subjects studied by Web Scientists in practice. The results of this analysis are presented in Table 45.

Table 42. Interviewees' Description of Research Areas

No.	Interviewees' Description of Research Area
1	Looking into online offline communities a new social reality provide an instrument to measure the new reality of online/offline ||also provide evidence to show that social change has occurred |||developed an instrument enables conventional surveying and pseudo surveying and analysis of online social communities and social media ||common spine of metrics which link to two survey approaches |||two areas of London small local areas case studies of these areas use of social media||examine their Twitter activity|||
2	looking at open data gap in the research lack of research as to how people are using it ||process happening ||open source platforms -
3	looking at maintaining the relevance of knowledge based hypertext over time now that we have it comes from having it long enough||ownership issues ||maintain systems that people can use ||arrive at something without having to ask ||need of the curator maintainer |of hypertext ||assumed ownership |authorship of online content |||
4	looking at how the green party uses the internet in their design making processes equality. ||Computer science part - looking at what tools are used and how ||sociological aspect - how can you improve participation in these tools
5	interdisciplinary - value exchange can be economical, the web is a trading space not only bits and bytes its about behaviour and knowledge ||people may choose to do things that work for them not only for the value but for their cultural ethical values - aim to understand human action and decisions that we make on the web are guided by values. do things we do on the web have worth? Network effects of people actions on the web and the implications. Effect of the web culture on peoples values.
6	I looked at the intersection between computer science and law - the idea that criminals only just make enough money to be profitable by treating security like public health it would be possible to annihilate cyber attacks
7	Looking at the tools and platforms that exist for crowdsourcing online citizen science platforms||research output and accuracy of results
8	Looking at the mobile web accessibility of mobile web text for print impaired users. Users might be blind or have cognitive impairments - dyslexia or physical impairments such as arthritis
9	looking at how the co-construction of mooc and educator design and practices and roles producing mooc sin higher education how are the learning designers and educators involves with them and how they shapes the moocs and how the moocs shape the practices of the learning designers
10	I'm interested in what a group of teaching professional people connected with the profession of teaching - what do they talk about in connection with professional issues on the blogosphere and twitter - particularly interested with the changes in discussion around the current government from when they were first elected in 2010. ||
11	in the past decade social media is increasingly being used by governments to warn about natural disasters at the moment its not reliable suggesting ways to improve this an standardise
12	The sharing of knowledge aspect of the constructivist theory of learning supported through social media
13	my background is in teaching language/literacy - people who have missed out on learning opportunities ||people with fewer opportunities don't access the web as much or gain as much from it ||how are these people supported to get on the web - would it be possible to use technologies applied to disabled people - can these be used to help - |sociology as well as technology |web access
14	archiving social media how organizations store the information the legal aspects the technology behind it
15	web science is interdisciplinary - trying to look at it using the technical skills. How the emergency responders are using social media to understand how and why - what impact does the use of social media have on emergency responders
16	Background is in biological anthropology coming at this from an evolutionary perspective that tends to be traditionally used by evolutionary anthropologists. Also looking to combine the computer science and technical aspects that I am learning about as part of web science. ||I am trying to investigate the evolution of technology using an empirical investigation ||lifespan of the web - analyse it using biological archaeological techniques - biological anthropology - has tools that would be applicable for investigating the development of the web as an entity.||web science can offer a lot to traditional evolutionary scientists

8.7.4 What are your research questions?

Participants were asked to describe their research questions in order to provide additional detailed insight into their chosen area of study.

Table 43Interviewees' Research Questions

No.	Interviewees' Research Questions
1	There's a new social reality online/offline communities ||Online/offline community is more than the sum of its parts
2	critical success factors for usage/ how can we see demand for open data ||key aspects of open source that are missing for open data that if they existed for open data greater level of innovation
3	I don't have a research question - articulating the role of the skills required ||ownership and authorship |
4	what party officials and members want from decision making processes how to people see democracy?
5	What are the personal values that people take into account in order to act on the web?
6	how can public health be reimagined given the current threat landscape, which stake holder is best to target in order to minimise the damage caused by drive by downloads are intermediary obligations to combat drive by downloads appropriate can actions by a single country or group of countries have a significant effect on the worldwide prevalence of infections from drive-by downloads
7	what is success in citizen science |what kind of forms does sociability and interaction take in this sort of project |how do these types of interaction influence or fail to influence the success criteria metrics that arise |
8	What are the functional categories that will support a framework that achieves the title ||how can mobile web accessibility be evaluated in the context of the provision of academic e-texts
9	what are the socio-technical networks in which learning designers are involved when they are designer moocs ||how to educators and learning designer perceive the impact of moocs on their practices ||what are the roles of educators and learning designers in mooc development
10	What are the common topics of conversation and how have they changed over time and is there any evidence of impact on ofstead/department of education.
11	how is disaster social media currently being used and by whom? ||what are the limitations associated with that? ||Can analysis of case studies be used to improve policy and standardisation
12	how is social media perceived and used by school children (11 to 18)?||How do stakeholder at the operational level perceive pupil use of social media? ||can social media be used to support non-formal learning?
13	how does the way the society conceptualises literacy impact on the digital inequality of adults with basic literacy skills/disabilities?
14	why would organisations |British library congress store information? |what are their policies on data storage?
15	How are the responders using social media why are the responders using social media who are their followers and what impact does social media bring to emergency management
16	is it possible to find evolutionary transitions in web technology?||we hear a lot about how the web can be defined by web 1.0 web 2.0.....there a need for a more detailed empirical approach

As anticipated from the keywords listed by participants, the research subjects and research questions described by participants represent a broad range of socio-technical constituent areas of Web Science. A common element of many of the research projects described, is the fact that the projects described are generally applied subjects, which represent the application of technology in order to provide solutions which benefit society. This is consistent with the definition of Web Science, according to (Berners-Lee, Weitzner, et al. 2006) and (Halford et al. 2010).

8.7.5 How would you (briefly) describe the subject 'Web Science'?

Many of the descriptions of Web Science provided by the prior survey of Web Science Students and academics were fairly brief and lacked any detail. Therefore, the same question was posed to the interviewees in order to ascertain whether the lack of detail in answers was just down to the brief time allowed by participants when contributing to a survey, or if participants feel unable to provide a more detailed description of Web Science due to lack of knowledge about the subject.

Table 44 Interviewees descriptions of Web Science

No.	Interviewees' Descriptions of Web Science
1	interdisciplinary continuum social science comp sci being either end of the spectrum ||
2	standard - looking at the web as a social technical construct using this as a basis for explaining its various phenomena||very much at the heart of what I'm doing - using organisation business constructs - how do you get people and tools/ platforms together to create innovation
3	struggle with a definition - its about how the web functions ||its about the engineer and the end user - who makes the web and who uses it |
4	how society influences the web and how does the web influence society
5	web science is an experiment because the proposal is really interesting its about an interdisciplinary approach - incorporates many disciplines but the problem is the research method implies to focus in only one direction - so the interdisciplinary approach is just a buffer - when people conduct focused research they only specialise in one area - they loose the interdisciplinary.
6	its not a single subject - its an imagination of many different subjects - the theme of the web is many subjects - the web is a thing and web science is not a subject or a discipline on its own - for me its a part of law and computer science - its an application of various disciplines
7	The socio technical study of the effect of the web on people and people on the web can't have one without the other and all the tools and processes involved
8	Co-constitution the web informs society and society informs the construction of the web - interdisciplinary
9	Exploring the interaction between technical aspects of the web and the social context in which its used
10	To me web science is inter - or multi-disciplinary, and it combines methods from one discipline e.g. sociology and applies that often using methods from computer science to e.g. individuals behaviour on the web - or the way the web has influenced human behaviour or vice versa.
11	The application of understanding how social interaction happens on the web how the web is shaped by social interactions and how social interactions are shaped by the web – 'a socio-technical construct'||
12	It's a multidisciplinary approach to understanding the many aspects of technology and in particular the web and society. the is a very varied field so covers a very wide range of topics. ||What we are trying do is understand the web so that in 20 years we have the web we want.
13	It's the study of the web in society - the web's relationship with society||there is and should be an emphasis on trying to understand the web in order to develop it to be the best it can be ||the web is for everyone!
14	The study of how the web shapes society and how society shapes the web
15	We look at how the web impacts society and how society impacts the web - not just how the web shapes us but how we shape the web. ||
16	its an interdisciplinary field that's looking to gather as many disciplines as possible to investigate the social and technical processes that shape the web ||helping to shape what the web becomes - i think perhaps its a good thing that we're not constrained - we can shape the web....

The descriptions of Web Science provided by the participants suggested that all participants had a grasp of the fundamental sociotechnical nature of Web Science. Similar to the results of the survey of Web Science students and academics, nearly all of the explanations included the terms 'interdisciplinary' and 'social-technical', emphasising the interdisciplinary nature of the subject, as previously outlined by (Berners-Lee, Weitzner, et al. 2006) and (Halford et al. 2010).

However, despite all participants having completed the University of Southampton's Web Science MSc, some still admitted to struggling to provide answers to this question. While they were aware of the socio-technical nature of Web Science, a good number of participants still failed or struggled to provide any greater level of detail to their answers. This highlights the potential benefits of a reference resource which communicates the scope of Web Science.

Participant 5 suggests that Web Scientists lose their interdisciplinary approach, focusing their research on subjects from only one perspective, either socio- or technical. This reflects and may be a consequence of the fact that some survey participants felt that there was a lack of interdisciplinary focus between Web Science taught modules, with modules comprising of either social or technical content, but lacking a combination of both approaches necessary for Web Science.

8.7.6 Coding the Interview Material

Table 45 shows the top 15 most frequently coded key topics and paper topics identified within the interview material. This includes coding of data from question(s) 8.7.1 to 8.7.4, in which participants describe their chosen subject of study within Web Science.

Table 45 Top 15 most frequently coded key topics and paper topics coded from interview material.

Key Topics Coded	Freq.	Paper Topics Coded	Freq.
Social Networking	19	Disaster Response	7
Networks	9	Online Offline Community	6
Accessibility	9	Open Data	5
Web Society	7	Digital literacy	5
Law	7	Government	4
Security	6	Microblogging	4
Web Technologies	5	User Behaviour	4
Hypertext	5	Computer Science	4
Education	5	Web Archiving	4
Economics	4	Social Science	3
Mobile Web	4	MOOCs	3
NetworkTheory	3	Virtual Community	2
Biology	3	Health	2
Big_Data	3	Online Engagement	2
HCI Human Computer Interaction	3	Crowdsourcing	2

The results of the Nvivio analysis suggests that within the interviews, there is a fairly even spread between social and technical topics identified. Social Networking is the most commonly identified key topic amongst the interviews, which corresponds with it being the third most frequently occurring topic from the prior analysis of Web Science modules and papers. This is unsurprising, given that many of those interviewed were utilising social networking data. In a number of cases the studies specifically related to use of social networking, for example the 'use of social networking by political parties'. The extensive use of social networking sites, consequently means that it is the biggest data source for monitoring user behaviour online, and provides scope for measuring people's interactions and behaviour which was previously impossible. Therefore, it is unsurprising that it always features prominently, however, has it become so prominent, that it is the most dominant single subject within Web Science? Has Web Science become the science of social networking?

'Disaster response' also featured fairly prominently, two of the studies out of the 16 interviews conducted, focused on use of the Web for disaster response, which is a fairly large percentage given the size of the sample.

8.8 Conclusions/Future Work

This chapter has outlined the process of analysing the data sources collated within this thesis, and has also revealed an initial presentation of results. The chapter documents the stage of analysis which can be described as the Grounded Theory process of 'Open Coding' and 'Axial Coding', which, in the context of this thesis, involves the identification of key topics and paper topics, described by (Strauss & Corbin 1998) as themes. The survey of Web Science academics and students also provided supporting information, which may inform the creation of the Web Science Lexicon

The resulting key topics and paper topics identified during the analyses outlined in this chapter, will be refined utilising a process of further Axial and Selective coding, in order to structure and refine the topics identified, for inclusion within the Web Science Lexicon. The following chapter details the process of structuring the proposed Web Science Lexicon.

9 Structuring the Lexicon

9.1 Introduction

The term 'Lexicon' was chosen to denote the proposed taxonomy of key topics representing the scope of Web Science, as the definition of Lexicon in English is '*branch of knowledge*'[35], and in Greek, it refers to '*book of words'.* The following chapter outlines the process of structuring the Web Science Lexicon, including providing an account of the WAIS Fest 2017 Lexicon project (9.2), in which ideas for the Lexicon structure were refined. The process of structuring the topics for inclusion in the Web Science Lexicon can be attributed to the stages of Axial and Selective coding. The themes or 'topics' identified during the analyses in section 8, are refined, and a brief description is provided of each.

In their publication, Visualising Knowledge Domains, (Börner et al. 2003) suggest that painting the 'big picture' of a multidisciplinary field is challenging, stating that: "*When it comes to a multidisciplinary field of study, it is rather difficult to maintain an overview of what is going on*." This is very true of Web Science. Although not writing about Web Science directly, (Börner et al. 2003) describe 'Domain Visualisation' as an emerging subject, also recognising a new generation of 'Information Scientists', as described by (White & McCain 1997). Although this reference dates back to 1997, it is still post internet, and could therefore in some senses be seen as a precursor to Web Science itself.

(Börner et al. 2003) propose that "*each research field can be characterized by a list of the most important keywords*." It is this theory that forms the basis for the proposed Web Science Lexicon. While (Börner et al. 2003) utilised a form of Network Analysis in their study, the work conducted within this thesis was collated utilising qualitative methods, (See chapter 6). Having identified the series of Key Topics from Web Science taught programmes and another series of Paper Topics from Web Science conference publications, the next stage in the analysis process, is to endeavour to classify these topics, grouping like with like. This will then form the structure of the proposed Web Science Lexicon. Biglan suggests in his research (Anthony Biglan 1973b) that academics studying a given subject are best placed to define and understand the nature of the subject. They conducted an early pre-web form of crowdsourcing in order to rate the attributes of various academic disciplines. Therefore, the final version of the Web Science Lexicon will be hosted in an online format, with the aim of making content accessible to the wider Web Science community, and providing a resource to which others can contribute.

The following sections of this chapter include an exploration of different ways in which the Web Science Lexicon might be structured, utilising the key and paper topics identified from the prior analyses of Web Science key and paper topics.

[35] https://en.oxforddictionaries.com/definition/lexicon

9.2 Web Science Lexicon: WAIS Fest Project

The University of Southampton's Web and Internet Science Research Group holds and annual research week nicknamed 'WAIS Fest'. This annual event provides researchers with the opportunity to conduct short group research projects involving the collaboration of other researchers from the group. The event is largely considered an opportunity for researchers to work on short projects which are of interest to them, but are unrelated to their main research topic. It is also an opportunity for researchers who need to recruit short-term help and collaboration from others within their existing research projects. As seen previously in section 7.4, this event was utilised for the process of interviewing Web Scientists. The 2017 WAIS Fest event [36], was utilised as an opportunity to do two things; firstly, to crowdsource content for the Web Science Lexicon. Secondly, to obtain feedback from Web Science researchers, regarding the structure and subject headings to be included within the Lexicon.

9.2.1 WAIS Fest Planning Meeting

The initial idea and proposal for WAIS Fest was to gain expert feedback on the Web Science key topics identified during the analysis of Web Science taught programmes and Web Science conference proceedings. The refined key and paper topics structure, could then contribute to a Wiki outlining the scope of Web Science. This idea was pitched to Lab managers during a WAIS Fest planning meeting. Chris Gutteridge, member of the University of Southampton's IT Innovation team one of the iSolutions staff in charge of the WAIS website, suggested that instead of the initial plan for the creation of a wiki, it would be beneficial to provide a resource on the WAIS website itself which outlines the scope of Web Science. Stemming from the initial concept of the Web Science subject categorization, he proposed that as an alternative to this author's initial idea of creating a wiki, the identified key and paper topics might be used to create a 'Web Science Lexicon' hosted within a University of Southampton web domain, and developed using the University's content management platform, Drupal.

9.2.2 The Focus Group

It was decided to convene a focus group as the main objective of a 'Web Science Lexicon' WAIS Fest project theme. The purpose of this focus group was to gain the feedback of people in the Web Science community regarding the structure and content of the proposed Web Science Lexicon. An informal focus group format was chosen due to the expected size of the group (estimated at anywhere between 2 or 3 to a dozen participants according to typical engagement with WAIS Fest projects). An informal discussion format was decided upon as opposed to structured interviews, as the goal was to foster a two-way discussion between the host and participants, and gain the views of the participants themselves, as opposed to influencing their feedback with structured questions. The secondary focus of the WAIS Fest Lexicon project theme, was to collate content for the Web Science Lexicon, via a survey of Web Science knowledge areas. The resulting content would then form part of a Lexicon topic webpage.

[36] https://secure.ecs.soton.ac.uk/wiki/w/WAISFest2017

9.2.3 The Google Survey

A second part of the theme involved a Google survey intended to crowdsource content for the Lexicon, to be included within a subject page for a specific topic. Created using a Google form, the short survey asked contributors to choose an area of Web Science that is of interest to them from the list of Web Science programme identified key topics. (They were also provided with an 'other' option and asked to specify if their area of interest within Web Science was not listed). Contributors were then asked to briefly describe their area of interest within Web Science, followed by a brief explanation of why their chosen topic is of interest to and relevant to Web Science. They were then asked to give some keywords that they feel describe their chosen topic. Finally, they were asked to provide one or more links to a publication which described their area of interest within Web Science. It was decided that no ethics application would be required for this survey, instead contributors would simply tick a box at the start of the survey agreeing to contribute content under a Creative Commons 3.0 license. Version 3 was the current version of this license during 2017, and is the same license used by Wikipedia allowing users to freely contribute content.[37]

9.2.4 Recruiting Project Participants

WAIS Fest begins with an introductory presentation event, in which a representative of each research theme is invited to give a brief presentation or project pitch to the entire WAIS research group. Following the initial theme pitches, people are then given the opportunity to mingle and chat to theme leaders and decide which projects they wish to participate in. As a theme leader, the author spoke to members of the WAIS research group, and verbally enlisted the help of several participants, who agreed to attend a focus group the following day, with the aim of exploring and brainstorming the structure and content of the proposed Web Science Lexicon. All members of WAIS spoken to by the author were also invited to participate in the online survey for crowdsourcing Lexicon content, which was also promoted via group-wide email.

9.2.5 The Focus Group: Discussion

The focus group discussion took place at 2:30 on Wednesday 7th June 2017. (Participants involved in the group are also credited in the acknowledgements section.) Contributors included:

- Ash Smith
- Rafael Melgarejo Heredia
- Bartosz Paszcza
- Caroline Harcow

Contributors were shown the existing two attempts to create outlines of Web Science, which include the Web Science Cluster diagram and Web Science Subject Categorization (Vafopoulos 2010). They were provided with paper handouts of these to browse through, as well as a handout listing the key and paper topics identified during the analysis of Web Science taught programmes and conference literature. They were also given a handout of a possible proposed structure for the taxonomy and asked to discuss and give their views on the structure and the list of topics. They were also asked for their views on how they would structure the taxonomy if they themselves were creating it.

---

[37] https://creativecommons.org/licenses/

9.2.6 Web Science Lexicon Draft Structure - High Level Key and Paper Topics

This list represents a draft semi-structured list of the majority of higher-level key and paper topics which were identified as part of Web Science. The list is ordered alphabetically. This structured list represents actual disciplines identified within Web Science, e.g. 'Computer Science'. This was an attempt to identify potential headings under which to structure the majority of key and paper topics. It was though that these higher-level headings could represent majority of lower level key and paper topics identified, although a problem would be potential overlaps between some topics. For example, would 'social networking' be classified under 'Psychology' for user behaviour, 'Sociology' for the societal impact, or Computer Science to represent the technological platform.

Table 46 Draft Concept for Lexicon Headings

Draft Lexicon Main Heading Structure Featuring High Level Key and Paper Topics
Arts and Culture
Biology
Business
Marketing
Entrepreneurship
Criminology
Computer Science
Economics
Education
Geography
Health
Law
Media
Network Science
Politics
Social Sciences
Digital Sociology
Sociology
Psychology
Theology

The consensus was that the high level topics example shown in Table 46, although comprehensive, were not necessarily specific enough to Web Science, and could just as easily represent, for example, a list of all the disciplines taught by a university. While the example indeed illustrates that Web Science is multi-disciplinary, the list shown in Table 46, is not specifically identifiable as Web Science. It was instead suggested that the three main headings of the Web Science Subject Categorization; (Web Technologies, Web Analysis and Web Society), might serve more effectively as high-level headings under which to group lower level key and paper topic headings. Interestingly, the headings of the WSSC, (Web Technology, Web Analysis and Web Society), largely correspond with three key section headings outlined within the (Berners-Lee, Weitzner, et al. 2006) original proposals for Web Science, which include: Engineering the Web, The Analysis of the Web and Social Aspects. The WSSC headings are also high-level enough to provide a structure which is identifiably Web Science, adequately representing the sociotechnical nature of the subject, without there being too much overlap between the lower level topics included under each heading. Rafael Melgarejo Heredia proposed the creation of a matrix structure in order to demonstrate the overlap between Web Science topics. This related to a prior work and publication he had co-authored, (Heredia & Vinueza

2015). He suggested that the three main headings from the Web Science Subject Categorization might be used to group the headings. He also pointed to the 'Outline of Science'[38] a Wikipedia based taxonomy listing all the sciences. He suggested that this might serve as an example of how a similar Web Science taxonomy could be structured, and also that some of the topics listed within the 'Outline of Science', might be relevant to Web Science.

9.2.7 Focus Group Outcomes

To conclude the focus group, contributors were provided with copies of the handouts to take away from the group, as several had expressed a desire to attempt to structure the Web Science key and paper topics themselves in their own time. Rafael Melgarejo Heredia followed up on the discussion, by providing an example of the Matrix structure that he had suggested during the focus group, utilising the Outline of Science higher level headings as an example, grouping these under the three main headings of the Web Science Subject Categorization and illustrating the overlap between subjects. This is described in further detail below and his example is included in Table 47.

In addition to discussing the structure of the Lexicon, participants were also asked to comment on the Key topics and Paper topics identified. Participants were provided with printouts of the lists of key topics and paper topics, as well as a copy of the WSSC for comparison. Limited feedback was achieved regarding the key topics and paper topics. Participants in the group did not disagree with any of the key or paper topics, but they did not feel able to provide any further feedback. It was decided that this was because while participants were confident of their own individual research areas, they did not feel equipped to assess the content of Web Science as an entire subject. This is why students are an excellent knowledge base for their own research subject, and why Web Science is an interdisciplinary subject requiring the contribution of people from across multiple disciplines. No one person can be an expert in the entirety of such a large interdisciplinary subject. This highlights the need for a resource such as the proposed Web Science Lexicon, in order to collate knowledge from multiple disciplinary perspectives into a centralised resource.

9.2.8 Matrix Idea for the Web Science Lexicon

A notable outcome of the focus group run during WAIS Fest, was the contribution of Rafael Melgarejo Heredia. He suggested the idea of utilising a Matrix table structure for the Web Science Lexicon, in order to illustrate the interdisciplinary nature of Web Science. Rafael had utilised a matrix previously, in his publication entitled 'A Proposal Model to Monitor Interdisciplinary Research Projects in Latin American Universities' (Heredia & Vinueza 2015). Rafael produced an example of his proposed visualisation method, which is shown in Table 47, which utilises the Outline of Science topics as an example, structured under the three main headings of the Web Science Subject Categorisation. The table illustrates how it is possible to utilise a matrix structure to show different classifications and represent the interdisciplinary nature and overlaps between related subjects utilising a table format. Rafael Melgarejo Heredia explained that his example matrix presented in Table 47, showed only two levels of headings, (e.g. Level 1 'Natural', level 2 'Physics') but that it could be adapted to show three to four levels if needed. Where additional levels included, this would allow further 'lower level' topics to be represented. While this is an excellent format for representing the overlap between the three main WSSC categories, the 'book structure' within the platform of the Drupal site used for hosting the online version of the Lexicon does not facilitate the

[38] https://en.wikipedia.org/wiki/Outline_of_science

representation of a matrix structure such as this. Therefore, the online version of the Lexicon will need to feature a more linear menu structure, more similar to that of the original WSSC format, which was a list of subjects, structured under higher level headings; although it will be possible to include the idea of multiple levels within the heading menu structure.

Table 47 Example Lexicon structure featuring Outline of Science subjects (Rafael Melgarejo Heredia)

		Web Technologies	Web Analysis	Web Social
Natural	Physics			
	Chemistry			
	Earth		X	
	Biology			
Formal	Computer Science	X		
	Mathematics		X	?X
	Statistics		X	?X
	Systems Science		X	X
Social	Anthropology			X
	Business			X
	Civics		X	X
	Cognitive			X
	Criminology	X	X	X
	Cultural		X	X
	Demography		X	
	Development		X	
	Economics		X	X
	Education		X	X
	Environmental		X	X
	Gender and sexuality		X	X
	Geography		X	X
	Gerontology		X	X
	History			
	Industrial relations		X	X
	Information Science	X		
	International		X	X
	Law		X	X
	Legal management			X
	Library		X	
	Linguistics		X	
	Management		X	X
	Media		X	X
	Philosophy		X	X
	Planning		X	X
	Political		X	X
	Psychology		X	X
	Public administration		X	X
	Social work		X	X
	Sociology		X	X
	Sustainable development		X	X
	Sustainability		X	X
Applied	Agronomy			
	Architecture			
	Applied chemistry			
	Computing technology	X		
	Education		X	X
	Electronics	X		
	Energy	X		
	Engineering	X		
	Environmental		X	X
	Forensic		X	X
	Health		X	X
	Applied linguistics			X
	Management		X	X
	Applied mathematics		X	X
	microtechnology	X		
	military		X	X
	applied physics			
	spatial			

9.3 Utilising Biglan's Classification Systems to Categorize Web Science Topics

The following section examines the feasibility of utilising Biglan's classification of academic disciplines (Antony Biglan 1973) (Anthony Biglan 1973a) to categorise the topics identified as part of Web Science. This is not only an alternative means of realising the matrix structure proposed by (Heredia & Vinueza 2015), it would also provide an indication of the bias of topics within Web Science according to Biglan's classification. The creation of Table 49 was considered a useful experiment, in order to determine the bias of Web Science according to Biglan's classifications. Table 48 shows Biglan's original classification table for academic disciplines. Topics are structured under six categories, Non-life and Life, Hard and Soft, Pure and Applied.

Table 48 Biglan's Original Classification of Academic Disciplines

	Non-Life		**Life**	
	Pure	**Applied**	**Pure**	**Applied**
Hard	Physics Chemistry Mathematics Engineering	Engineering (e.g., Chemical, Mechanical, Civil, Nuclear)	Biology Zoology Bacteriology Philosophy Life	Medicine Agriculture Physical Education
Soft	(e.g., Chemical, Mechanical, Civil, Nuclear)	Library Science Finance Operations Research Foreign Languages Law Architecture	Psychology Sociology Anthropology Literature	Education Organizational Behaviour Social Work Performing Arts

In this experiment categorizing Web Science subjects according to Biglan's classification, the (Anthony Biglan 1973a) table was used as a basis for the creation of a Web Science version of the matrix classification table. Based on the analysis of the Web Science conference material, nearly all coded content falls under one of the three headings, Web Technologies, Web Analysis and Web Society. Therefore, these headings were added as an extra layer of classification within the (Anthony Biglan 1973a) table.

Whereas Biglan sought the opinions of academics in YEAR 1973, this was before the advent of the Web, and it is now possible to research categories of disciplines without the need for a time-consuming and in-depth survey. However, use of the internet to research subjects lacks the insight and detailed expertise of academics. During this experimental version of a 'Biglan's classification applied to Web Science', Web searches were utilised as a reference to ascertain the nature of subjects, determining whether they were hard, soft, pure or applied etc. Sources used as a reference to classify topics include a more detailed version of the Biglan diagram[39], (Alise 1980), (Clark 2003), and (Becher et al. 2001). The WSSC (Vafopoulos 2011), was also utilised to aid in the classification of the Web Science topics.

[39] https://goelsan.wordpress.com/2010/07/27/biglans-classification-of-disciplines/

Table 49 shows the restructured version of the (Anthony Biglan 1973a) classification table. The three main WSSC headings of Web Technologies, Web Analysis and Web Society were added as an extra layer. In addition, the original categories were rearranged to facilitate coverage of topics, according to the three new headings. For example, it was decided that Web Technologies, and Web Society featured examples of 'pure' subjects, yet Web Analysis was solely 'applied' (McGrath 1978). Therefore, it was possible to maintain mostly the same basic structure as Biglan, only with rearranged categories, in order to accommodate the Web Technology, Analysis and Society Lexicon Headings.

Table 49 Biglan's Classification Applied to Web Science

	Hard			Soft	
	Pure	**Applied**			**Pure**
	Web Technologies		**Web Analysis**	**Web Society**	
Life	Biology Mobile Web Web 2.0	Geo-tagging Internet of Things Knowledge Patterns Open Data	Content Analysis Narrative	Arts and Culture Crowdsourcing Education Health Online Engagement Social Networking	Digital Sociology Digital Ethnology Demographics Politics Psychology
Non-Life	Linked Data Semantic Web Web Architecture Web Languages Web Ontologies	Algorithms Computer Science Quality Management Web Modelling	Data Mining Network Science Statistical Analysis Visualisation Web Archiving Web Graph Web Search	Business Economics Law Marketing Media	Communication Science Geography Research Methodology Security

The inclusion of some key and paper topics within this example was problematic, due to the fact that they represent 'application areas', methods of analysis or technologies, which are a key feature of Web Science, as opposed to academic disciplines. For the purpose of this example, such 'application area' topics are only included when they occur frequently within the analysis of key and paper topics, and can be attributed with the traits of Biglan's classification system.

The experimental Biglan Classification applied to Web Science, as shown in Table 49 indicates that, perhaps unsurprisingly, the topics included under the heading 'Web Society' were the most evenly spread between life and non-life. A consequence of the way in which this model is designed, means that there are less topics contained within the 'Hard' and 'Applied' and 'Life' categories in the 'Web Technologies' and 'Web Analysis' sections, as the majority of topics within the 'Life' section are classified under the 'Web Society' headings, as 'life' generally pertains to 'society' within the Web Science Lexicon structure. However, there are a number of exceptions; for example, 'narrative'. The construction of 'narratives' from online content, generally relates to the complication of life 'stories' such as the example given in (Wienberg & Gordon 2015), which focuses upon the ethical issues involved within the public documentation of users' private lives in online narratives, (Refactorings et al. 2007). Another example is that of 'Knowledge Patterns', which relates to the recording of user knowledge (hence the life category) within a knowledge management systems (Web Technology), in order that others my benefit from the recorded knowledge, (Rech et al. 2009).The rationale for including geo-tagging under the Web Technology, Hard, Applied and Life category, is that geo-tagging generally involves people tagging themselves at a location, hence the 'life' classification, while Web.2.0 Technologies utilised to geo-tag facilitate socio-technical interaction and user generation of Web content.

The results of this experiment superficially seem to suggest, that more subjects within Web Science fall under the 'non-life' category. While for example, economics and marketing represent parts of 'society', they are not classified in the original Biglan table as 'life' subjects because they do not strictly relate to human behaviour. One aspect that this diagram fails to represent, is the frequency at which each of the topics are coded. So, while the spread of subject superficially indicates slightly less subjects within in the 'life' category, it is known from the results of the prior analysis of key and paper topics, that some of the 'life' subjects represent some of the most frequently coded key and paper topics. While this classification of Web Science disciplines is only a provisional experiment, it still provides an interesting starting point for comparison and further study. Future work could include a more formal study centred on the Biglan classification system, also involving the participation of academics to rate subjects.

9.4 Structuring the Lexicon Topics

The topics to be included within the Web Science Lexicon are those identified within the prior analyses of Web Science taught programmes and conference literature. (See sections 8.2.1 and 8.3.2 for a more detailed description of the criteria for the identification of key and paper topics). A combination of these key and paper topics represents the common themes identified within Web Science as it is taught, and as it is published. This provides the content for the proposed Lexicon.

The progression of structuring the key and paper topics into categories within the Lexicon, represents the visible application of axial and selective coding. The process of deciding the structure of topics within the Lexicon is as follows. Because it is not possible to utilise a matrix table structure within the menu of the site produced using the Drupal content management system provided by the University for designing the Lexicon, the topics are displayed in a structured list format similar to that of the WSSC. Content is formed from the key and paper topics identified during the prior analyses of Web Science taught modules and conference publications.

Various sources, including papers and websites (included as footnotes), Wikipedia's Outline of Science, and the existing WSSC topics are utilised as a reference when grouping topics and subtopics, in order to decide which topics are related. Another reference utilised when deciding the structure of topics is the co-occurrence of key topics and paper topics within the analyses of Web Science conference material, as well as predefined author keywords within papers. If a key or paper topic is frequently identified as occurring with another key topic of paper topic, the likelihood that they are related and should be grouped together is higher. For example, Social Networking, and Web Society were frequently coded together, and therefore, the likelihood is that Social networking is a subtopic of the heading, Web Society.

9.5 Use of the WSSC Headings

The majority of survey participants who rated the WSSC, felt that the WSSC heading areas were either 'relevant' or 'very relevant', this was confirmed by the outcome of the WAIS Fest 2017 focus group. Additionally, the results of the analysis of Web Science modules and papers show 'Web Analysis' and 'Web Society' to be two of the most frequently coded topics. Therefore to this end, it was decided to use the three main headings of the WSSC as a basis for structuring Lexicon content.

The WSSC headings were also chosen, as due to the diverse range of topics, it is difficult to provide a structure under which all the topics can be classified without overlapping, especially given the need to represent topics in a list format suitable for inclusion on the Drupal site.

9.6 Overlapping Topics

During the process of structuring the topics, it was evident that some of the lower level topics relate to more than one higher-level topic. This corresponds with the findings of (Hooper et al. 2012). When mapping the Web Science discipline utilising the Web Science Cluster diagram, they discovered that topics did "*not necessarily directly map to disciplines*." Therefore, low level topics within the Lexicon are grouped under the higher level topic of 'best fit', and the subject pages within the online version of the Lexicon will include a list of 'related topics', showing which other topics a topic relates to.

9.7 The Web Science Lexicon Topics

The following section outlines **the structure and topics to be included within the Web Science Lexicon.** These topics represent the key and paper topics or themes, which were coded during the prior stages of analysis. These topics are now structured into different levels of topics. Each topic includes a description, and where a subtopic is grouped under a higher-level topic, a brief justification is provided. It was decided to only include the more frequently coded subjects as headings within the Lexicon, each with its own dedicated page. However, where a topic included covers a lower level topic not included, the lesser coded topic is referenced within the subject page. The format of the Web version of the Lexicon will facilitate further contributions and changes to the structure beyond the timescale of this thesis, therefore, were any of the lesser coded subjects deemed of greater importance at a later time, they can still be given a heading/allocated as a higher level topic within the Lexicon.

9.8 Web Technologies

Relates to Web technology, infrastructure and Web Architecture. Can also include web languages and standards, as well as the hardware on which the web is based.

9.8.1 Big Data

As may be obvious from the name, 'Big Data' refers to very large datasets. These datasets usually require some form of querying or analysis in order to derive useful information. Topic also relates to: Data Mining, Web Analysis. Web Archiving, Knowledge Patterns

9.8.2 Cloud Computing

This refers to the distribution of computing resources over a network. This can refer to the use of online applications, or the remote storage of online data. There are three main types or models of cloud, these include public, private and hybrid.

9.8.3 Open Data

Open Data refers to data which has been made publicly available via the Web.

9.8.4 Semantic Web

The term "Semantic web" refers to the vision of Web of linked data", as described by (Miarka 2013).Therefore, Linked data is grouped as a subtopic of Semantic Web.

9.8.4.1 Linked Data

(Miarka 2013) state that the main objective of *"Linked data is to allow users to share structured data in the web environment*."

9.8.4.1.1 RDF

RDF is a sub-topic of Linked Data, as RDF is used by linked data. According to (Miarka 2013), RDF data model is used to "*publish structured data on the web*." Additionally, RDF links are used to "*link data sources"*.

9.8.4.1.1.1 Knowledge Patterns

According to (Miarka 2013), RDF is used to identify knowledge patterns in linked data. Therefore, 'knowledge patterns' is grouped under RDF. Knowledge patterns are also related to 'linked data'.

9.8.4.2 Web Ontologies

Another subtopic of Semantic Web, Web Ontologies refers to "*many ontologies, vocabularies or knowledge bases"* which are *"created by web ontological languages. Languages RDF+RDFS and OWL*," which are described as "*the most important web ontological languages.*" (Miarka 2013)

9.8.5 Web Applications

Web application refers to an application "*in which all or some parts of the software are downloaded from the Web each time it is run*."[40] Different types of Web application include: Browser based, Client Based and Mobile Web Applications.

9.8.5.1 Web Modelling

Web modelling refers to the design of large-scale Web applications[41]. Therefore, it is grouped under 'Web Applications'. Unified Modelling Language (UML) is an example of a Web Modelling language. (Wimmer et al. 2007)

9.8.6 Web Architecture

Refers to the infrastructure of the Web. Web Architecture "*focuses on the foundation technologies and principles which sustain the Web, including URIs and HTTP*."[42]

9.8.7 Web Development

The creation of websites, Web based content and Web based applications.

9.8.8 Web Languages

Refers to the languages in which Web content is authored.

[40] https://www.pcmag.com/encyclopedia/term/54272/web-application

[41] https://en.wikipedia.org/wiki/Web_modeling

[42] https://www.w3.org/standards/webarch/

9.8.8.1 Hypertext

Hypertext is an example of a language used to create the World Wide Web, and is therefore grouped under the higher level topic 'Web languages'.

9.8.9 Web 2.0

Web 2.0 describes Web technology which enables the authoring of Web content by Web users. Web 2.0 enables "*large numbers of people to come together to work, share and build*" web content. (Shuen 2008)

9.9 Web Analysis

Relates to the analysis and study of the Web, or Web data, in order to provide insight into a topic or reveal patterns of behaviour on the Web.

9.9.1 Algorithms

An algorithm may be described as what can be "*a simple function that takes a set of parameters as input and returns an output*"[43].

9.9.2 Content Analysis

This topic refers to the analysis of Web based content.

9.9.3 Data Mining

The process of analysing large datasets in order to extract usable data and find patterns. This category was coded on a great number of occasions, as the majority of Web Science papers analysed utilised some form of dataset collated from existing web data.

9.9.4 Narrative

The process of deriving a 'story' based on an analysis of facts and data. This can be achieved either manually or utilising an algorithm

9.9.5 Network Science

Network Science is grouped under the heading 'Web Analysis', because this is the heading under which the WSSC grouped the topic: 'Networks'. In addition, the majority of the papers coded with this topic refer to the analysis of network content and behaviour on networks, as well as the analysis of network structure. However, Network Science can also refer to the technological platforms which constitute network structure, including for example, TCP/IP.

9.9.5.1 Network Theory

"*Network theory is the study of complex interacting systems that can be represented as graphs equipped with extra structure*." (Baez n.d.)

9.9.6 Statistical Analysis

Statistical analysis can be described as the "*science of collecting, exploring and presenting large amounts of data to discover underlying patterns and trends*." [44]

---

[43] http://giocc.com/a-gentle-introduction-to-algorithms-for-web-developers.html

[44] https://www.sas.com/en_gb/insights/analytics/statistical-analysis.html

9.9.6.1 Topic Modelling

"*In machine learning and natural language processing, a topic model is a type of statistical model for discovering the abstract "topics" that occur in a collection of documents. Topic modelling is a frequently used text-mining tool for discovery of hidden semantic structures in a text body*."[45] Therefore, topic modelling is grouped under the higher level topic, statistical analysis.

9.9.7 Visualisation

The process of displaying data in a visual format, in order to communicate it effectively. "*A visualization method is a systematic, rule-based, external, permanent, and graphic representation that depicts information in a way that is conducive to acquiring insights, developing an elaborate understanding, or communicating experiences*." (Lengler & Eppler 2007)

9.9.8 Web Archiving

The process of storing data on the web. Also involves the process of searching for and retrieving data stored.

9.9.9 Web Graph

(Hendler et al. 2008) describe Web Graph as one way of examining the structure of the Web as a whole. They describe it as an "*abstraction of the Web based on one part of the processing and protocols underlying its function*."

9.9.10 Web Search

The process of searching for web based content. Can refer to Web search engines such as Google.

9.10 Web Society

The study of one or more aspects of humans or society in general, in the context of the Web.

9.10.1 Academic Publishing

"*Academic publishing is the subfield of publishing which distributes academic research and scholarship*." [46]

9.10.2 Accessibility

Relates to ensuring Web content is usable and viewable by all, including those with disabilities. Accessibility also refers to the availability of Web content to all, regardless of their geographical world location or language. Related topics include: Human Computer Interaction, and the 'Digital Divide'.

9.10.3 Arts and Culture

Refers to art and cultural content on the Web. In at least one instance, this refers to online availability and use of museum content. (Mulholland 2015)

[45] https://en.wikipedia.org/wiki/Topic_model

[46] https://en.wikipedia.org/wiki/Academic_publishing

9.10.4 Business

Can be described as a commercial activity or the act of trading. Also refers to organisations. Also relates to entrepreneurship.

9.10.5 Cognitive Science

As stated by Carnegie Mellon University, Cognitive Science is an interdisciplinary subject, encompassing: "*philosophy, cognitive psychology, computer science, and neuroscience*"[47] While the programme they describe is offered by the department of philosophy, suggesting that Cognitive Science might be a sub-topic of philosophy, it also encompasses psychology and Computer Science. Therefore it is not grouped as a sub-topic and is listed as a higher level topic.

9.10.6 Communication Science

Refers to the study of "*content, use and effects of communication and of various forms of media, from radio and television to newspapers and the internet*"[48]. There are four main types of Communication Science, as described by the University of Amsterdam. These include: "*Persuasive Communication, Entertainment Communication, Corporate Communication and Political Communication*".

9.10.7 Crowdsourcing

The Web facilitated process of enlisting the collaboration of a large number of people, in order to accomplish a task that would otherwise be unachievable. One example of this includes Amazon's Mechanical Turk, a platform which was referred to within a number of Web Science conference papers.

9.10.8 Cybercrime

Refers to any form of crime committed online. This most commonly occurs hacking, but can also refer to situations such as online bullying or abuse.

9.10.9 Digital Anthropology

Both Sociology and Anthropology focus upon "*studying the behaviour of humans within their societies*"[49]. "*The key difference between the two social sciences is that sociology concentrates on society while anthropology focuses on culture*." Digital Anthropology focuses on the study of online culture, and also uses online digital media to observe general culture. Survey Respondent 11 suggested that Anthropology is a subject which is omitted from the WSSC subjects. (Phethean et al. 2016) state the following, reinforcing the association between 'Anthropology' and Web Science:

> "*The Web's impact on society, and vice versa, receive equal importance and focus. Sociology, politics, law, economics, and anthropology all provide invaluable contributions to the field and are fundamental in ensuring a holistic and societally beneficial analysis of the Web*."

[47] https://www.cmu.edu/dietrich/philosophy/research/areas/science-methodology/cognitive-science.html

[48] http://www.uva.nl/en/programmes/bachelors/communication-science/communication-science.html

[49] https://web.archive.org/web/20180225050701/http://education.seattlepi.com/anthropology-vs-sociology-1536.html

9.11 Digital Ethnography

The first-hand observation of a social situation or setting. Digital ethnography specifically refers to the use of technology and digital resources to observe society.

9.11.1 Digital Sociology

"Digital sociology is a sub-discipline of sociology that focuses on understanding the use of digital media as part of everyday life, and how these various technologies contribute to patterns of human behaviour, social relationships and concepts of the self." [50]

9.11.1.1 Demographics

Demographics refer to the "quantifiable characteristics of a given population"[51], these include attributes such as gender, age, race etc. Demographics is referred to as a sub-field of sociology, and is therefore grouped under 'Digital Sociology'

9.11.2 Disaster Response

This topic relates to the use of online data to analyse user behaviour during and following disasters. It also refers to attempts to use online data to predict disaster occurrence.

9.11.3 Economics

It was considered whether the topics 'economics' and 'ecommerce' should be merged; however, some papers discus economic issues which may not refer to ecommerce specifically. Therefore ecommerce was allocated as a subtopics of economics.

9.11.3.1 eCommerce

eCommerce refers to commerce conducted online. This includes activities such as online shopping, and also online markets, online accounting systems, financial management and online banking.

9.11.3.1.1 Recommender Systems

A sub-topic of eCommerce, recommender systems analyse user behaviour based on past purchases and recommend purchases customised to users' perceived tastes.

9.12 Education

Refers to "*the process of facilitating learning, or the acquisition of knowledge, skills, values, beliefs, and habits*." [52]

9.12.1.1 e-learning

Refers to technology facilitated learning. This can include distance learning via online platforms. Also related to the provision of online courses known as MOOCs (Massive Open Online Course).

9.12.2 Ethics

Refers to moral principles applied in a web context.

[50] https://en.wikipedia.org/wiki/Digital_sociology

[51] https://en.wikipedia.org/wiki/Demography

[52] https://en.wikipedia.org/wiki/Education

9.12.3 Health

Within the context of Web Science, papers coded with 'health' may involve the use of health data to study the spread of epidemics. Telemedicine and online health provision is an additional feature of health in relation to Web Science. Another increasing trend, is the use of health websites facilitating online self-diagnosis of health conditions.

9.12.4 Human Computer Interaction

HCI refers to the interaction between humans and technology, and the design of accessible interfaces which better facilitate this interaction.

9.12.5 Law

In the context of Web Science, law not only refers to legal issues government statutes, it also relates to standards governing online behaviour.

9.12.6 Marketing

The promotion of products or services.

9.12.6.1 Online Advertising

Refers to advertising in an online context in order to promote good and services. Includes recommender systems.

9.12.7 Media

Mass communication; can include: books, films, music, news, etc.

9.12.8 Geography

Is defined as the study of the "*lands, the features, the inhabitants, and the phenomena of Earth*." [53]In the context of the analysis of Web Science publications, this topic was coded if papers referred prominently to the analysis of geographical locations within studies.

9.12.8.1 Geo-tagging

Can refer to the act of tagging at a location and sharing that location via the Web, often via a social media platform. Geo-tagging was allocated as a sub-topic of geography, as it refers to geographical location.

9.12.9 Philosophy

"*The use of reason in understanding such things as the nature of the real world and existence. The use and limits of knowledge, and the principles of moral judgment*." [54]

9.13 Politics

Politics is defined as the "*process of making decisions applying to all members of each group*". Relates to government, democracy and law.

9.13.1.1 Government

The group of people in a society with the authority to govern, and tasked with the process of deciding and implementing laws. [55]

[53] https://en.wikipedia.org/wiki/Geography

[54] http://dictionary.cambridge.org/dictionary/english/philosophy

9.13.1.2 Democracy

The process of electing government representatives.

9.13.2 Privacy and Trust

In the context of Web Science, this relates to the sharing and protection of personal data online.

9.13.3 Research Methodology

The methods by which research is conducted, for example qualitative or quantitative.

9.13.4 Security

Security relates to the protection of systems and personal data from cyberattacks and other online threats.

9.13.5 Social Networking

Relates to the study of social networks and their properties. Also encompasses the use of social networking data in order to address a reach question, for example looking at some form of user behaviour utilising data from social networks.

9.13.5.1 Microblogging

The act of sharing content to a microblog site. Microblogs are a short form of blogging limited by content size. The most common example is Twitter. Because many tweets are public, Twitter data is utilised in a large number of the studies discussed in the Web Science conference papers, and the topic was consequently coded frequently.

9.14 User Behaviour

Activity demonstrated by users of the Web. Many of the Web Science conference papers analysed feature studies of some aspect of online user behaviour.

9.14.1.1 Online Engagement

The topic relates to users use of the Web. This includes the popularity of online services, and users desire to utilise online platforms, as well as how they choose to use them.

9.14.1.2 Online Identity

Online user profiles and patterns of behaviour, often on social networking sties.

9.14.1.3 Online-Offline Community

The relationships between real-life communities and online communities.

9.14.1.4 Psychology

Psychology can be described as "the science of the mind" [56] and relates to human motivations and behaviour. In the context of Web Science, it refers to user motivations and activities demonstrated in an online environment.

9.14.2 Virtual Community

The creation of an online community, often via a social networking platform. Virtual communities can also develop within online multiplayer gaming platforms.

---

[55] https://en.oxforddictionaries.com/definition/government

[56] http://www.bbc.co.uk/science/humanbody/mind/articles/psychology/what_is_psychology.shtml

9.15 Web Publishing

This topic refers to the publishing of content on the Web. Can include website content, and also ebooks and multimedia.

9.16 Eliminating Subjects

A number of topics were omitted from the final structure of the Lexicon. While this included topics which were coded less frequently, a number of more frequently coded topics were also omitted. These are explained below.

9.16.1 Sociology vs Digital Sociology

It was decided to merge these two topics, as the majority of examples of the two topics were co-coded within the paper analysis. Also, the majority of examples of sociology within the context of Web Science refer to digital sociology.

9.16.2 Social Science

The decision was made to drop 'Social Science', as the majority of social science topics were already represented by existing specific social science topics, such as for example, 'digital sociology'.

9.16.3 Computer Science and Information Commutation Technology

Two topics which were in the midrange of coding frequency, yet were omitted from the taxonomy include: Information Communication Technology and Computer Science. Neither of these subjects fitted squarely under one of the three main headings. For example, Computer Science could not be classified into a single category, as it features both technologies and analysis. Information Technology includes elements of all three main headings. Despite many Web Scientists originating from Computer Science, the subject was not coded as frequently as might have been expected during the paper analysis. This is evidence to support the fact that Web Science and Computer Science are indeed two separate and different disciplines. The comparison between Web Science modules and Computer Science modules in Table 12 of Chapter 4 also suggests that there is limited overlap between the two subjects in a taught context.

9.16.4 Social Machines

While social networks are example of social machines, there are types of social machines which are not specifically social networks. Social machine refers to a socio-technical system, which encompasses both humans and technology, which mutually shape each other. One description of Web Science is the study of social machines. It is perhaps surprising, therefore, that the topic 'social machine' was seldom discussed as the specific subject of Web Science conference publications. While it is frequently referred and eluded to in relation to the study of social networks, the specific topic of 'social machines' is referred to by name far less frequently than might be expected.

9.17 WSSC and Key and Paper Topic Comparison

As seen during the review of Web Science literature in chapter 3, the Web Science Subject Categorisation and the Web Science Cluster diagrams are the only current attempts to outline the scope of Web Science. The WSSC is the more detailed of the two, and constitutes a useful benchmark against which to compare the key and paper topics identified as part of this study. While differing significantly to the range of key and paper topics identified during this thesis, the WSSC does share some common topics, and was therefore utilised as a point of reference when deciding which of the three headings topics should be grouped under. It was especially useful in the case of the more technical topics, as there was generally a greater overlap between the key and paper topics and the WSSC in this area.

Table 50 Structuring the Key and Paper Topics According to WSSC Categories

WSSC Topics	Key and Paper Topics
Web Technologies	
C.1 General Web Technologies C.2 Web Milieux Document technologies; Hypertext technologies; Internet technologies; Mobile Web technologies; Grid and Cloud computing technologies C.3 Basic Web Architecture HTTP and related technologies; URIs; HTML; XML; CSS and related technologies; Interfaces and Browsers; Servers Web Services C.4 Web 2.0 technologies C.5 Semantic Web/Linked Data Metadata; Knowledge Representation; Ontology Languages; Linked Data; Natural Language Processing; Provenance systems in the Web C.6 Internet/Web of Things D.	Big Data Open Data Semantic Web Linked Data RDF Knowledge Patterns Web Ontologies Web Applications Web Modelling Web Architecture Web Languages Hypertext Web 2.0
Web Analysis	
D.1 General Web Analysis D.2 Mathematical Methods of Web analysis Web data sampling and analytics; Logic and Inference in the Web; Statistical Inference in the Web; Statistical Analysis of the Web; Web as a Complex System; Graphs; Networks; Mathematical methods for describing Web services; Crawling; Indexing and Searching; Data Mining; Information Retrieval and Machine Learning; Other Algorithms for the Web	Algorithms Content Analysis Data Mining Narrative Network Science Network Theory Statistical Analysis Topic Modelling Web Archiving Web Graph Web Search
Web Society	
E. Web Society E.1a Economics Goods in the Web; The Web economy; Antitrust Issues and Policies in the Web; Intellectual property and digital rights management; Web-based economic development E.1b Business E-commerce Business models in the Web; Advertising in the Web; sponsored search E.2 Social Engagement and Social Science Social networks; Mass phenomena; Collective intelligence; Peer production; Globalization; Systems; Social structures and processes; Virtual communities, groups and identity; Social capital and power inequality in the Web; On-line lives, intergenerational differences; Mass media E.3 Personal Engagement and Psychology System Psychology and Behaviour; Child and adolescent psychiatry; Tele-working E.4 Philosophy, Philosophy of information; Objects; Reference and Cognition in the Web; Ethics in the Web E.5 Law Intellectual Property in the Web; Digital Rights Management; Digital crime; Laws for Web access; Antitrust Law E.6 Politics and Governance Political science; E-Government; E-Politics; E- Democracy; Policy and Regulation; Web Governance; Privacy; Trust; Security; Network neutrality; E-Inclusion	Academic Publishing Cognitive Science Communication Science Crowdsourcing Cybercrime Digital Anthropology Digital Ethnography Digital Sociology Demographics Economics Education e-learning Health Human Computer Interaction Law Media Marketing Geography Geo-tagging Philosophy Politics Government Privacy and Trust Research Methodology Social Networking Microblogging User Behaviour Online Engagement Online Identity Online-Offline Community Psychology Virtual Community Web Publishing

9.18 Online Lexicon: The Drupal Site

The online version of the Lexicon was hosted as part of a Drupal site. This setup was used because it is one of the available tools used by the University of Southampton Web team, and includes a content management system which enables both non-developers and developers to create a website. Anyone who wishes to contribute content can be provided with a login to the site. The Lexicon website can be accessed from this live link: https://webscience-lexicon.southampton.ac.uk/

Figure 7, shows the homepage of the Web Science Lexicon. The 'Book Navigation' menu seen on the left hand-side represents the three main WSSC headings are utilised as the highest level categories under which all other topics are grouped.

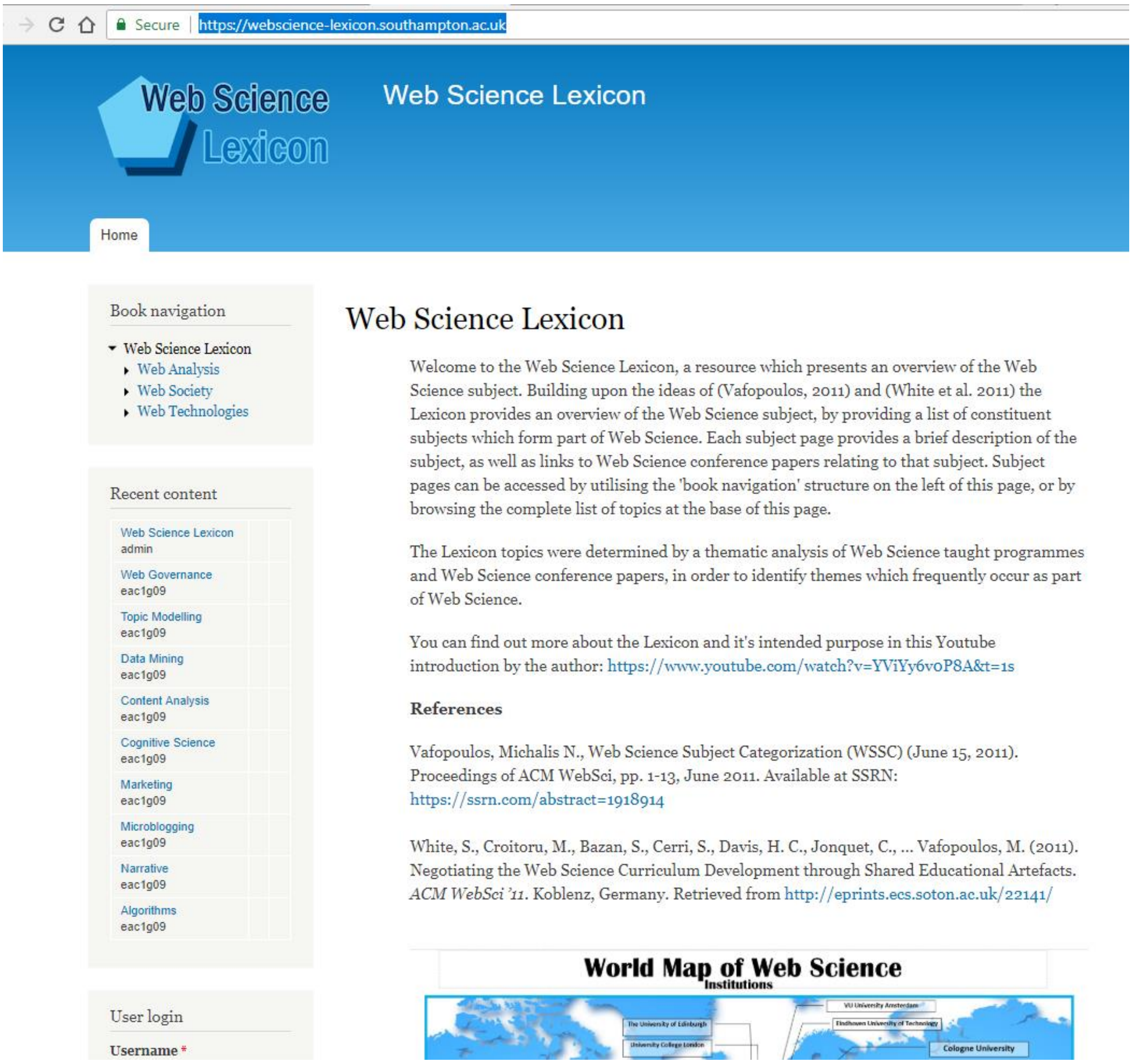


Figure 7. Preview of the Web Science Lexicon 'main menu'

The content of each individual topic page within the Lexicon page is as follows:

- Subject heading
- Brief description of content
- Recommendations reading for several papers relating to the topic
- Related Lexicon topics

The menu structure consists of topics identified as higher level topics, under which lower level topics are be grouped, according to the structure outlined in sections 9.7 to 9.15. Where possible, subject pages will be created for more frequently coded topics. Each of the headings listed in sections 9.7 to 9.15 represent a subject page within the online Lexicon. As well as a description of each topic, each page within the Lexicon also features links to papers relating to each of the topics, as well as other topics to which each topic relates.

9.19 Conclusions

While the Lexicon is a representation of Web Science, it must be stressed that this is just one possible representation of the subject Web Science. It is unrealistic to expect to create a fully comprehensive and watertight representation of such a rapidly evolving subject. The Lexicon will provide one possible representation of the discipline, and will serve as a starting point from which to stimulate further discussion and research.

The menu structure of the Lexicon included on the website was unable to be resented using the matrix table structure suggested by (Heredia & Vinueza 2015), due to the limitations of the editing tools for site content. However, future work could involve migrating the site to a more flexible platform that is able to facilitate the concept of a matrix structure, similar to those explored in Table 47 and Table 49.

The live version of the Lexicon serves as a starting point for facilitating further contributions of crowdsourced content from Web Science academics and students at the University of Southampton. This should facilitate expert input to the project, and continue the development of the Lexicon beyond the initial lifespan of this thesis development. This vision for crowdsourcing additional Lexicon content builds on the ideas of (White et al. 2011) for crowdsourcing Web Science content, and also the suggestion of (Anthony Biglan 1973b) that academics in a given area are the best source of knowledge about a given subject.

10 Conclusions

10.1 Introduction

This chapter will feature a summary of the final project outcomes, addressing how each of the four research questions was answered, and also outlining the contributions made by this thesis.

10.2 Answering the Research Questions

The following section outlines how each of the research question outlined in chapter 5 were answered, based on the analysis findings. The questions are listed in the order that they were addressed during the study.

10.3 2. Where is Web Science Taught?

The first question to be answered was query: 'Where is Web Science taught?' This question was addressed first, as it was necessary to begin with an identification of Web Science taught programmes in order to begin to address the following question: What is taught as Web Science? Research began with an investigation of the list of Web Science teaching intuitions hosted on the Web Science Trust Website[57]. The list of Web Science teaching institutions included on the Web Science Trust at the time of investigation in late 2013/early 2014, included a number of outdated links. When followed up, it appeared that a number of the institutions for which broken links were listed were no longer offering current Web Science taught programmes. Having eliminated the links which were no longer live, a desk survey was then conducted in order to identify additional Web Science taught programmes. This desk survey was supplemented by a Web Crawler, designed to search the .ac.uk domain for a combination of the keywords 'Web' and 'Science'. The results produced by the crawler lead to the identification of an additional Web Science programme missed by the manual desk survey; a Web Science MSc taught by the University of Liverpool. The resulting Web Science teaching intuitions located using a combination of the desk survey and the crawler are shown in Table 51.

[57] https://web.archive.org/web/20161019215623/http://www.webscience.org/web-science/studying-web-science/

Table 51 Web Science Teaching Institutions

Web Science Teaching Institutions
1. Aristotle University of Thessaloniki - Thessaloniki, Greece
2. "British university in Egypt - EL SHEROUK CITY, Cairo, Egypt"
3. Cologne University, Germany
4. Eindhoven University of Technology: Netherlands
5. Georgia Institute of Technology
6. Goldsmiths, University of London
7. Johannes Kepler University Linz
8. Korea Advanced Institute of Science and Technology (KAIST)
9. MIT - Massachusetts Institute of Technology
10. Northwestern University School of Communication
11. Oxford Internet Institute
12. Rensselaer Polytechnic Institute
13. RWTH Aachen University
14. Saint-Joseph University of Beirut
15. The University of Edinburgh: School of Social and Political Science
16. UAH MediaLab, University of Alcalá (Spain)
17. University College London
18. University of Erlangen-Nürnberg
19. University of Koblenz-Landau, Institute for Web Science and Technologies,
20. University of Liverpool
21. University of Southampton, UK
22. VU University Amsterdam; the Network Institute

In addition to the Desk survey and the Web Crawler, the survey of Web Science Academics and students included asked questions aimed at identifying participants' affiliated academic institutions, with the aim of discovering additional institutions, which taught Web Science programmes. This failed to return any new results, and only featured results for institutions which were previously identified by the original desk survey.

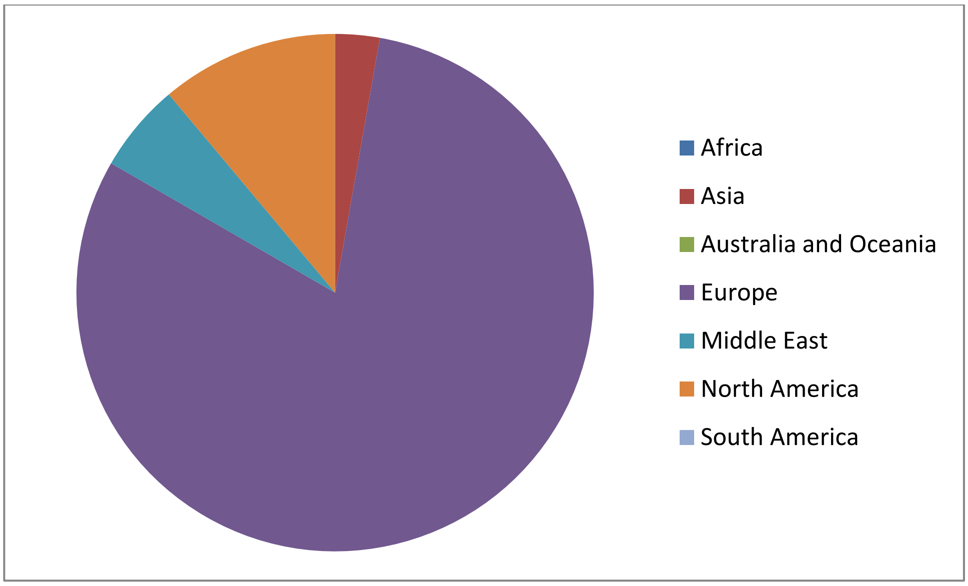


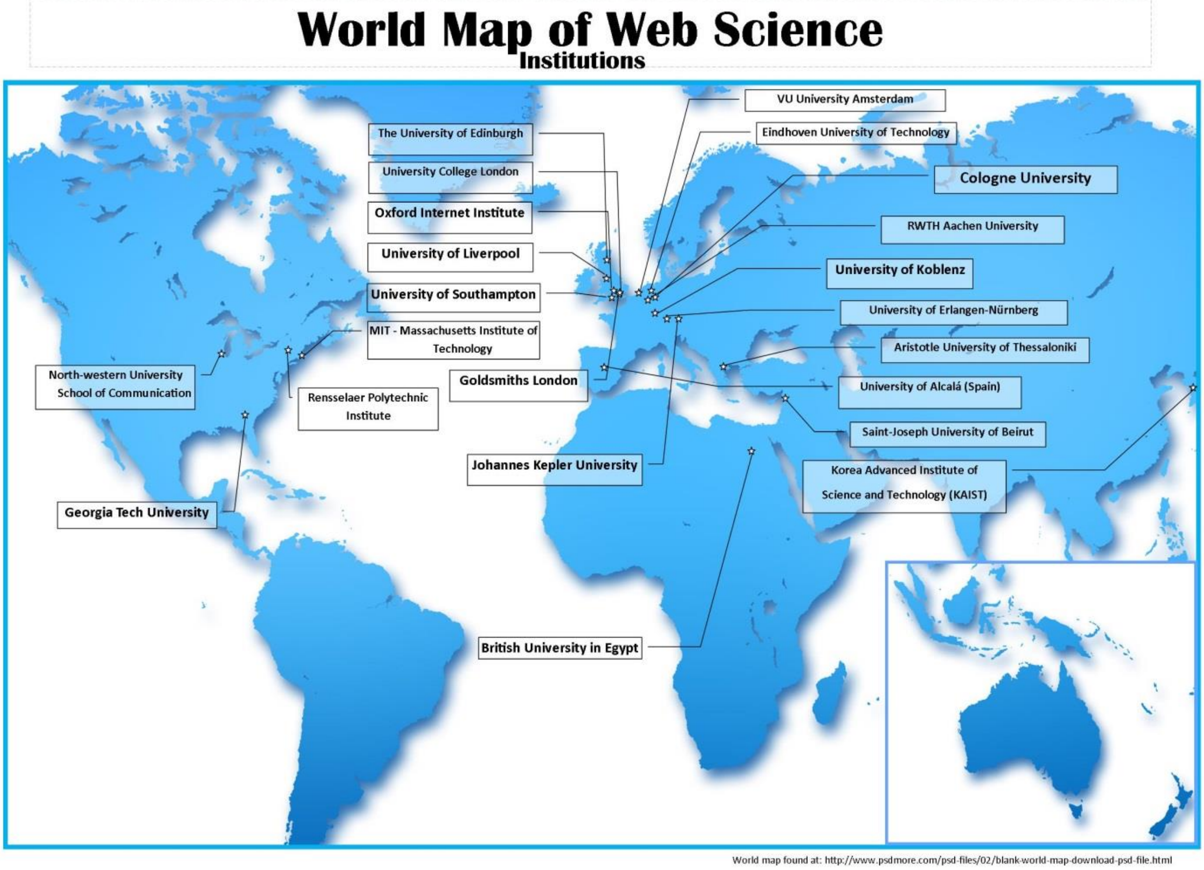
Figure 8 World locations of Participants' Academic Institutions

According to the survey results, the participating Web Scientists originate from the world locations illustrated in Figure 8. This largely reflects the findings of the Desk survey of Web Science Teaching institutions, which indicates that the majority of Web Science teaching institutions are located within Europe.

10.3.1 World Map of Web Science teaching Institutions

The resulting Web Science teaching institutions identified by the desk survey and the Web Crawler were utilised to create a visualisation depicting where the world locations in which Web Science programmes are taught. Due to the scale of the map, a preview of the map can be seen in Figure 8, however, **a bigger scale version with larger text is included in appendix section 16.4.**

Figure 9 Small-scale preview of the world map of Web Science teaching institutions.



A draft version of the map was presented during the Web Education workshop at the 2015 Web Science conference, and was well received.

10.4 2. What is taught as Web Science

The answer to this question was addressed with a more detailed survey of taught modules, which built on the findings of the preliminary desk survey of Web Science teaching institutions. Module data was manually gathered from the Websites of the institutions identified, which offered programmes featuring taught modules. This information was then stored in an Excel spreadsheet. A total of 146 modules were identified from this analysis. The resulting modules identified from the survey of Web Science taught modules were then analysed and coded using Nvivo, identifying key topics taught. The results of this analysis are presented in Table 52. These key topics represent the range of subjects which are taught as part of the Web Science programmes included in the analysis.

Table 52 Key Topics identified from the Web Science Taught Modules Analysis

Name	Sources	Refs	Name	Sources	Refs
Accessibility	3	3	Network Theory	9	16
Artificial_Intelligence	3	3	Operating Systems	1	1
Big_Data	7	9	Performance Management	1	1
Biological Networks	2	3	Politics	5	6
Biology	3	3	Privacy and Trust	8	9
Business	7	11	Programming	4	6
Cloud Computing	5	6	Project Management	3	3
Computer Graphics	2	2	Psychology	3	6
Computer Vision	2	4	Quality Management	1	1
Criminology	1	1	Research Methodology	8	16
Data Mining	7	8	Risk Management	2	2
Digital Copyright	4	5	Security	5	9
Digital Ethnography	1	1	Semantic Web	10	21
Digital Sociology	4	6	Social Networking	11	14
Discrete Mathematics	1	1	Sociology	6	17
Economics	7	11	Software Agents	1	1
Education	1	1	Software Engineering	4	8
e-Learning	1	1	Statistical Analysis	4	9
Entrepreneurship	2	2	Systems Design	4	6
Ethics	2	3	Virtual Environments	1	1
Game Theory	3	3	Visualisation	6	6
General Web Science	5	7	Web 2.0	6	6
Hardware	3	5	Web Analytics	5	7
HCI Human Computer Interaction	3	3	Web Applications	4	4
Hypertext	2	3	Web Architecture	8	12
ICT Information Communications Technology	5	9	Web Design	3	4
Information Systems	2	2	Web Development	5	5
Information Theory	2	2	Web Governance	4	5
Internet of Things	3	5	Web Graph	4	7
Internet Protocols	1	1	Web History	3	3
Law	6	10	Web Languages	10	15
Linked Data	5	7	Web Modelling	5	8
Management	4	6	Web Ontologies	5	6
Marketing	1	2	Web Search	5	7
Media	5	7	Web Society	9	23
Mobile Web	3	3	Web Technologies	7	8
Networks	9	24			

The sources column indicates the number of sources across which a given topic was identified, and the 'references' column indicates the number of times each topic was identified across all resources.

10.5 4. What subjects are studied by Web Scientists in practice?

This study builds on the work of (Hooper et al. 2012)and (Hooper, Dix, et al. 2013), who at the time of their initial study, were only able to include the first three Web Science conferences in their analysis of the coverage of Web Science disciplines within the conferences. As indicated during the review of Web Science literature in chapter 3; (Hooper, Dix, et al. 2013) describe the fact that there has been ongoing discussion within the Web Science community regarding which disciplines are represented within Web Science and how Web Science should be defined as a subject, however, there is currently very little research in this area. A gap, which this thesis aims to address. (Hooper, Dix, et al. 2013) proceed to then explain that it is important to identify the spread and representation of disciplines within Web Science and to ensure that all disciplines within Web Science are accurately represented. The key data source for providing an answer to this research question is the analysis of Web Science conference publications from 2009 to 2017. Publications from the Web Science conference series were manually coded and with the keywords previously identified during the analysis of Web Science taught modules. A series of dedicated paper topics were also identified, representing topics which did not directly correspond to key topics already coded. These paper topics are shown below in Table 53. A list of the manually coded papers from this analysis can be found in Appendix 19.

Table 53 Web Science Paper Topics identified (II)

Name	Freq.	Name	Freq.
Academic publishing	11	MOOCs	2
Algorithms	31	Narrative	14
Archaeology	1	Natural Language Processing	4
Arts and Culture	9	Network Science	47
Bibliometrics	2	Nutrition	1
Cognitive Science	8	Online Advertising	6
Communication Science	23	Online Engagement	41
Computer Science	15	Online Identity	14
Content Analysis	121	Online Offline Community	10
Crowdsourcing	23	Open Access	4
Cyber Bullying	5	Open Data	18
Cybercrime	11	Philosophy	7
Democracy	7	Provenance	2
Demographics	15	Quality Control	5
Digital literacy	4	RDF	10
Digital Native	2	Real-time data	3
Disaster Response	5	Recommender Systems	8
eCommerce	9	Religion	3
Environmental Science	1	Scientific Method	5
Folksonomy	4	Sentiment Analysis	15
Gamification	4	Social Machines	12
Geographic Information Systems GIS	3	Social Science	79
Geography	28	Socialbot	2
Geo-tagging	12	Synthesis Ranking	1
Government	16	Theology	1
Health	15	Topic Modelling	19
Information Theory	1	User Behaviour	145
Journalism	4	Virtual Community	42
Knowledge patterns	10	Web 3.0	2
Machine Learning	2	Web Archiving	25
Microblogging	77	Web Publishing	21

Table 53 shows the paper topics which were identified as part of the analysis of Web Science conference papers. The frequency column indicates the number of times each topic was coded. This range of paper topics, along with the key topics coded, represents the main answer to the question: 'What subjects are studied by Web Scientists in Practice', as the analysis included a total of 357 publications across 9 Web Science conferences, featuring the work of hundreds of researchers from the Web Science community.

The most frequently coded paper topics were content analysis and user behaviour. This corresponds with the observation that the majority of papers focus on the utilisation of some form of content analysis in order to analyse the behaviour of users, in order to address a research question in a given area. While the frequency of coding of paper topics is interesting, it is the diversity of topics present which is of greatest interest. This is because these are the topics which represent the scope of Web Science as it is published, and are therefore representative of what is being researched and published by Web Scientists in practice. These topics contribute to the Web Science Lexicon, in which they are sorted into categories in order to represent the scope of Web Science.

10.5.1 Interviews

A series of interviews were conducted involving Web Science researchers based at the University of Southampton's WAIS Group. While only 16 people participated in the exercise, it still provided valuable data regarding the subjects studied by Web Scientists in practice. The following list shows the range of projects studied by those interviewed.

- Measuring online/offline community
- Demand and innovation for open data
- Longitudinal maintenance of hypertext
- use of the internet in decision making processes in the green party
- Value values exchange process on the web
- The public health analogy in web security
- The impact of community and sociability
- Evaluating the mobile web accessibility of electronic text for print impaired user in higher education
- Socio-technical construction of MOOCS and educator and learning designer roles and practices in higher education
- What do teachers talk about on social media
- Looking at ways to improve the information spread and warnings natural disasters using social media
- Perceptions and use of social media and support of non-formal learning by secondary school pupils
- A technological solution to adults with basic literacy skills
- Archiving social media
- Exploring emergency responders use of social media
- Looking at the evolution of web technology

The subjects studied by the Web Science Researchers interviewed, largely correspond with and support the findings of the analysis of Web Science taught modules and Web Science conference proceedings.

10.6 Observations made from the analysis of Web Science Conference Series

The following section provides an overview and discussion of the findings and observations made of the analysis of the Web Science conference series.

10.6.1 A lack of Web Science Context and Vision within Research?

A trend which is evident from the earlier years of Web Science conference publications, is the fact that a greater number of the early Web Science papers relate their detailed topics of study to the context of Web Science. Because during the first few years of the Web Science conference, the Web Science subject was still 'young', it is possible that authors felt more of a need to define the subject and justify why their research was relevant to Web Science. During the later conferences, papers have a tendency not to relate their work to the wider context of Web Science as a subject, and are generally far more niche in their focus.

10.6.1.1 The original vision for Web Science

The 'mission statement' of the original papers outlining the scope of Web Science, state that the key goal of Web Science is to bring researchers from multiple disciplines together, in order to provide an interdisciplinary perspective, in order to engineer the future of the Web. To better understand the Web, in order to secure the future evolution of the Web, and influence it in a way that benefits society, (Berners-Lee, Weitzner, et al. 2006) and creating "*the Web we want*". While there are a small number of papers within the Web Science conference series so far, which examine the development of the Web and Web Science from this perspective, these are the exception rather than the rule. Papers analysed do focus on the way that the Web can be used to make existing real life problems easier, but on a fairly small scale.

There are few papers which focus upon the bettering of the Web itself, and the realisation of the (Berners-Lee, Weitzner, et al. 2006) vision of engineering "the Web we want". The general theme, especially of the later few years' conferences, is that the majority of papers seem more concerned with using some form of Web analysis to better understand one specific aspect of user behaviour on the Web. While this is still useful and relevant research, it does not specifically address the original vision for Web Science.

Only 43 papers are coded with the key topic 'General Web Science', the topic specified for indicating that papers relate their content to the context of Web Science in some way. There were more examples of General Web Science coded during the earlier years of the Web Science conference. This could be due to the fact that the conference was new and novel, and perhaps consequently, authors felt more of a need to justify the relevance of their work to the wider context of Web Science. Another notable difference in the case of the three initial three years of the Web Science conference, is that during these three years, the conference was not yet sponsored by the ACM. This change in organisation may have had in impact upon conference content; however further, more in depth research in this area would be required in order to determine the extent to which this may be true.

10.6.2 Web Science – the Science of Social Networking?

The majority of papers in the Web Science conference series examine ways in which social media can be used to gain insight about current issues. While there are some examples of work which does not focus on social media based research, the bulk of papers do centre on this area. It is true that social media represents the majority of people's interactions on the Web, therefore it is logical to use this as the basis of research into user behaviour. However, there are less examples of studying how the Web impacts people's actual lives outside the Web. It could be argued that this is because Web Science is concerned with how people use the Web itself. This is true to an extent, however if we are to ensure "the Web we want" as described by Berners-Lee and ensure that the Web has a positive effect on society (Hall), then we also need to consider how the Web effects society in general, not focusing solely on how people are using social media platforms.

This raises the question, has social networking become the primary focus and scope of Web Science research? Does the scope of Web Science research no longer extend beyond social networking research? As stated by (Berners-Lee, Weitzner, et al. 2006) and (Halford et al. 2010), the primary goal of Web Science is to engineer a better web for the benefit of society. Does Web Science research published indicate that the community has lost this focus? During the analysis of conference papers, it was observed that more of the early Web Science conference papers appear to ask philosophical questions about the impact of the Web as a whole, and its effect upon people. If Web Science is to succeed as a field of study and engineer a better future for the Web, (Berners-Lee, Weitzner, et al. 2006) and (Halford et al. 2010) suggest that it is key that the Web Science community does not lose sight of this original mission statement.

One of the reasons for the increasing usage of social media data to analyse real world issues, is likely to be a consequence of the fact that the Web and social media data now presents easily accessible datasets which facilitate social research, which was previously not possible, or substantially more difficult before the advent of the Web. While some of these studies would previously be conducted via manual surveys, it was extremely difficult to gather worldwide data similar to that seen in many of these studies examined in the Web Science conference series, prior to the Web. The Web is a platform connecting the whole world, allowing not only worldwide participation, but also worldwide research in a way previously unseen. A key reason for the focus on social networking, is partly due to the easy accessibility of data from platforms such as Twitter, which are publicly available, and therefore relatively easily to analyse.

10.7 1. What is the extent of the Web Science Subject?

One of the features of qualitative research, is that unlike with quantitative methods, there is no one definitive answer to a given research question. Instead, findings are often subjective. The same is arguably true of Web Science; it is a vast discipline, which is constantly evolving, and while it is possible to provide an example of the extent of the Web Science subject, it is unrealistic to suggest that this will be a 'definitive' answer. It will merely be one example answer to the question.

As previously explained during the review of Web Science literature seen earlier in chapter 3, Web Science is an emerging socio-technical subject which despite existing for over 10 years, is still largely undefined. The most comprehensive attempt to provide a definition for Web Science was the WSSC, however, this resource can no longer be found on the Web Science Trust Website. There are a number of Web Science curricula in existence, and there is also a dedicated conference for Web Science which has run for 9 years. The existence of these resources facilitate a reverse engineering approach to research into what is taught and published as part of Web Science, identifying the scope of the subject in a reverse engineering fashion. The results of this analysis was also supported by surveying and interviewing Web Science academics and students, in order to ascertain people's perceptions of Web Science, as well as the topics which are studied as part of the discipline in practice.

The analysis of Web Science taught modules and papers led to the identification and compilation of a series of key topics (modules) and paper topics (papers) detailing the topics which are most frequently taught Web Science programmes and published as part of the Web Science conference series. The complete combined list of key topics and paper is displayed below in Table 54. This list of topics represents the scope of Web Science according to the results of the analysis described.

As seen during chapter 3, (Hooper et al. 2012) evaluate the two current examples of attempts to depict the Web Science curriculum; the Web Science Cluster diagram and the WSSC. They state that the Web Science Cluster diagram represents more of a "vision" of Web Science than a full and comprehensive reality. They also critique the Web Science Subject Categorization as only providing "*vision and structure, not information on subjects' prevalence within the community*". Research conducted during this thesis addresses this gap, by showing the frequency at which subjects are coded within the analysis. The frequency column of Table 54 indicates the number of times each topic was identified during the analysis. Topics identified are ordered according to frequency, with the most frequently occurring topics beginning in the left-hand column. Topics with a coding frequency of only one or two are omitted for ease of presentation.

Table 54.Combined list of key and paper topics for all modules and papers

Topic Name	Freq.	Topic Name	Freq.
Web Society	231	Geo-tagging	12
Web Analytics	218	Information and Communications	12
Social Networking	192	Security	12
User Behaviour	145	Social Machines	12
Content Analysis	121	Visualisation	12
Data Mining	104	Web Development	12
Digital Sociology	91	Academic publishing	11
Psychology	84	Accessibility	11
Social Science	79	Cybercrime	11
Microblogging	77	Ethics	11
Linked Data	59	Business	10
Big_Data	57	Cloud Computing	10
Privacy and Trust	52	Digital Ethnography	10
Media	50	Knowledge patterns	10
Semantic Web	49	Online Offline Community	10
Network Science	47	RDF	10
Research Methodology	45	Arts and Culture	9
General Web Science	43	eCommerce	9
Networks	42	Artificial Intelligence	8
Virtual Community	42	Cognitive Science	8
Online Engagement	41	Mobile Web	8
Web Modelling	41	Recommender Systems	8
NetworkTheory	36	Web Governance	8
Web Technologies	34	Democracy	7
Politics	32	e-Learning	7
Web 2.0	32	Information Systems	7
Web Graph	32	Philosophy	7
Algorithms	31	Web Design	7
Sociology	31	Digital Copyright	6
Geography	28	Internet of Things	6
Web Architecture	28	Online Advertising	6
Web Search	28	Risk Management	6
Web Ontologies	27	Systems Design	6
HCI Human Computer Interaction	26	Computer Graphics	5
Web Archiving	25	Cyber Bullying	5
Statistical Analysis	24	Disaster Response	5
Communication Science	23	Hardware	5
Crowdsourcing	23	Project Management	5
Quality Management	23	Quality Control	5
Web Publishing	21	Scientific Method	5
Topic Modelling	19	Criminology	4
Web Applications	19	Digital literacy	4
Open Data	18	Folksonomy	4
Government	16	Game Theory	4
Computer Science	15	Gamification	4
Demographics	15	Internet Protocols	4
Economics	15	Journalism	4
Health	15	Management	4
Sentiment Analysis	15	Natural Language Processing	4
Education	14	Open Access	4
Hypertext	14	Programming	4
Marketing	14	Software Engineering	4
Narrative	14	Biology	3
Online Identity	14	Entrepreneurship	3
Web History	14	Geographic Information Systems GIS	3
Web Languages	14	Real-time data	3
Law	13	Religion	3

10.8 The Web Science Lexicon: a definition for Web Science

The topics shown in Table 54 form the basis of the online Web Science Lexicon, a resource outlining the scope of Web Science, designed to provde a form of Web Science encyclopedia.

Table 55 Web Science Lexicon Topic Listing

Web Science Lexicon Topic Listing
Web Technologies
Big Data
Open Data
Semantic Web
Linked Data
RDF
Knowledge Patterns
Web Ontologies
Web Applications
Web Modelling
Web Architecture
Web Languages
Hypertext
Web 2.0
Web Analysis
Algorithms
Content Analysis
Data Mining
Narrative
Network Science
Network Theory
Statistical Analysis
Topic Modelling
Web Archiving
Web Graph
Web Search
Web Society
Academic Publishing
Cognitive Science
Communication Science
Crowdsourcing
Cybercrime
Digital Anthropology
Digital Ethnography
Digital Sociology
Demographics
Economics
Education
e-learning
Health
Human Computer Interaction
Law
Media
Marketing
Geography
Geo-tagging
Philosophy
Politics
Government
Privacy and Trust
Research Methodology
Social Networking
Microblogging
User Behaviour
Online Engagement
Online Identity
Online-Offline Community
Psychology
Virtual Community
Web Publishing

The most frequently coded topics from Table 54 were included in Table 55. The topics shown in Table 55 were stuctured under the three headings, Web Technology, Web Analysis and Web Society. While the topic list viewed in isolation provides an immediate overview of the subjects included within Web Science, each topic in the online version of the Lexicon features a dedicated page including the following: A description of the topic, and further reading, in the form of a list of Web Science conference papers relating to the topic. Also included, are a links to other sub-topics to which the given topic relates.

10.8.1 Lack of Interdisciplinary collaboration within Web Science?

(Halford et al. 2010) state in their 'Manifesto for Web Science', that discipline specific expertise relating to the Web has a tendency to remain within the discipline in which it originates, and that there is little sharing between disciplines, especially between the "embedded binary divide" between engineering and the social sciences. Respondents 7 and 20 of the online survey make the point that they would like to see greater interdisciplinary representation within Web Science courses. One states that: "*Web Science individual research projects within Web Science still rarely transcends the boundaries of one disciplinary perspective*." (see Table 38 Subjects omitted by the WSSC) Therefore, this suggests that there is still a lack of interdisciplinary collaboration within Web Science teaching.

The analysis of Web Science conference proceedings would seem to confirm this. Web Science papers do include examples of research from the perspectives of diverse disciplines such as law, economics and psychology which widen the scope and input of Web Science research beyond it's often Computer Science dominated foundations. However, most Web Science conference papers included in the analysis, even though representing diverse subjects, still generally only address their given topic of research from one disciplinary perspective.

Therefore, it would seem that more still needs to be done in order to promote genuinely interdisciplinary collaborations and sharing of expertise within Web Science if the vision of (Berners-Lee, Weitzner, et al. 2006) and (Halford et al. 2010) is to be realised . Therefore any resource such as the Lexicon, which encourages collaboration between and contributions from different disciplinary areas should be beneficial in achieving the original mission statement for Web Science.

10.9 Contributions

This thesis has delivered the following contributions to the Web Science community.

10.10 Web Science Taxonomy of Key Topics

The Web Science Lexicon is a resource outlining the scope of Web Science, as it is taught, and as it is published according to the Web Science conferences series. The need for the Lexicon is highlighted by the lack of detailed knowledge regarding the scope of Web Science demonstrated by the Web Science students and academics surveyed and interviewed. (Börner et al. 2003) propose that "*each research field can be characterized by a list of the most important keywords*." It is this theory that forms the basis for the proposed Web Science Lexicon, which consists of a taxonomy of topics representing the scope of Web Science. The structured Lexicon topics provide a similar format of resource to the 'body of knowledge' topics outlined within the (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013). The body of knowledge topics within that report, are not designed to give detailed learning outcomes for Computer Science, rather, they are intended to provide an overview of the subjects which represent the scope of the Computer Science curriculum. The Web Science Lexicon is a similar concept, except it includes greater detail than the CS 'body of knowledge' headings.

The Web Science Lexicon resource is included in an online format, suitable for use as part of the University of Southampton's Web and Internet Science department's Website. It is hoped that the Lexicon should serve as an interactive encyclopaedia for Web Science, listing the interdisciplinary subjects which are part of Web Science, and providing a resource page for each, which includes a description of each subject, and related publications for 'further reading' on the subject. The resource could also be edited by anyone with permission, facilitating the addition of new content.

The Lexicon could potentially benefit anyone wanting to know more about the Web Science subject, from academics teaching the subject, to students studying the subject. It also serves as a reference resource for any intelligent layperson wanting to know more about the Web Science. The Lexicon could aid with recruiting students to Web Science programmes, by informing prospective students about the scope and nature of the subject. The Lexicon could also have useful applications for the design of teaching and learning objectives for Web Science material. This could include Universities wishing to begin teaching a new Web Science programme, the redesign of existing Web Science content, or reporting to exam boards. When designing learning objectives, it is beneficial to have access to a 'typical' example of what falls under the scope of a subject, in order to derive learning objectives accordingly, which is the aim of examples such as the ACM recommendations for Computer Science, (The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society) 2013). As there are currently few examples of such guidelines for Web Science, the Lexicon represents a community approach to fill the gap. It serves as an overview of the scope of Web Science, and also lists Web Science related conference material relating to each topic listed.

It is important to recognise that the Lexicon is not a fully watertight definitive definition for Web Science; this would be impractical due to the huge scope and rapid evolution of the subject. The Lexicon serves and one possible example of a definition for the Web Science subject, according to the scope of the Web Science taught programmes and conference material analysed. However, the aim of the Web Science Lexicon, is to serve as a reference tool and starting a point facilitating further discussion regarding the scope of Web Science. It is also a resource which has the potential to be further developed and contributed to by the wider Web Science community.

10.11 A Bottom up Approach to Defining an Emerging Subject

Web Science represents an emerging subject for which there is a lack of definition. There are very few resources available which outline the scope of Web Science, the WSSC comprising the only detailed example. This lack of definition is not only the case for Web Science, it is likely to also likely to be true for other emerging subjects, as such as, for example, the relatively new discipline of Digital Sociology, which is a relatively recent branch of sociology.

This thesis proposes that it is possible to gain an overview of an emerging subject for which limited information is available, by conducting a bottom up analysis of what is taught and published as part of the given subject. Similar to the example of Web Science, there may be no curriculum or subject guidelines available for an emerging subject, yet there may a number of taught programmes. If this is the case, it is possible to gain an insight into the scope of the subject by analysing what is taught and published as part of the given subject in a reverse engineering fashion. While this method is useful for Web Science, it may also be useful and applicable for any emerging discipline.

If a similar methodology were to be reused, the research methods outlined in chapter 6 and 7 could be applied. In summary, the process would involve a desk survey to identify where the subject is taught, followed by a more detailed analysis of module content and material published within the scope of the subject, within related conferences or journals. This process could be implemented with ether qualitative or quantitative methods, however at least some manual qualitative analysis is useful for the researcher to gain an understanding of the subject, and the grounded theory method utilised during this work has the added benefit of incorporating mixed methods.

10.11.1 Datasets

The following data sets were produced during the lifespan of this project.

- Web Science modules survey data
- Web Science paper topic data

The analysis of Web Science taught modules was recorded in a spreadsheet, and therefore can easily be converted to a coma separated file for storage in the University of Southampton Web Observatory, where it can then be accessed and utilised by other research projects.

The results of the analysis of Web Science conference publications were originally produced in a Word document. Therefore in order to make them machine readable, it is possible to code the data in JSON, allowing the data to be easily queried. JSON was chosen over XML, due to the fact that JSON facilitates arrays facilitating the inclusion of lists of multiple key and paper topics within each paper entry. Figure 10 shows an example of a JSON coded paper entry.

```
{

"WebSci2016":[

 { "Title":" Beyond the MOOC platform: gaining insights about learners from the social
web",

            "KeyTopic":["e-learning,”," Web Analytics "," Web Society",”Web Society”,”Data
Mining”,” Big Data”,” Research Methodology”],

            "Papertopic":["MOOC,”," Content Analysis"," User Behaviour ",”Online
Engagement”, ” Gamification”]

            "AuthorKeyword":["Human-centred computing,”,"User models","Social networks",”
Applied computing”,” Learning management systems”],

]

}
```

Figure 10 Example of a JSON entry for a Web Science paper

This dataset will also be submitted to the University of Southampton Web Observatory, so that it may be utilised in other research projects beyond the life of this study. Once uploaded, both datasets can then be utilised by anyone wishing to conduct further Web Science related research.

10.12 A Framework for Web Science: Comparison

As part of the conclusions for this study on Web Science, it is useful to consider, how does the original vision set out for Web Science by Berners-Lee et al compare with the findings of this research and the Web Science Lexicon? The following section compares the findings of this study with the original proposals outlined by (Berners-Lee, Weitzner, et al. 2006) in their original 'Framework for Web Science' work. A brief summary of each chapter of (Berners-Lee, Weitzner, et al. 2006) is provided, along with some key Lexicon topics which relate to the content of each chapter.

Table 56 illustrates that the high-level headings utilised to structure the content of the Web Science Lexicon, largely correspond with the key section headings presented within (Berners-Lee, Weitzner, et al. 2006). The first section of (Berners-Lee, Weitzner, et al. 2006) 'The Web and its Science' is essentially an overview section. Therefore, it is the following four sections which are of most interest. Each of the WSSC headings have been placed within the table adjacent to the corresponding (Berners-Lee, Weitzner, et al. 2006) heading that they relate to. Consequently, one can see that 'Engineering the Web' corresponds with 'Web Technologies', as the 'Engineering the Web' section of (Berners-Lee, Weitzner, et al. 2006) largely focuses upon proposals relating to the technological infrastructure of the Web. The following two headings of (Berners-Lee, Weitzner, et al. 2006) directly correspond with the two Lexicon headings, Web Analysis and Web Society. The exception within this example, is the final key heading of (Berners-Lee, Weitzner, et al. 2006), 'Web Governance and Standards'. 'Web Governance Security and Standards' is not directly represented by a high-level heading within the Lexicon, however, it is represented within several sub-headings of the Lexicon, under the higher level, 'Web Society' heading. Also, as discovered during the following analysis, the 'Web Governance Security and Standards' potentially has more content which directly relates to the Lexicon heading 'Web Society' than the 'Social Aspects' chapter of (Berners-Lee, Weitzner, et al. 2006).

Table 56 Framework for Web Science and Lexicon Heading Comparison

Key Sections of the Framework Paper	Corresponding Lexicon Headings
The Web and its Science (Overview)	
Engineering the Web	Web Technologies
The Analysis of the Web	Web Analysis
Social Aspects	Web Society
Web Governance Security and Standards	

10.12.1 A Framework for Web Science – Sections

The following sections provides a brief summary of each key section of (Berners-Lee, Weitzner, et al. 2006) and how it is reflected within the Web Science Lexicon content.

10.12.1.1 Engineering the Web vs Web Technologies

This section discusses the fact that it is the goal of Web Science to track the development of the infrastructure of the Web, "determining which innovations are good", with the aim of influencing the Web's development in a positive way. Key aspects of Web Technologies and Architecture reviewed within this section include: The Semantic Web, Web Services, P2P and grid computing. This chapter content is reflected within the following example Lexicon sub-topics:

Table 57 Web Technology Lexicon Topics relating to 'Engineering the Web'

Web Technology Lexicon Topics	Freq.
Semantic Web	49
Web Ontologies	27
Web Languages	14
RDF	10
Cloud Computing	10
Recommender Systems	8

Judging by the coding frequency by which these topics were identified within the analysis of taught programmes and conference publications, there is less of a focus on these technological aspects within the Lexicon topics than might be expected, given their importance within (Berners-Lee, Weitzner, et al. 2006). However, this might reflect the fact that the technology, such as for example, 'Cloud Computing', is more well established and is now considered less of an innovation warranting research in itself, and more of a tool to be utilised in other research. These Web Technology topics may also represent areas which should be the focus of more research, in order to ascertain whether they are being utilised in the best possible way for the benefit of the wider Web and Society.

XML is a topic frequently referred to by (Berners-Lee, Weitzner, et al. 2006). While not directly referenced within the Lexicon, it is represented by the broader topic, 'Web Languages'. Also mentioned prominently within (Berners-Lee, Weitzner, et al. 2006) under the 'Engineering the Web' heading, was the topic 'Folksonomies'. This was discussed by (Berners-Lee, Weitzner, et al. 2006) in relation to Web Ontologies, with the differences between the two compared. While the term 'Web Ontologies' is fairly prominently coded within the key and paper topics, 'Folksonomy' is not represented by that specific term. It is instead represented by the term 'Social Networking', and is located under the 'Web Society' Lexicon heading. This was the case because the majority or Web Science publications tagged with 'Social Networking' refers to the behaviour of Social Network users, as opposed to the underpinning technology. This may reflect a shift in the environment during time between the publication of (Berners-Lee, Weitzner, et al. 2006), to the present day. Similar to the Cloud Computing example, it could be the case that the technology itself is more well established, and less of a focus for research. It is conceivable that researchers are now more interested with how the platform is being used, as they are already familiar with the concept of how the underpinning technology works.

10.12.1.2 The Analysis of the Web vs Web Analysis

The Web Analysis chapter is largely concerned with mapping and analysing the structure of the Web and charting its development. This features methods such as graph theory and mathematical methods for describing the structure of the Web. Web Graph and Network Science are both frequently coded topics, indicating that they are well represented within the Lexicon in terms of related publications.

Table 58 Web Analysis Lexicon Topics relating to 'The Analysis of the Web'

Web Analysis Lexicon Topics	Freq.
Network Science	47
Web Modelling	41
Network Theory	36
Web Graph	32
Web Search	28

There are a considerable number of Web Science publications focusing on the Network and Graph theory aspects of the Web, suggesting that this area is still considered key in terms of Web Science research. (Berners-Lee, Weitzner, et al. 2006) describe Graph theory, or 'Web Graph' as the best-known paradigm for Web analysis. Its popularity is reflected within the Lexicon subjects; as one methodology, a coding frequency of 32 (Table 58) shows that it features prominently in the Web Science publications analysed. Many Web Science publications analysed also feature some form of Web Network analysis, represented within the Lexicon as 'Network Science'. Also included within (Berners-Lee, Weitzner, et al. 2006), 'The Analysis of the Web' section is 'structure-based search' which refers to Web search features such as the Google page-rank algorithm, which is represented in the Lexicon with the topic 'Web Search'.

10.12.1.3 Social Aspects vs Web Society

The 'Social Aspects' chapter of (Berners-Lee, Weitzner, et al. 2006) focuses upon the exploration of how the 'social requirements' of the Web can 'feed into engineering decisions'. During the introduction, it was stated that governance of the Web in order to benefit society was separated and included within a final dedicated chapter. Despite what the heading 'Social Aspects' would seem to suggest, the majority of this chapter still appears to place more emphasis upon the technological platform, as opposed to the impact of the Web upon society itself.

A key focus of this chapter of (Berners-Lee, Weitzner, et al. 2006) is the logic and usability of Web structure, which seemingly relates more to Web Analysis and Technology and it does to social issues on the Web. The focus appears to be largely on Human Computer Interaction in terms of the technical structure and usability of the Web. This is likely to be a reflection of the fact that this paper originates before Social Networking giants such as Facebook became prominent. Therefore, the authors were more concerned with the Human Computer Interaction and usability of the technology as opposed to online interaction between users.

Table 59 shows the Lexicon topics which relate to topics mentioned within this section of (Berners-Lee, Weitzner, et al. 2006). There is a reference by (Berners-Lee, Weitzner, et al. 2006) to 'Web Sociology', which is represented within the Lexicon as 'Digital Sociology'. Digital Sociology is a subject which was very prominent within the key and paper topic analyses, with a coding frequency of 91. Online Gaming is referred to by (Berners-Lee, Weitzner, et al. 2006) as an increasing example of growing online communities. This is reflected within the Lexicon Topic 'Virtual Community', and also 'Social Networking'. The fact that the topics of Web Sociology and online interaction are discussed less than the usability of the Web as an interface, is likely to be a reflection of the fact that this paper was written over a decade ago. Were it written in the present day, more emphasis would likely be given to online communities and the user interaction facilitated by the technology, as opposed to focusing largely upon user interaction with the technology itself.

Table 59 Web Society Lexicon Topics relating to 'Social Aspects'

Web Society Lexicon Topics	Freq.
Social Networking	192
Digital Sociology	91
Virtual Community	42
Human Computer Interaction	26

It is proposed during the conclusion of the section that "successful participative technologies such as RSS, folksonomies, wikis and blogs" should be explored, in order to understand the common factor which makes them successful, and also what interface features are important. This is a foreshadow and in some ways a prediction of things to come, with the future evolution of social networks. It would seem that the following (Berners-Lee, Weitzner, et al. 2006) section on 'Web governance' relates more directly to the Lexicon version of Web Society than this 'Social Aspects' section, despite the implication of the section title. Once again, this may be a reflection of the fact that (Berners-Lee, Weitzner, et al. 2006) was written over 10 years ago.

10.12.1.4 Web Governance Security and Standards

The 'Web Governance Security and Standards' section of (Berners-Lee, Weitzner, et al. 2006) was the only example not represented by a high-level heading within the Web Science Lexicon. However, as indicated during the previous section, the content of this section of (Berners-Lee, Weitzner, et al. 2006) potentially relates more closely to the Web Society Lexicon heading, than the 'Social Aspects' (Berners-Lee, Weitzner, et al. 2006) section.

When outlining their vision for Web Science, (Berners-Lee, Weitzner, et al. 2006) include 'Web Governance Security and Standards' as their final key section for creating a science of the Web. This section mainly relates to providing governance and standards for use of the Web. The recommendations made within section of the document are still surprisingly accurate and relevant given that this document was written before social media giants such as Facebook came to prominence. The following Table 60 lists the key and paper topics identified within the analysis of Web Science taught programmes and conference publications within this study, which are relevant to the subject areas mentioned within the 'Web Governance Security and Standards' section of (Berners-Lee, Weitzner, et al. 2006).

Table 60 Web Society Lexicon topics relating to 'Web Governance Security and Standards'

Web Society Lexicon Topics	Freq.
Privacy and Trust	52
Web Publishing	21
Online Identity	14
Security	12
Cybercrime	11
Web Governance	8
Cyberbullying	5

Privacy and Trust is one of the more frequently coded topics within the Lexicon key and paper topic list. Table 60 shows that it was identified within a total of 52 sources. With a coding frequency of 8, the term 'Web Governance' itself is not one of the most frequently coded topics. However, it is also reflected within related Lexicon topics such as Privacy and Trust, which is well represented at a coding frequency of 52. (Berners-Lee, Weitzner, et al. 2006) suggest that Governance of the Web is an important issue which needs to be given more prominence and attention. They highlight the fact that users should have control over their personal data online, and that this control needs to be simple enough for the 'untrained' or 'uninterested' to utilise easily, but still comprehensive enough to provide users with protection online. They also suggest that such measures for data control should not interfere with the general Web usage experience. They also highlight the difficulty with imposing standards within such a large decentralized structure such as the Web. They suggest that this issue of Web Governance should be a key focus within Web Science, and this is still very much the case for the current Web.

(Berners-Lee, Weitzner, et al. 2006) suggest that a lack of 'sophisticated information controls' have hindered people who want to share information via the Web. Given that information sharing is now very commonplace, it has either it has become less of an issue over 10 years later, with more standards in place for protecting information online, which is to some extent the case; or people have simply become more culturally accustomed to the idea of sharing content via the Web, as the Web is more mature, and has become a more prominent feature in people's lives for longer. More everyday activities are now conducted via the internet, such as banking and shopping, education and so on. For example, many academic publications such as conference proceedings are now shared via the Web, whereas previously they would only be accessible in paper form. Procedures have been put in place for protecting the use of these publications, such as access via an academic institution email login, which is consistent with the (Berners-Lee, Weitzner, et al. 2006) suggestion for the protection and appropriate use of content. The recommendations that (Berners-Lee, Weitzner, et al. 2006) set out for privacy and security of personal data is also consistent with the introduction of GDPR regulations within Europe (Information Commissioner's Office 2018). However, this is only a relatively recent introduction, suggesting that governments have been relatively slow to adopt and heed the suggestion of the (Berners-Lee, Weitzner, et al. 2006) original agenda for Web Governance and standards for data security. As well as ensuring the protection of users' personal data on the Web, (Berners-Lee, Weitzner, et al. 2006) are also concerned with the behaviour of users themselves on the Web. This is in order to ensure that the Web is a safe environment in which people can interact and share content without fear of bullying or misuse of their content. While Web giants such as Facebook do take measures to ensure that fraudulent content or inappropriate comments can be reported, privacy and trust remains a complex and prominent issue within social media and Web usage today, as also reflected by the Lexicon topics listed.

10.12.2 Key changes in Web Use since the Publication of the Framework Paper

(Berners-Lee, Weitzner, et al. 2006) state the fact that most Semantic Web data is sitting "within stand-alone repositories". They propose the need for data to be more connected and integrated in order to facilitate easier access and reuse. It's arguable that this is something which has changed since the publication of (Berners-Lee, Weitzner, et al. 2006). A considerable amount of data is now stored within online cloud-based repositories. Although not all of this is publicly accessible, many initiatives such as the Open Government data intuitive in the UK, (HMG 2012) are promoting the accessibility of data.

The use of Social Networking platforms has also greatly increased since the publication of (Berners-Lee, Weitzner, et al. 2006). While Facebook was originally founded in 2004, when (Berners-Lee, Weitzner, et al. 2006) was published in 2006, it was still relatively new. Twitter was also founded in 2006, so that had not yet come to prominence. Despite this, (Berners-Lee, Weitzner, et al. 2006) still accurately predict the need for the development of Web Governance designed to protect users of such platforms. (Berners-Lee, Weitzner, et al. 2006) represents a comprehensive effort to outline a manifesto for the development of the Web. Despite being written over a decade ago, their predictions and proposals for Web Science are still largely relevant for the Web today.

10.13 Conclusions

This thesis has provided an overview of current Web Science related content, evaluating existing attempts to define Web Science, and the problems faced when using traditional methods to define such a rapidly evolving field.

This thesis has proposed a research method suited to the study of Web Science. This involves utilising bottom up approach, in order to evolve a definition for Web Science from available taught module data, and Web Science conference material. This reverse engineering approach is a departure from traditional methods of research, however, it is arguably more suited to Web Science, due to the fact that it is an emerging subject with little current definition. Web Science is also an example of a rapidly evolving field, and traditional methods of defining curriculum guidelines would potentially become outdated by the time the evaluation process were completed.

The ultimate aim of this project was to emerge a framework for a clearer definition of Web Science. This was realised with the creation of the Web Science Lexicon, a taxonomy of topics representing the scope of Web Science as it is taught and as it is published within the context of the Web Science conference. The purpose of the Lexicon is to provide a resource for anyone wishing to study or teach Web Science, or the intelligent layperson wishing to gain an overview of the subject. The online version of the Lexicon also features the potential for future crowdsourcing of additional content and input from the wider Web Science community.

The Lexicon represents just one example of a definition for the field of Web Science, the scope of which is vast. While the Lexicon is just one possible representation of Web Science, it serves as a starting point for further discussion. It occupies a current gap in Web Science research, where, as (Hooper et al. 2012) point out, little work exists.

11 Future Work

11.1 Introduction

The following section outlines proposals for future work, and areas for further development which were considered to be beyond the scope or time available for this thesis.

11.2 Crowdsourcing Additional Lexicon Content

One of the reasons behind the idea for hosting the Lexicon as part of the University of Southampton's WAIS Website, is in order to facilitate the crowdsourcing of additional content by Web Science academics and students within the WAIS research group. The Drupal site on which the Lexicon is hosted, can be accessed by anyone provided with a login. They would then be able to add additional Lexicon topics or headings, as well as page content.

An initial experiment involving the crowdsourcing of content for the Lexicon was conducted during the 2016 WAIS Fest event. A quick Google survey was constructed; participants were asked to select an area of interest to them from a list of key and paper topics, and then to write short descriptions about their chosen subject. While this yielded a number of promising responses, there were only a small number of contributions. However, it did demonstrate the feasibility of using a survey to gather content, as an alternative to participants directly editing the Website. Therefore, the use of a survey to manually crowdsource additional Lexicon data, could be a viable alternative for collating Lexicon content in the future. This would be most effective if promoted as part of an event such as WAIS Fest, at a conference workshop, or as part of a lecture exercise for the Web Science MSc students.

11.3 Incorporating Nvivo Project functionality in the Lexicon using JSON

The analysis of Web Science conference papers resulted in a set of coded papers stored within an Nvivo project. If a particular topic is of interest, it can be searched for within the Nvivo project by selecting the topic, for example, 'Digital Sociology' or 'User Behaviour' and asking Nvivo to search for all coding relating to the given topic. Nvivo will then produce results showing all papers which are coded under the search topic. This is a valuable resource, but the functionality it is currently limited to within the Nvivo project.

It would be beneficial to incorporate a similar feature within the Web Science Lexicon, whereby if a topic is selected, for example, 'Digital Literacy', all the papers relating to that subject are automatically included in the page describing that subject. Currently papers are only manually listed in subject content pages, but given that the Web Science paper analysis has already tagged Web Science papers with the relevant topics, automation would be a relatively straightforward addition. The results of the analysis of Web Science conference publications formatted in JSON, could be utilised as part of the Web Science Lexicon in order to realise some of this potential of the Nvivo project within the Lexicon. The paper analysis coded in JSON format would facilitate quires based upon keywords. This could then allow all papers coded with a specific keyword to be automatically searched.

11.4 Widening Research to include additional Publications

It is recognised that the perspective of the content included within the Web Science Lexicon is biased to the outlook of Web Science as it is presented by the Web Science conference, as this was the one of the key data sources used to gather information on Web Science content. The process of manual analysis was time-consuming; therefore the manual analysis of Web Science publications was limited to the Web Science conference series. This was because the Web Science conference is the recognised conference dedicated to Web Science, and has been running since 2009, providing a comprehensive data source. Were there opportunity for further study, it would be beneficial to widen the scope of research to Include additional data sources, such as the WWW conference and Journal of Web Science, as these are additional sources of publications which are relevant to Web Science.

11.5 Further Investigation of Related Disciplines

The scope of investigating the overlaps between Web Science and related subjects is huge, and could provide enough material for an entire further study dedicated to such an analysis. The following subjects were identified during the process of reviewing literature, as demonstrating similarities with Web Science, and therefore may warrant further study.

11.5.1.1 Digital Sociology and Internet Science

Digital Sociology is an example of a discipline which demonstrates significant overlap with Web Science, as demonstrated by the Goldsmiths MSc in Digital Sociology, which was identified and included as part of the analyses of Web Science taught programmes, as it was deemed to include a great number of teaching topics relevant to Web Science. Research also identified a CTD Digital Sociology research group based at Cardiff University[58], specialising in Digital Sociology research. A repository containing research papers published by an associated Social Data Science Lab was also identified[59]. Given the opportunity for further study, it would be beneficial to conduct a similar analysis of these papers to that conducted of the Web Science conference series, and compare the findings of the two analyses, in order to determine the extent to which topics covered by the publications of the two disciplines overlap. The same is true for the papers published within the Internet Science conferences. A more detailed analysis and comparison of Internet Science and Web Science may also yield interesting results indicating the overlap between the two disciplines, as well as the differences between the two.

Other possible disciplines warranting investigation, include Network Science, as this was another discipline identified in a study by (Tiropanis et al. 2015) into the overlaps between Web Science and related disciplines. Human Computer Interaction (Hooper & Dix 2012), Digital Anthropology, Information Science and Communication Science also potentially demonstrate overlaps with Web Science, and the nature of this could be investigated further.

---

[58] http://www.cardiff.ac.uk/research/explore/research-units/digital-sociology-research-group

[59] http://socialdatalab.net/publications

11.5.2 Web Science at Research at Interdisciplinary Centres for Doctoral Training

Results of the survey of Web Science academics and students suggested that there might be a correlation between Web Science, and the work of other interdisciplinary focused centres for doctoral training such as Highwire CDT at Lancaster University[60]. The correlation between Web Science, and related subject research conducted by institutes such as these, is worthy of further study.

11.5.2.1.1 Research Methods employed by Web Scientists

(Halford et al. 2010) suggest that Web Science requires a different research approach to that of standard subjects. They advise that: "*if we are to follow the all actors implicated in the web we need to adopt both interdisciplinarity and mixed methods and open up web science to the ontological, epistemological and methodological possibilities offered by the social sciences and humanities*." Future work could constitute of a more detailed study to discover the methods used by Web Science researchers. Some respondents of the survey and interview suggest that Web Scientists fail to utilise enough interdisciplinary methods, and that they focus on one subject and its associated methods, as opposed to incorporating interdisciplinary methods and taking a wider view of the subject.

11.6 Web Science ebook

The idea for the creation of a Web Science ebook emerged during the intermediate stages of this project. During the 2015 Web Science Education workshop, run during the 2015 Web Science conference in Oxford, discussions between members of the Web Science community lead to the suggestion that it would be beneficial for the community to contribute to a Web Science textbook. The difficulty faced by such an effort, is the fact that Web Science is a rapidly evolving subject, and the time taken to contribute material and edit a standard textbook is considerable. Material would therefore become outdated soon after publication. In the case of rapidly evolving subjects such as Web Science, a different approach to that of traditional publications would therefore be beneficial.

The concept of creating a Web Science ebook was inspired by (Cohen & Scheinfeldt 2010) 'Hacking the Academy'; a successful attempt to crowdsource the content for an ebook within a week. Material was contributed by numerous authors, and then edited together to create the final ebook. The ebook format allows for self-publication, which significantly speeds up distribution of the completed publication. The proposed format of the Web Science ebook was also inspired by the concept of (Newman 2012); an Amazon Kindle ebook, which includes a series of questions relating to computer gaming. Following each question, there are a set of responses from participants, each describing their perspective and experiences in answer questions. A similar approach could be used to crowdsource a Web Science ebook with contributions from Web Science students. This was the initial idea behind the series of interviews of Web Science students, however, not enough responses were obtained for a full ebook. However, the idea for the creation of a crowdsourced Web Science eBook is still one which has potential, and could be investigated further. The concept of crowdsourcing content is an excellent way in which to collate expert knowledge about Web Science, and could potentially be achieved in a much shorter timescale than a standard academic textbook.

[60] http://highwire.lancaster.ac.uk/

11.7 Future potential of the Web Crawler

As outlined in chapter 7.2, a simple Web crawler was created to monitor the .ac.uk domain for a combination of the keywords 'Web' and 'Science'. This initial prototype was effective, but fairly limited in its scope. The idea has potential to be expanded and developed further, in order to facilitate searching for the keywords 'Web' and 'Science' worldwide. The search could also be widened to include the key and paper topics identified during the manual analyses. However, either widening the scope to include worldwide searches or additional keywords, would require considerable computing power.

11.8 A Community Approach to Defining Web Science

An alternative approach to defining the scope of Web Science could involve the identification of a network of academics who are conducting Web Science research in practice. This is an alternative approach which has not yet been attempted by any other research into Web Science. (Bastow et al. 2014) propose that studying the number of academic staff and research students associated with Social Science subjects can be a method of ascertaining the size and prominence of a discipline. Research by (Hooper, Bordea, et al. 2013) focused on the identification of subject-based communities within Web Science according to the prevalence of publishing for difference subject areas within the Web Science conference. However, this did not extend to the identification of the specific members of those communities.

Identification of the academics associated with Web Science research could provide a way in which the discipline could be mapped. For example, the number of academics associated with a specific area of Web Science could provide an indication of a subject area's prevalence within Web Science. It should be possible to identify interdisciplinary connections between Web Science and related disciplines, and the disciplines that most commonly co-author content could be identified. So if for example, an academic publishing a Web Science paper originated from a geography department or disciplinary background, then this could indicate a connection between Web Science and geography. The paper 'Predicting Trust Relations Within a Social Network: A Case Study on Emergency Response' (Vedula & Shalin 2017), is an example of a paper from the 2017 Web Science conference which features interdisciplinary collaboration. This work features a collaboration between two academics from the Department of Computer Science and Engineering at Ohio state University, and an academic from the Department of Psychology at Wright State University. This provides evidence of interdisciplinary collaboration within the Web Science community, originating from Computer Science and Psychology backgrounds. Additionally, although neither of these institutions feature a dedicated Web Science department, this paper is evidence of Web Science related research being conducted at these institutions. Therefore, institutions conducting Web Science research can potentially be identified from the academics involved, as in addition to institutions which specifically identify themselves as researching or teaching Web Science.

11.9 Research Missing from Web Science?

In addition to analysing Web Science taught programmes or content published as part of Web Science, it may be beneficial to consider what additional data and research is available within the public domain, which might be relevant to Web Science. There is undoubtedly research and data which is relevant to Web Science, but is not represented in the Web Science conference series. The following sections consider some potential sources of additional data and research which may be applicable to a future analysis of Web Science. This includes Web Observatory datasets, Open Government Data, and commercial data and research which may not be available within the public domain, such as user research conducted by Google and Facebook. While such information is not necessarily freely accessible for analysis, its existence and relevance to Web Science could still be taken into consideration.

11.9.1 Web Observatories and Open Government Datasets

Research project datasets such as the examples which are deposited within the University of Southampton Web Observatory[61], could be relevant to Web Science. Many of the datasets contained within such Web observatories originate from projects which research some aspect of Web usage. Freedom of public information datasets, including Open Government Data (HMG 2012) collated and made public under the Reuse of Public Sector Information Regulations 2005, etc. could also provide datasets which could contribute to a definition of Web Science. While Open Data itself is a key topic relevant to Web Science, not all of these datasets made available are necessarily directly related to the scope of Web Science. For example, road traffic data regarding the use of a motorway is not necessarily directly relevant to the Web Science curriculum. However, if this this was then applied to a Web-based tool for monitoring motorway usage such as that described in a case study within (HMG 2012), it could become relevant.

Some government datasets may be directly relevant to Web Science, especially those which may in some way relate to society's utilisation of online resources. (Davies & Frank 2013) refer to a study which examines the use by the public of online government resources. This would be an example of research which would be directly relevant to Web Science, as it relates to society's use of the Web. The lines between whether a research project is relevant to Web Science or not are potentially blurred, and determining specific criteria for deciding whether a project is relevant or not would be a key part of any future data driven analysis of Web Science.

11.9.2 Google and Facebook Data

Many examples of Web Science research conducted and published within the Web Science conference, consists of analyses of Web data, for example, Twitter data, in order to ascertain the motivations or opinions of Web users relating to a specific subject. Just one of many examples, is (Bhattacharya et al. 2012), which utilises Twitter data to measure users' levels of trust in online content. Twitter data is the dominant form of social networking data utilised, because the majority of Twitter accounts are public. However, internet giants such as Facebook and Google collect and Google store millions of gigabytes of data on individuals' Web usage. This data is largely unpublished and unavailable for research purposes, and is retained by Facebook and Google for advertising purposes. This data can in theory be requested and downloaded by the individual themselves, however, it is largely unavailable for general research purposes.

11.9.3 What data do Google and Facebook Collect?

As highlighted by (Engber 2015) and (Schmidt 2018), in practice Facebook and Google gather far more data than the information that users choose to publish. Facebook collects a huge amount of data on users, beyond what they knowingly upload to the site. For example, Facebook collects time and geolocation information on when and where users login to the site. When users login to

---

[61] https://webobservatory.soton.ac.uk/

Facebook via an app, they collect information on the type of device used. The Facebook like button offered on websites also provides Facebook with an indication of what outside websites users interact with. These are just some example of the ways in which Facebook collates information on its users. Google is identified as the biggest collector of individuals' personal web usage data. They gather information which far transcends that of mere web searches. A key example is the fact that Google utilises the Android OS used on millions of phones to collect data on individuals. Information they gather on individuals includes: location of users, email content, contacts, calendar usage, music preferences, videos watched, and much more.

11.9.3.1 Facebook Data Research

Facebook locked down it's API in 2015 in order to protect personal data. In 2016, user profile data was leaked[62].Third party apps originally had access to Facebook data such as the friends of app users. This lead to the Aleksandr Kogan scandal, which allegedly lead to the data from 270,000 Facebook users being leaked to a marketing company 'Cambridge Analytica' which then allegedly targeted voters with political advertising during the 2016 presidential election. This scandal lead to Facebook restricting their API and research to only sponsored researchers[63], creating a big gap in big data research.

In 2018, Facebook made limited information available to independent researchers, with the purpose of monitoring social media usage in regards to elections[64]. This data made available by Facebook included a constantly updated list of all public links shared on Facebook. Research utilising this data must be pre-planned, and permission must first be sought from an anonymous peer review panel. This unitive was due to Facebook itself wanting help with analysing the data, and is just one intuitive specifically targeted at understanding social media usage during elections. Despite this example, Facebook remains highly restrictive in terms of allowing access to data for research purposes. Although this intuitive opened up research a little regarding this specific research topic, it still does not permit all research opportunities.

11.9.3.2 Example of study involving Facebook Data

Despite Facebook limiting access to user permission data via it's API, some research is still possible. A study by (Dragan & Zota 2017) explains how it is still possible (as of 2017) to gather Facebook user data by utilising the Facebook graph API v2.0. They explain that in order to gather information utilising this interface, user permission is required. Therefore, it is still possible to run surveys, however the user must first consent to participate, and allow app permission. If users are then provided with the facility to invite friends to participate in the study, this can then be used to determine their connections with other Facebook users. This confirms that some studies utilising Facebook data are still possible, although limited.

11.9.3.3 Conclusions

The amount of personal data collection by Facebook and Google raises series ethical concerns, and is itself an important justification for the need for Web Science. Identification of this data represents not only a key research topic for Web Scientists, the data itself also represents an aspect of Web

[62] http://theconversation.com/facebooks-data-lockdown-is-a-disaster-for-academic-researchers-94533

[63] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EIAACS

[64] https://www.cnet.com/news/zuckerberg-announces-facebook-election-research-commission/

Science in terms of content, as it relates to society's use of the Web. The relevance of social networking data to Web Science can also be confirmed by the number of Web Science conference papers tagged with the key topic 'social networking', as shown in Table 54 of section 10.7. All examples of the data gathered by Facebook ad Google constitutes information on the way in which Web users, members of society, utilise the Web. This is inherently part of Web Science. Even if the information that Facebook and Google collect is not publicly available for research, its existence should still be considered in relation to Web Science, as it does exist, and represents a key part of society's interaction with and via the Web.

12 References

Ackland, R., 2013. *Web Social Science: Concepts Data and Tools for Social Scientists in the Digital Age*, Sage Publications.

Akkermans, H. et al., 2011. Is (Web) Science Ready for Empowerment ?

Alise, M., 2008. *a Comparison of Groups in the Biglan Classification Scheme*.

Alise, M.A., 1980. *DISCIPLINARY DIFFERENCES IN PREFERRED RESEARCH METHODS: A COMPARISON OF GROUPS IN THE BIGLAN CLASSIFICATION SCHEME*,

Arbor, A., Teng, C. & Lin, Y., 2012. Recipe recommendation using ingredient networks.

Baez, J., Network theory Presentation. Available at: http://math.ucr.edu/home/baez/econ.pdf.

Bargh, Catherine; Scott, Peter; Smith, D., 1996. *Governing Universities. Changing the Culture?*,

Barker, J., Basic search tips and advances Boolean explained. *Teaching Library*, pp.1–2. Available at: http://learning.fresno.edu/file.php/3385/Boolean.pdf.

Bastow, S., Dunleavy, P. & Tinckler, J., 2014. The social sciences in modern reserch. *The Impact of the Social sciences: How academics and their research make a difference*, p.344. Available at: https://uk.sagepub.com/en-gb/eur/the-impact-of-the-social-sciences/book241492#description.

Bazan, S.B. & Vafopoulos, M., 2013. The Web Science Curriculum at Work: The Digital Economy Master Program at USJ-Beirut. In *Proceedings of the 5th Annual ACM Web Science Conference*. WebSci ’13. New York, NY, USA: ACM, pp. 19–23. Available at: http://doi.acm.org/10.1145/2464464.2464477.

Becher, T. & Trowler, P., 2001. *Academic tribes and territories : intellectual enquiry and the cultures of disciplines* 2nd ed., Buckingham: Open University Press in association with The Society for Research into Higher Education.

Becher, T., Trowler, P. & Becher, T.; Trowler, P., 2001. *Academic tribes and territories : intellectual enquiry and the cultures of disciplines* 2nd ed., Buckingham: Open University Press in association with The Society for Research into Higher Education. Available at: http://ww.mcgraw-hill.co.uk/openup/chapters/0335206271.pdf [Accessed February 16, 2014].

Berg, B.L., 2006. *Qualitative Research Methods for the Social Sciences*, Pearson.

Berners-Lee, T., Weitzner, D.J., et al., 2006. A Framework for Web Science. *Foundations and Trends® in Web Science*, 1(1), pp.1–130. Available at: http://www.nowpublishers.com/product.aspx?product=WEB&doi=1800000001 [Accessed June 13, 2011].

Berners-Lee, T., Hall, W., et al., 2006. Creating a Science of the Web. *Science*, 313(5788), pp.769–771. Available at: DOI: 10.1126/science.1126902.

Berners-Lee, T. & Fischetti, M., 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor* 1st ed., Harper San Francisco.

Bhattacharya, S. et al., 2012. Belief surveillance with Twitter. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci ’12*, pp.43–46. Available at: http://dl.acm.org/citation.cfm?doid=2380718.2380724.

Biggs, J., 2007. *Teaching for Quality Learning at University Third Edition Teaching for Quality Learning at University*, Open University Press. Available at: http://teaching.polyu.edu.hk/datafiles/R131.pdf.

Biglan, A., 1973. Biglan ' s Model of Differences among Academic Disciplines *. , p.1973.

Biglan, A., 1973a. Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, 57(3), pp.204–213.

Biglan, A., 1973b. The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57(3), pp.195–203. Available at: http://content.apa.org/journals/apl/57/3/195.

Bloom, B.S., 1956. *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*, New York, NY, USA: David McKay Co Inc.

Bloomberg, L. & Volpe, M., 2012. *Completing Your Qualitative Dissertation: A Road Map From Beginning to End*, Sage Publications.

Börner, K., Chen, C. & Boyack, K.W., 2003. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), pp.179–255. Available at: http://dx.doi.org/10.1002/aris.1440370106.

Branscomb, H.E., 2000. *Casting Your Net: A Student's Guide to Research on the Internet* 2nd ed., Needham Heights, MA, USA: Allyn & Bacon, Inc. Available at: http://dl.acm.org/citation.cfm?id=557665.

Braun, V. & Clarke, V., 2013. *Successful Qualitative Research: A Practical Guide for Beginners*, Sage Publications.

Bryant, A, & Charmaz, K., 2007. *The Sage handbook of grounded theory*, Sage Publications. Available at: https://books.google.co.uk/books?id=OrgZjp9CoN8C&printsec=frontcover#v=onepage&q&f=false.

Bryman, A., 2001. *Quantity and quality in social research*, Routledge.

Bryman, A., 2011. Triangulation. *Encyclopedia of Social Science Research Methods*, (1966), pp.1–5. Available at: http://www.sagepub.com/chambliss4e/study/chapter/encyc_pdfs/4.2_Triangulation.pdf.

Buter, R.K., Noyons, E.C.M. & Van Raan, A.F.J., 2004. A combination of quantitative and qualitative maps in an evaluative bibliometric context. *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, pp.978–982. Available at: http://ezproxy.rowan.edu:2157/search/freesrchabstract.jsp?tp=&arnumber=1320260&searchWithin=Search+Index+Terms:.QT.+bibliometric+maps.QT.&openedRefinements=*&searchField=Search+All&abstractAccess=no&userType=inst\nhttp://ieeexplore.ieee.org/lpdocs/epic03/.

Carr, L.A., 2010. Web Science Big Wins: Information Big Bang & Fundamental Constants. In *Satellite Meeting of the Royal Society Discussion Meeting "Web Science: A New Frontier" , Kavli Royal Society International Centre, UK30 - 29 Sep 2010.* London, UK.

Chin, K.L. et al., 2007. Ontology-based IT Pedagogical Knowledge Framework. *Informing Science*, pp.155–166.

Chomic, G., 1988. Change Over Time : Research Approaches to ICT - Enabled Emergent Change. , pp.28–31.

Cîmpan, S. & Salamatian, K., Internet Science Curriculum Construction in EINS.

Clark, D.D. et al., 2002. Tussle in Cyberspace: Defining Tomorrow's Internet. *SIGCOMM Comput. Commun. Rev.*, 32(4), pp.347–356. Available at: http://doi.acm.org/10.1145/964725.633059.

Clark, M., 2003. Computer Science: A hard-applied discipline? *Teaching in Higher Education*, 8(1), pp.71–87. Available at: http://www.tandfonline.com/action/journalInformation?journalCode=cthe20.

Clarke, V. & Braun, V., 2013. Teaching thematic analysis : Overcoming challenges and developing strategies for effective learning. *The Psychologist*, 26(2), pp.120–123.

Cohen, D.J. & Scheinfeldt, J.T., 2010. Hacking the Academy. A Book Crowdsourced in One Week (May 21-28, 2010). , (September).

Coskun, E., Smith, A. & White, S., 2015. Mapping the Web Science Curriculum (s). Available at: http://webscience-education-workshop.blogs.usj.edu.lb/files/2015/03/WSEDUCATION_2015_submission_3-elisabeth.pdf.

Davies, T. & Frank, M., 2013. There's no such thing as raw data: Exploring the socio-technical life of a government dataset. , (November 2012).

Day, M.J., Carr, L. & Halford, S., 2015. Developing the "Pro-human " Web. In *WebSci '15*. pp. 1–10.

Denning, P.J. et al., 1989. *COMPUTING AS A DISCIPLINE*,

Denning, P.J., Ralston, A. & Hemmendinger, D., 1999. *COMPUTER SCIENCE: THE DISCIPLINE*,

Denzin, N. & Lincoln, Y., 2011. *The SAGE Handbook of Qualitative Research*, Sage Publications. Available at: https://books.google.co.uk/books?id=Z-5yAwAAQBAJ&printsec=frontcover#v=onepage&q&f=false.

Dragan, I. & Zota, R., 2017. Collecting Facebook data for big data research. *2017 16th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, pp.1–3. Available at: http://ieeexplore.ieee.org/document/8123757/.

Easley, D. & Kleinberg, J., 2010. Chapter 13 The Structure of the Web. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, pp.375–395. Available at: http://www.cs.cornell.edu/home/kleinber/networks-book/\nChapter.

Engber, A., 2015. Data Collection by Facebook and Google. *Consumers' Research Winter 2015*, pp.14–15. Available at: http://consumersresearch.org/wp-content/uploads/2016/01/Data-Collection-by-Facebook-and-Google-Alec-Engber.pdf.

Feibleman, J.K., 1961. Pure Science , Science , Technology , Applied at Definitions An Attempt Engineering : *Technology and Culture*, 2(4), pp.305–317.

Flick, U., 2009. *An Introduction To Qualitative Research*, London: Sage Publications.

Glaser, B.G. & Strauss, A.L., 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Available at: http://www.amazon.com/dp/0202302601.

Guevara, K. & Blackwell, A.F., 2013. A Reflective Examination of a Process for Innovation and Collaboration in Internet Science. In *1st International Conference on Internet Science, 9-11 April*.

Halford, S., Pope, C. & Carr, L., 2010. A Manifesto for Web Science? In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*. Raleigh, NC: US.: Web Science Trust, pp. 1–6. Available at: http://eprints.ecs.soton.ac.uk/21033/\nhttp://journal.webscience.org/297/.

Hall, W. & Tiropanis, T., 2012. Web evolution and Web Science. *Computer Networks*, 56(18), pp.3859–3865. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1389128612003581 [Accessed November 12, 2013].

Hendler, J. et al., 2008. Web science: An Interdisciplinary Approach to Understanding the Web. *Communications of the ACM*, 51(7), pp.60–69. Available at:

http://dl.acm.org/citation.cfm?id=1364798 [Accessed May 15, 2015].

Heredia, R.M. & Vinueza, P.C., 2015. A Proposal Model to Monitor Interdisciplinary Research Projects in Latin American Universities. *Revista Iberoamericana de Tecnologias del Aprendizaje*, 10(3), pp.102–108.

HMG, 2012. *Open Data White Paper: Unleashing the Potential Unleashing the Potential Open Data White Paper*,

Hooper, C., 2014. The Role of Disciplinary Analysis in Web Science Education. , pp.0–1.

Hooper, C. & Dix, A., 2012. Web Science and Human-Computer Interaction: When Disciplines Collide. In *ACM Web Science 2012*.

Hooper, C.J., Dix, A., et al., 2013. Web Science and the Two (Hundred) Cultures: Representation of Disciplines Publishing in Web Science. In *Proceedings of the 5th Annual ACM Web Science Conference*. {WebSci} '12. New York, NY, USA: ACM, pp. 128–136. Available at: http://doi.acm.org/10.1145/2464464.2464478.

Hooper, C.J., Bordea, G. & Buitelaar, P., 2013. *Web Science and the two (hundred) cultures: representation of disciplines publishing in Web Science*, Available at: http://eprints.soton.ac.uk/354137/.

Hooper, C.J., Marie, N. & Kalampokis, E., 2012. Dissecting the butterfly: representation of disciplines publishing at the web science conference series. *WebSci*, pp.137–140. Available at: http://dblp.uni-trier.de/db/conf/websci/websci2012.html#HooperMK12 [Accessed February 16, 2014].

Hooper, C.J., Neves, B. & Bordea, G., 2015. A Disciplinary Analysis of Internet Science. *Traffic*, 1, pp.1–51.

Information Commissioner's Office, 2018. An Introduction to the Data Protection Bill. , p.77. Available at: https://ico.org.uk/media/for-organisations/documents/2258303/ico-introduction-to-the-data-protection-bill.pdf.

Kolb, D.A., 1984. Experiential Learning: Experience as The Source of Learning and Development. *Prentice Hall, Inc.*, (1984), pp.20–38.

KRATHWOHL, D. & ANDERSON, L., 2001. *A Taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives* L. Anderson & D. Krathwohl, eds., New York: Pearson.

Krishnan, A., 2009. *What are Academic Disciplines?*,

Lengler, R. & Eppler, M.J.M.M.J., 2007. Towards a periodic table of visualization methods for management. In *IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007)*. Clearwater, Florida, USA. Available at: http://www.visual-literacy.org/periodic_table/periodic_table.pdf [Accessed February 16, 2014].

Lin, A.C., 1998. Bridging positivist and interpretivist approaches to qualitative methods. *Policy Studies Journal*, 26(1), pp.162–180. Available at: Full Text HTML HTML: http://vnweb.hwwilsonweb.com.proxy.library.carleton.ca/hww/jumpstart.jhtml?recid=0bc05f7a67b1790eff29d56cacadaa60e6dbcf52760f81ad624a7e3db85177445beea3f845f9d0d0&fmt=H.

McGrath, W.E., 1978. Relationships between hard/soft, pure/applied, and life/nonlife disciplines and subject book use in a university library. *Information Processing and Management*, 14(1), pp.17–28.

Miarka, R., 2013. Representation of Knowledge patterns for Semantic Web. , 10(1), pp.201–207.

Mulholland, P., 2015. Storyscope : Supporting the authoring and reading of museum stories using online data sources.

Newman, J., 2012. *Difficult Questions About Videogames* Kindle., modernplaying.

Oinas-kukkonen, H., 2010. Behavior Change Support Systems : The Next Frontier for Web Science.

Oinas-Kukkonen, H., 2010. The Discipline of Information Systems : A Natural Strategic Alliance for Web Science MOST PROMINENT SUB-DISCIPLINE OF. In *Web Science 2010*.

Orlikowski, W.J., 2008. Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. *Organization Science*, 11(4), pp.404–428.

Phethean, C. et al., 2016. The Role of Data Science in Web Science.

QSR, 2010. THEMATIC ANALYSIS : Initial level Coding Using NVivo. , pp.1–21.

Rech, J., Feldmann, R.L. & Ras, E., 2009. Knowledge patterns. *Encyclopedia of Knowledge Management*.

Refactorings, K. et al., 2007. The Quality of Knowledge : Knowledge Patterns The Quality of Knowledge : Knowledge Patterns and Knowledge Refactorings. , (January).

Riera, D., 2008. Science, the Web and Web Science. *Voc Papers*. Available at: http://uocpapers.uoc.edu.

Schmidt, D., 2018. *Google Data Collection*, Available at: https://digitalcontentnext.org/wp-content/uploads/2018/08/DCN-Google-Data-Collection-Paper.pdf.

Shadbolt, N., What Is Web Science? (Talk). Available at: http://webscience.org/professor-nigel-shadbolt-explains-web-science/.

Shadbolt, N. & Berners-Lee, T., 2008. Web Science Emerges. *Scientific American*, 299(4), p.76. Available at: http://eprints.ecs.soton.ac.uk/17143/.

Shadbolt, N., Berners-Lee, T. & Shadbolt, N., 2008. Web Science Emerges. *Scientific American*, 299(October), p.76. Available at: http://eprints.ecs.soton.ac.uk/17143/.

Shin, J.C. & Arimoto, A., 2014. *Teaching and Research in Contemporary Higher Education: Systems, Activities and Rewards: 9 (The Changing Academy – The Changing Academic Profession* 9th ed., Springer. Available at: https://books.google.co.uk/books?id=K1jFBAAAQBAJ&pg=PA183&lpg=PA183&dq=life+science+vs+non+life+science&source=bl&ots=HUyWoF_Bsb&sig=2GEWxn87hux3QvU4UtG-l5N85XQ&hl=en&sa=X&ved=0ahUKEwjDl5Gu09rWAhWBF8AKHT6yD40Q6AEIggEwEQ#v=onepage&q=life science vs non lif.

Shneiderman, B., 2007. Web Science : A Provocative Invitation to Computer Science. *Communications of the ACM*, 50(6), pp.25–27.

Shuen, A., 2008. *Web 2.0: A Strategy Guide*, O'Reilly.

Strauss, A. & Corbin, J., 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Available at: http://srmo.sagepub.com/view/basics-of-qualitative-research/SAGE.xml.

Tang, A., A COMPUTING CURRICULUM WIKI : ANALYSIS AND MODELLING USING THE MAS-CommonKADS AGENT- ORIENTED METHODOLOGY. *Engineering*.

The Joint Task Force for Computing Curricula, 2005. *Computing curricula 2005: The Overview Report*, Available at: http://www.acm.org/education/education/curric_vols/CC200 5-March06Final.pdf.

The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society), 2013. *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*, Available at: http://ai.stanford.edu/users/sahami/CS2013/strawman-draft/cs2013-strawman.pdf\nhttp://www.acm.org/education/CS2013-final-report.pdf\nhttp://dl.acm.org/citation.cfm?id=2534860.

Tinati, R. et al., 2012. Mixing methods and theory to explore web activity. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, pp.308–316. Available at: http://dl.acm.org/citation.cfm?doid=2380718.2380758.

Tiropanis, T. et al., 2015. Network science, web science, and internet science. *Communications of the ACM*, 58(8), pp.76–82. Available at: http://dl.acm.org/citation.cfm?doid=2808213.2699416.

Tiropanis, T. & Hall, W., 2015. Network Science, Web Science and Internet Science: comparing interdisciplinary areas. *Communications of the ACM*, 58(8), pp.76–58. Available at: http://eprints.soton.ac.uk/376676/1/three-sciences-paper.pdf.

Topi, H. et al., 2010. IS 2010: Curriculum guidelines for undergraduate degree programs in information systems. *Communications of the Association for Information Systems*, 26(1), pp.359–428.

University of Leicester, 2015. *Student and Academic Services CODE OF PRACTICE FOR THE DEVELOPMENT , APPROVAL , AND Modification of taught provision*, Available at: https://www2.le.ac.uk/offices/sas2/quality/codes/documents/programmeapproval.pdf.

Vafopoulos, M., 2011. The Web Science Subject Categorization (WSSC). *ACM WebSci '11*.

Vafopoulos, M., 2010. Web Science Subject Categorization (WSSC). *ACM WebSci '11*.

Vedula, N. & Shalin, V.L., 2017. Predicting Trust Relations Within a Social Network : A Case Study on Emergency Response. , pp.53–62.

White, H.D. & McCain, K.W., 1997. Visualization of literatures. *Annual Review of Information Science and Technology*, Vol. 32, pp.99–168.

White, S. et al., 2011. *Negotiating the Web Science Curriculum Development through Shared Educational Artefacts*, Koblenz, Germany, NY, USA: ACM. Available at: http://eprints.ecs.soton.ac.uk/22141/.

White, S. & Vafopoulos, M.N., 2012. Web Science: expanding the notion of Computer Science. In *43rd ACM Technical Symposium on Computer Science Education, 28th February-3rd March 2012*. Rayleigh, NC. Available at: http://eprints.ecs.soton.ac.uk/22710/.

Wienberg, C. & Gordon, A.S., 2015. Insights on Privacy and Ethics from the Web ' s Most Prolific Storytellers.

Wimmer, M., Schauerhuber, A. & Kargl, H., 2007. On the Integration of Web Modeling Languages: Preliminary Results and Future Challenges. *Proceedings of the 3rd International Workshop on Model-Driven Web Engineering*. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.68.5508.

Wright, A., 2011. Web Science Meets Network Science. *Commun. ACM*, 54(5), p.23. Available at: http://doi.acm.org/10.1145/1941487.1941497.

Bibliography

Abel, F., Gao, Q., Houben, G., & Tao, K. (2011). Analyzing Temporal Dynamics in Twitter Profiles for Personalized Recommendations in the Social Web.

Abokhodair, N., Abbar, S., Vieweg, S., & Mejova, Y. (2016). Privacy and Twitter in Qatar : Traditional Values in the Digital World. *In Proceedings of the 8th ACM Conference on Web Science*, 66–77. doi:10.1145/2908131.2908146

Aiello, L. M., Schifanella, R., & State, B. (2014). Reading the source code of social ties. *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 139–148. doi:10.1145/2615569.2615672

Akker, C. Van Den, Legêne, S., Erp, M. Van, Segers, R., Wielinga, B., & Oomen, J. (2011). Digital Hermeneutics : Agora and the Online Understanding of Cultural Heritage Categories and Subject Descriptors.

Akkermans, H. (2012). Web Dynamics as a Random Walk : How and Why Power Laws Occur, 1–10.

Akkermans, H., & Bakhshi, R. (2013). Toward a next generation of network models for the web. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 1–10. doi:10.1145/2464464.2464517

Akkermans, H., Bon, A., Tuyp, W., & Allen, M. (2011). Is (Web) Science Ready for Empowerment ?

Alexandrov, D., Gorgadze, A., & Musabirov, I. (2016). Virtual caucasus on VK social networking site. *WebSci 2016 - Proceedings of the 2016 ACM Web Science Conference*, 215–217. doi:10.1145/2908131.2908205

Almaatouq, A., Alabdulkareem, A., & Nouh, M. (2013). Twitter : Who gets Caught ? Observed Trends in Social Micro-blogging Spam, 33–41.

AlNoamany, Y., Weigle, M. C., & Nelson, M. L. (2017). Generating stories from archived collections. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 309–318. doi:10.1145/3091478.3091508

Alonso, O., & Diaz, F. (2017). Automatic Generation of Event Timelines from Social Data, 207–211.

Alonso, O., Kandylas, V., Tremblay, S., Hofman, J. M., & Sen, S. (2017). What ' s Happening and What Happened : Searching the Social Web, 191–200. doi:10.1145/3091478.3091484

Alrajebah, N., Carr, L., Luczak-roesch, M., & Tiropanis, T. (2017). Deconstructing Diffusion on Tumblr : Structural and Temporal Aspects. *Proceedings of the 9th ACM Conference on Web Science*, 319–328. doi:10.1145/3091478.3091491

An, J., Quercia, D., Cha, M., Gummadi, K., & Crowcroft, J. (2013). Traditional media seen from social media. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 11–14. doi:10.1145/2464464.2464492

An, J., Quercia, D., & Crowcroft, J. (2013). Why individuals seek diverse opinions (or why they don't). *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 15–18. doi:10.1145/2464464.2464493

Anantharam, P., Thirunarayan, K., & Sheth, A. (2012). Topical anomaly detection from Twitter stream. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 11–14. doi:10.1145/2380718.2380720

Annexstein, F. S., & Berman, K. a. (2012). Synthesis ranking with critic resonance. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 15–23. doi:10.1145/2380718.2380721

Arantes, M., Figueiredo, F., & Almeida, J. M. (2016). Understanding video-ad consumption on YouTube: A measurement study on user behavior, popularity, and content properties. *WebSci 2016 - Proceedings of the 2016 ACM Web Science Conference*, 25–34. doi:10.1145/2908131.2908159

Araujo, M., Mejova, Y., Weber, I., & Benevenuto, F. (2017). Using Facebook Ads Audiences for Global Lifestyle Disease Surveillance. *Proceedings of the 2017 ACM on Web Science Conference - WebSci '17*, 253–257. doi:10.1145/3091478.3091513

Arbor, A., Fabrikant, A., & Lefevre, K. (2012). Look Who I Found : Understanding the Effects of Sharing Curated Friend Groups, 95–104.

Arbor, A., Teng, C., & Lin, Y. (n.d.). Recipe recommendation using ingredient networks.

Aroyo, L., & Welty, C. (n.d.). Crowd Truth : Harnessing disagreement in crowdsourcing a relation extraction gold standard.

Ashktorab, Z., Golbeck, J., Haber, E., & Vitak, J. (2017). Beyond cyberbullying: Self-disclosure, harm and social support on ASKfm. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 3–12. doi:10.1145/3091478.3091499

Ayala-Gomez, F., Daróczy, B., Mathioudakis, M., Benczúr, A., & Gionis, A. (2017). Where could we go? Recommendations for groups in location-based social networks. *[WebSci2017]Proceedings of the 2017 ACM on Web Science Conference*, 93–102. doi:10.1145/3091478.3091485

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012). Personality and patterns of Facebook usage. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 24–32. doi:10.1145/2380718.2380722

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012). Four degrees of separation. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 33–42. doi:10.1145/2380718.2380723

Baken, N. H. G., Wiegel, V., & Oortmerssen, G. Van. (n.d.). THE VALUE (DRIVEN) WEB, (d).

Bauer, J. M., & Eeten, M. Van. (2009). Ars Technica , 1–16.

Baumgartner, R., Kr, B., Max, C. G., Klein, D., & Gattringer, R. (2010). Web 2 . 0 Vision for the Blind.

Bazan, S. B., & Vafopoulos, M. (2013). The Web Science Curriculum at Work: The Digital Economy Master Program at USJ-Beirut. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 19–23). New York, NY, USA: ACM. doi:10.1145/2464464.2464477

Bazan, S. B., & Varin, C. (2011). Sharing innovative teaching experience in higher education on the Web . An interdisciplinary study on a contextualized Web 2 . 0 application for community building and teacher training.

Beasley, A., & Mason, W. (n.d.). Emotional States vs . Emotional Words in Social Media.

Berendsen, R., & Meij, E. (2011). Classifying Queries Submitted to a Vertical Search Engine.

Berners-lee, S. T., Hall, W., & Hendler, J. (2005). Hosted by the British Computer Society , London Report Executive summary.

Berners-lee, T., Hall, W., Hendler, J., & Shadbolt, N. (2005). Creating a Science of the Web. *Artificial Intelligence*, *2005*(September).

Berners-Lee, T., Weitzner, D. J., Hall, W., O'Hara, K., Shadbolt, N., & Hendler, J. a. (2006). A Framework for Web Science. *Foundations and Trends® in Web Science*, *1*(1), 1–130. doi:10.1561/1800000001

Bernstein, M. (2011). Flocks , Herds , and Stories temporal coherence and the long tail.

Bhattacharya, S., Suls, J., Tran, H., & Srinivasan, P. (2012). Belief surveillance with Twitter. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 43–46. doi:10.1145/2380718.2380724

Birkbak, A. (2013). From networked publics to issue publics : Reconsidering the public / private distinction in web science.

Bischoff, K. (2012). We Love Rock ' n ' Roll : Analyzing and Predicting Friendship Links in Last . fm, 47–56.

Boldi, P., & Monti, C. (2016). LlamaFur: Learning Latent Category Matrix to Find Unexpected Relations in Wikipedia, 218–222. doi:10.1145/2908131.2908153

Boldyrev, N., Spaniol, M., & Weikum, G. (2016). ACROSS : A Framework for Multi-Cultural Interlinking of Web Taxonomies. *Web Science 2016, Hannover, Germany*, 127–136. doi:10.1145/2908131.2908164

Bourlai, E., & Herring, S. C. (n.d.). Multimodal Communication on Tumblr : " I Have So Many Feels !," 171–175.

Brigadir, I., & Greene, D. (2015). Analyzing Discourse Communities with Distributional Semantic Models.

Bruns, A., & Weller, K. (2016). Twitter as a first draft of the present. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 183–189. doi:10.1145/2908131.2908174

Carr, L., & Harnad, S. (2009). Offloading Cognition onto the Web.

Carr, L., Pope, C., & Halford, S. (2010). Could the Web be a Temporary Glitch ? (pp. 1–6).

Casas, D. Las, Magno, G., Cunha, E., Gonçalves, M. A., Cambraia, C., & Almeida, V. (n.d.). Noticing the Other Gender on Google +, 156–160.

Ceolin, D., & Hage, W. R. Van. (n.d.). A Trust Model to Estimate the Quality of Annotations using the Web.

Chan, J., & Hayes, C. (n.d.). Decomposing Discussion Forums using User Roles.

Chandra, P., & Kalyanasundaram, A. (2012). A network pruning based approach for subset-specific influential detection. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 57–66. doi:10.1145/2380718.2380726

Chandrasekar, R., & Jain, K. (2010). Peer-to- Peer Human Computation & " Help Me Decide ":

Chatzakou, D., Kourtellis, N., Blackburn, J., De Cristofaro, E., Stringhini, G., & Vakali, A. (2017). Mean Birds: Detecting Aggression and Bullying on Twitter, 13–22. doi:10.1145/3091478.3091487

Cheng, L., Zhang, L., & Wang, J. (2012). A study of human flesh search with epidemic models. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 67–73. doi:10.1145/2380718.2380727

Cheng, X., Macaulay, L., & Zarifis, A. (2008). Individual trust development in computer mediated teamwork. *WebScience '08: Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence*, 1–5. doi:10.1145/1379157.1379159

Cîmpan, S., & Salamatian, K. (n.d.). Internet Science Curriculum Construction in EINS.

Cojan, J., Cabrio, E., Gandon, F., Lucioles, R., & Antipolis, S. (2013). Filling the Gaps Among DBpedia Multilingual Chapters for Question Answering.

Collins, T., Keynes, M., & Mulholland, P. (2012). Web supported emplotment : Using object and event descriptions to facilitate storytelling online and in galleries, 74–77.

Coskun, E. (2017). Web Science : Mapping the Curriculum, 241–245.

Coskun, E. (2014). Web Science Presentation : Student Experience.

Coskun, E., Smith, A., & White, S. (2015). Mapping the Web Science Curriculum (s). Retrieved from http://webscience-education-workshop.blogs.usj.edu.lb/files/2015/03/WSEDUCATION_2015_submission_3-elisabeth.pdf

Creese, S. (2008). No Title, 1–5.

Croitoru, M., Staab, S., Cerri, S., Jonquet, C., Scharffe, F., White, S., … Prini, G. (2011). Web Science Curriculum Development, (1), 1–2.

Cruz-albrecht, L. (2017). Characterizing Regional and Behavioral Device Variations Across the Twitter Timeline : a Longitudinal Study, 279–288.

Curie, M. (2013). Towards A Redefinition of Time in Information Networks ?, 1–4.

Daly, E. M., Dahlem, D., & Quercia, D. (2014). The New Blocs on the Block : Using Community Forums to Foster New Neighbourhoods, 52–61.

Dani, R. (2017). ASSIST : Automatic Summarization of SIgnificant Structural Changes in Large Temporal Graphs, 201–205. doi:10.1145/3091478.3091518

d'Aquin, M., Adamou, A., & Dietze, S. (2013). Assessing the educational linked data landscape. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 43–46. doi:10.1145/2464464.2464487

Das, A., Gollapudi, S., Kıcıman, E., & Varol, O. (2016). Information dissemination in heterogeneous-intent networks. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 259–268. doi:10.1145/2908131.2908161

Davies, H. C., Halford, S. J., & Gibbins, N. (2012). " Digital Natives ? Investigating young people " s critical skills in evaluating web based information ', 78–81.

Davies, T., & Frank, M. (2013). There's no such thing as raw data: Exploring the socio-technical life of a government dataset, (November 2012).

Day, M. J., Carr, L., & Halford, S. (2015). Developing the "Pro-human " Web. In *WebSci '15* (pp. 1–10). doi:10.1145/2786451.2786458

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Social media as a measurement tool of depression in populations. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 47–56. doi:10.1145/2464464.2464480

Degirmencioglu, E. A. (2010). Exploring Area-Specific Microblogging Social Networks.

Denardis, L. (2014). The Global War for Internet Governance, 2618146.

Diaz-Aviles, E., & Stewart, A. (2012). Tracking Twitter for epidemic intelligence. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 82–85. doi:10.1145/2380718.2380730

Difranzo, D. (2011). A Farm in Every Window : A Study into the Incentives for Participation in the Windowfarm Virtual Community.

Difranzo, D., Fernando, M., & Hendler, J. (2013). The Performativity of Data : Re-conceptualizing the Web of Data.

Dipple, A., Raymond, P. K., Prof, A., & Docherty, M. (2012). Stigmergy in Web 2 . 0 : a Model for Site Dynamics, 86–94.

Dix, A. (2011). A Shifting Boundary : the dynamics of internal cognition and the web as external representation.

Doosti, B., Crandall, D. J., & Su, N. M. (2017). A deep study into the history of web design. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 329–338. doi:10.1145/3091478.3091503

Dumitrache, A., Groth, P., & van den Besselaar, P. (2013). Identifying research talent using web-centric databases. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 57–60. doi:10.1145/2464464.2464507

Eikelboom, F., Groth, P., de Boer, V., & Hollink, L. (2013). A comparison between online and offline prayer. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, (2), 61–64. doi:10.1145/2464464.2464505

Erickson, J. S., Joan, M., Gloria, K. T., & Hughes, G. N. (2010). Expanding Web Science Integration in an ITWS Curriculum Categories and Subject Descriptors.

Feigenbaum, J., Haven, N., Hendler, J. A., Cs, R. P. I., Depts, C. S., Jaggard, A. D., … Wright, R. N. (2011). Accountability and Deterrence in Online Life (Extended Abstract).

Fernandez, M., Piccolo, L. S. G., Maynard, D., Wippoo, M., Meili, C., & Alani, H. (2016). Talking climate change via social media: communication, engagement and behaviour. *Proceedings of the 8th ACM Conference on Web Science*, 85–94. doi:10.1145/2908131.2908167

Fetahu, B., Anand, A., & Anand, A. (2007). How much is Wikipedia Lagging Behind News ?

Fiorio, L., Zagheni, E., Abel, G., Weber, I., Cai, J., & Vinué, G. (2017). Using twitter data to estimate the relationships between short-term mobility and long-term migration. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 103–110. doi:10.1145/3091478.3091496

Fitzsimmons, G., Weal, M. J., & Drieghe, D. (2014). Skim Reading : An Adaptive Strategy for Reading on the Web, 211–219.

Fitzsimmons, G., Weal, M., & Drieghe, D. (2013). On measuring the impact of hyperlinks on reading. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 65–74. doi:10.1145/2464464.2464501

Flöck, F., Vrandeˇ, D., & Simperl, E. (2011). Towards a diversity-minded Wikipedia.

Forever, B. F. (2009). BFFE (Be Friends Forever): the way in which young adolescents are using social networking sites to maintain friendship and exp, (February), 18–22.

Frank, M., & Allison, C. (n.d.). The Role of the Web Science MSc in Preparing For Research.

Fredheim, R., Moore, A., & Naughton, J. (n.d.). Anonymity and Online Commenting : The Broken Windows Effect and the End of Drive-by Commenting.

Fryer, H., & Simperl, E. (2017). Web Science Challenges in Researching Bug Bounties. *Proceedings of the 2017 ACM on Web Science Conference - WebSci '17*, 273–277. doi:10.1145/3091478.3091517

Fyson, R. W., Coles, S., & Carr, L. (2013). AltOA : A Framework for Dissemination Through Disintermediation.

Gadiraju, U., Dietze, S., & Diaz-aviles, E. (n.d.). Ranking Buildings and Mining the Web for Popular Architectural Patterns.

Gaffney, D. (2010). # iranElection : Quantifying Online Activism.

Gamble, M., & Goble, C. (n.d.). Quality , Trust , and Utility of Scientific Data on the Web : Towards a Joint Model Categories and Subject Descriptors.

Gamble, M., & Goble, C. (n.d.). Standing on the Shoulders of the Trusted Web : Trust , Scholarship and Linked Data .

Gao, M. (2012). EventShop : From Heterogeneous Web Streams to Personalized Situation Detection and Control, 105–108.

Garimella, K., Kostakis, O., & Mathioudakis, M. (2017). Ad-blocking: A Study on Performance, Privacy and Counter-measures, 259–262. doi:10.1145/3091478.3091514

Garimella, K., & Mathioudakis, M. (2017). The Effect of Collective Attention, 43–52.

Garimella, K., Morales, G. D. F., Gionis, A., & Mathioudakis, M. (2017). Factors in Recommending Contrarian Content on Social Media. *[WebSci2017]Proceedings of the 2017 ACM on Web Science Conference*, 10–13. doi:10.1145/3091478.3091515

Garimella, K., Weber, I., & De Choudhury, M. (2016). Quote RTs on Twitter: Usage of the New Feature for Political Discourse, 200–204. doi:10.1145/2908131.2908170

Gaskell, P., McGroarty, F., & Tiropanis, T. (2013). An investigation into correlations between financial sentiment and prices in financial markets. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 99–108. doi:10.1145/2464464.2464510

Gkotsis, G., Stepanyan, K., Pedrinaci, C., Domingue, J., & Keynes, M. (2014). It ' s all in the Content : State of the art Best Answer Prediction based on Discretisation of Shallow Linguistic Features, 202–210.

Golbeck, J., Gnanasekaran, R. K., Gunasekaran, R. R., Hoffman, K. M., Hottle, J., Jienjitlert, V., … Gergory, Q. (2017). A Large Labeled Corpus for Online Harassment Research. *Proceedings of the 2017 ACM on Web Science Conference - WebSci '17*, 229–233. doi:10.1145/3091478.3091509

Gossen, G., Demidova, E., & Risse, T. (2016). Analyzing web archives through topic and event focused sub-collections. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 291–295. doi:10.1145/2908131.2908175

Gottron, T., & Alhadi, A. C. (2011). Bad News Travel Fast : A Content-based Analysis of Interestingness on Twitter.

Greene, D., & Cross, J. P. (2015). Unveiling the Political Agenda of the European Parliament Plenary : A Topical Analysis.

Greene, D., & Cunningham, P. (2013). Producing a unified graph representation from multiple social network views. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 118–121. doi:10.1145/2464464.2464471

Groth, P., & Gurney, T. (2010). Studying Scientific Discourse on the Web Using Bibliometrics : A Chemistry Blogging Case Study.

Guanliang Chen, Dan Davis, Jun Lin, C. H. and G.-J. H. (2016). Beyond the MOOC platform: Gaining Insights aboutLearners from the Social Web. *ACM Web Science Conference, WebSci '16*, 15–24. doi:10.1145/2908131.2908145

Guimarães, R. L., Cesar, P., Bulterman, D. C. A., Kegel, I., Ljungstrand, P., & Park, A. (2011). Social Practices around Personal Videos using the Web.

Guntuku, S. C., Lin, W., Carpenter, J., Ng, W. K., Ungar, L. H., & Preoţiuc-Pietro, D. (2017). Studying Personality through the Content of Posted and Liked Images on Twiier. *Proceedings of the 9th International ACM Web Science Conference*, 223–227. doi:10.1145/3091478.3091522

Gupte, M., & Eliassi-Rad, T. (2012). Measuring tie strength in implicit social networks. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 109–118. doi:10.1145/2380718.2380734

Gyan, N. B., Bon, A., Aart, C. Van, Froumentin, M., Allen, M., Eco, S., & Rue, A. C. I. (2013). Voice-based Web access in rural Africa.

Hadgu, A. T. (n.d.). Identifying and Analyzing Researchers on Twitter, 23–32.

Hale, S. A. (2011). Multilinguals and Wikipedia Editing Categories and Subject Descriptors, 99–108.

Hale, S. a., Margetts, H., & Yasseri, T. (2013). Petition growth and success rates on the UK No. 10 Downing Street website. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, (10), 132–138. doi:10.1145/2464464.2464518

Hale, S. A., Meyer, E. T., Schroeder, R., & Margetts, H. (n.d.). Mapping the UK Webspace : Fifteen Years of British Universities on the Web Categories and Subject Descriptors Data preparation, 62–70.

Halford, S. (n.d.). Reflections on Developing a Cross Faculty Web Science Undergraduate Programme.

Halford, S., Pope, C. C., & Carr, L. L. (2006). A manifesto for Web Science? In *WebSci10: Extending the Frontiers of Society On-Line* (pp. 1–6). Raleigh, NC: US.: Web Science Trust. Retrieved from http://eprints.soton.ac.uk/271033/

Hall, W. (2016). Web Science, Social Media and Education. In *WWW2016*. Montreal. Retrieved from http://www2016.ca/

Hall, W., & Freng, D. B. E. F. R. S. (2014). Observing the Web, 2618143.

Hall, W., Hendler, J., & Staab, S. (2016). A Manifesto for Web Science @ 10, (December), 1–4.

Hall, W., & Tiropanis, T. (2012). Web evolution and Web Science. *Computer Networks*, *56*(18), 3859–3865. doi:10.1016/j.comnet.2012.10.004

Halpern, D., & Katz, J. E. (2012). From e-government to Social Network Government : Towards a Transition Model, 119–127.

Halpin, H. (n.d.). Social Meaning on the Web : From Wittgenstein To Search Engines, 1–2.

Halpin, H. (2013). Does the web extend the mind? *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 139–147. doi:10.1145/2464464.2479972

Halpin, H., Clark, A., & Wheeler, M. (2008). Towards a Philosophy of the Web : Representation , Enaction , Collective Intelligence, 1–5.

Han, K., Lee, S., Jang, J. Y., Jung, Y., & Lee, D. (2016). " Teens are from Mars , Adults are f rom Venus " : Analyzing and Predicting Age Groups with Behavioral Characteristics in Instagram. *ACM Web Science 2016*, 35–44. doi:10.1145/2908131.2908160

Hara, K. O. (2012). Transparency , Open Data and Trust in Government : Shaping the Infosphere, 223–232.

Hara, K. O. (2010). Intimacy 2 . 0 : Privacy Rights and Privacy Responsibilities on the World Wide Web.

Hara, K. O., Stevens, D., & Smith, A. (n.d.). The Devil ' s Long Tail : Religious Moderation and Extremism on the Web.

Hardman, L., Ossenbruggen, J. Van, Amin, A., & Hildebrand, M. (n.d.). Interactive Information Access on the Web of Data.

Hare, J., Osmond, A., Yang, Y., Wills, G., Weal, M., Roure, D. De, … Yardley, L. (n.d.). LifeGuide : A platform for performing web - based behavioural interventions 2 . The LifeGuide Software Suite.

Hargood, C., Millard, D., & Weal, M. (2008). A Thematic Approach to Emerging Narrative Structure. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence 08*, 41–45. doi:10.1145/1379157.1379168

Haslhofer, B., Robitza, W., Guimbretiere, F., & Lagoze, C. (2013). Semantic tagging on historical maps. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 148–157. doi:10.1145/2464464.2464506

He, X., Lu, D., Margolin, D., Wang, M., Idrissi, S. El, & Lin, Y. (2017). The Signals and Noise : Actionable Information in Improvised Social Media Channels During a Disaster, 33–42.

Hendler, J. A. (2010). Study Supreme Court Decision Making with Linked Data.

Hendler, J., Shadbolt, N., Hall, W., Berners-lee, T., & Weitzner, D. (2008). Web science: An Interdisciplinary Approach to Understanding the Web. *Communications of the ACM*, *51*(7), 60–69. doi:10.1145/1364782.1364798

Herdagdelen, A., Adamic, L., State, B., & Mason, W. (2016). The social ties of immigrant communities in the United States. *Facebook Research*, 78–84. doi:10.1145/2908131.2908163

Hoang, D. T., & Menczer, F. (n.d.). Crowdsourcing Scholarly Data.

Hockx-yu, H. (2011). The Past Issue of the Web, (December 2010).

Holzmann, H., Nejdl, W., & Anand, A. (2017). Exploring Web Archives Through Temporal Anchor Texts ∗, 289–298.

Hong, L., Torrens, P., Fu, C., & Frias-Martinez, V. (2017). Understanding citizens' and local governments' digital communications during natural disasters: The case of snowstorms. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 141–150. doi:10.1145/3091478.3091502

Hooper, C. (2014). The Role of Disciplinary Analysis in Web Science Education, 0–1.

Hooper, C. J. (2010). Using TAPT as an Analytical Method for Understanding Online Experiences.

Hooper, C. J., Bailey, B., Glaser, H., & Hendler, J. (2010). Social Machines in Practice : Solutions , Stakeholders and Scopes, 156–160. doi:10.1145/2908131.2908176

Hooper, C. J., Bordea, G., & Buitelaar, P. (2013). *Web Science and the two (hundred) cultures: representation of disciplines publishing in Web Science*. *WebSci'13*. Retrieved from http://eprints.soton.ac.uk/354137/

Hooper, C. J., Dix, A., Bordea, G., & Buitelaar, P. (2013). Web Science and the Two (Hundred) Cultures: Representation of Disciplines Publishing in Web Science. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 128–136). New York, NY, USA: ACM. doi:10.1145/2464464.2464478

Hooper, C. J., Marie, N., & Kalampokis, E. (2012). Dissecting the butterfly: representation of disciplines publishing at the web science conference series. *WebSci*, 137–140. doi:10.1145/2380718.2380737

Hooper, C. J., & Millard, D. E. (2010). Teasing Apart and Piecing Together : Towards Understanding Web-based Interactions.

Hooper, C. J., Nind, M., Parsons, S., Power, A., & Collis, A. (2015). Building a Social Machine : Co-designing a TimeBank for Inclusive Research.

Hooper, C., & Dix, A. (2012). Web Science and Human-Computer Interaction: When Disciplines Collide. In *ACM Web Science 2012*.

Huang, S., Peng, W., Li, J., & Lee, D. (2013). Sentiment and Topic Analysis on Social Media : A Multi-Task Multi-Label Classification Approach.

Hughes, C., & Lawrence, W. (n.d.). International interdisciplinary collaboration : Suggestions for improving project management.

Huurdeman, H. C., Ben-David, A., & Sammar, T. (2013). Sprint methods for web archive research. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 182–190. doi:10.1145/2464464.2464513

Introne, J., Lansing, E., & Goggins, S. (2015). Taming a Menagerie of Heavy Tails with Skew Path Analysis, 1–5.

Iren, D., Liem, C. C. S., Yang, J., & Bozzon, A. (2016). Using social media to reveal social and collective perspectives on music, 296–300. doi:10.1145/2908131.2908178

Isaksen, L., Simon, R., Barker, E. T. E., Hall, W., & Keynes, M. (2014). Pelagios and the Emerging Graph of Ancient World Data, 197–201.

Ishida, T. (n.d.). The Language Grid for Intercultural Collaboration.

Jacobs, A. Z., Way, S. F., Ugander, J., & Clauset, A. (n.d.). Assembling thefacebook : Using Heterogeneity to Understand Online Social Network Assembly Categories and Subject Descriptors.

Jesic, D. (2011). Web Technologies for Open Innovation.

Jewell, M. O. (2013). Beyond The Trowel ' s Edge : Provenance-Based Collective Archaeological Interpretations.

Jürgens, P., Jungherr, A., & Schoen, H. (2011). Small Worlds with a Difference : New Gatekeepers and the Filtering of Political Information on Twitter GATEKEEPERS DURING THE.

Kalo, J., Homoceanu, S., Rose, J., & Balke, W. (2015). Avoiding Chinese Whispers : Controlling End-to-End Join Quality in Linked Open Data Stores.

Kappler, K. E. (2011). Is there anybody out there ? – Social Media as a new social fetish.

Karnstedt, M., Rowe, M., Chan, J., Alani, H., & Hayes, C. (2011). The Effect of User Features on Churn in Social Networks.

Katsh, E. (2006). Online Dispute Resolution : Designing New Legal Processes for Cyberspace, *1367*(1996).

Kawase, R., Herderlsde, E., & Nejdl, W. (n.d.). Who Wants To Get Fired ?

Ke, Q. (2017). Sharing Means Renting?: An Entire-marketplace Analysis of Airbnb, 131–139. doi:10.1145/3091478.3091504

Keegan, B., Ahmed, M. A., & Williams, D. (2011). Sic Transit Gloria Mundi Virtuali ? Promise and Peril in the Computational Social Science of Clandestine Organizing.

Keneshloo, Y., Tech, V., Cadena, J., & Bioinformatics, V. (n.d.). Detecting and Forecasting Domestic Political Crises : A Graph-based Approach, 192–196.

Kershaw, D., & Rowe, M. (n.d.). Towards Tracking and Analysing Regional Alcohol Consumption Patterns in the UK through the use of Social Media Categories and Subject Descriptors, 220–228.

Kinder-kurlanda, K., & Weller, K. (2014). " I Always Feel It Must Be Great to Be a Hacker !" The Role of Interdisciplinary Work in Social Media Research, 91–98.

Kleek, M. Van, Murray-rust, D., Guy, A., Smith, D. A., & Shadbolt, N. R. (n.d.). Self Curation , Social Partitioning , Escaping from Prejudice and Harassment : the Many Dimensions of Lying Online.

Knobloch-westerwick, S., & Westerwick, A. (2011). Social Comparisons at Your Fingertips : The Importance of Ingroup / Outgroup Status.

Knud, M., Hausenblas, M., Cyganiak, R., Handschuh, S., Park, I. D. A. B., & Dangan, L. (2010). Learning from Linked Open Data Usage : Patterns & Metrics, 1–8.

Koczanski, A., Sabou, M., & Hunter, S. (n.d.). Sustainability Implications of Open Government Data : A Cross-Regional Study.

Kohli, P., Stillwell, D., Kearns, M., & Herbrich, R. (2012). Colonel Blotto On Facebook : The Effect of Social Relations On Strategic Interaction, 141–150.

Koltcov, S. (n.d.). Latent Dirichlet Allocation : Stability and Applications to Studies of User-Generated Content β α θ z, 161–165.

Koltsova, O., Koltcov, S., & Alexeeva, S. (2014). Do ordinary bloggers really differ from blog celebrities? *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 166–170. doi:10.1145/2615569.2615675

Kontostathis, A., Reynolds, K., Garron, A., & Edwards, L. (2013). Detecting Cyberbullying : Query Terms and Techniques.

Kosinski, M. (2012). Crowd IQ : Measuring the Intelligence of Crowdsourcing Platforms, 151–160.

Koumenides, C. L., Salvadores, M., Alani, H., & Shadbolt, N. R. (2010). Global Integration of Public Sector Information.

Krafft, D. B., Cappadona, N. A., Devare, B. M., Lowe, B. J., & Corson-rikert, J. (2010). VIVO : Enabling National Networking of Scientists.

Kraker, P. (2013). Head Start : Improving Academic Literature Search with Overview Visualizations based on Readership Statistics.

Krishnan, S., Butler, P., Tandon, R., Leskovec, J., & Ramakrishnan, N. (2016). Seeing the Forest for the Trees: New Approaches to Forecasting Cascades. *Proceedings of the 8th ACM Conference on Web Science*, 249–258. doi:10.1145/2908131.2908155

Kuhlman, C. J., Bioinformatics, V., Tech, V., & Ravi, S. S. (2012). Controlling Opinion Bias in Online Social Networks, 165–174.

Labs, N. B., Nicholson, P. K., Labs, N. B., & Sala, A. (2017). Breaking Cycles in Noisy Hierarchies ∗, 151–160.

Lagoze, C. (n.d.). The oreChem Project : Integrating Chemistry Scholarship with the Semantic Web, 1–4.

Laufer, P., Wagner, C., Koblenz, U., Flöck, F., Strohmaier, M., & Koblenz, U. (2015). Mining cross-cultural relations from Wikipedia - A study of 31 European food cultures.

Lazuras, L. (2012). Cyberbullying 2 . 0 : Implications of Web development for the prevention of cyberbullying.

Le, A. (n.d.). Country-Level Spatial Dynamics of User Activity : A Case Study in Location-Based Social Networks, 71–80.

Lee, M., Rha, J., & Kwon, T. T. (n.d.). Online Footsteps to Purchase : Exploring Consumer Behaviors on Online Shopping Sites Categories and Subject Descriptors.

Lehmberg, O., Meusel, R., & Bizer, C. (n.d.). Graph Structure in the Web – Aggregated by Pay-Level Domain, (3), 119–128.

Letierce, J., Passant, A., Decker, S., & Breslin, J. G. (2009). Understanding how Twitter is used to spread scientific messages.

Liao, H., & Hale, S. A. (n.d.). How much is said in a microblog ? A multilingual inquiry based on Weibo and Twitter.

Liao, Y., Lee, D., & Lee, K. (2017). Understanding Temporal Backing Patterns in Online Crowdfunding Communities, (May), 369–378. doi:10.1145/3091478.3091480

Lieto, A., & Vernero, F. (2013). Unveiling the link between logical fallacies and web persuasion. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 473–478. doi:10.1145/2464464.2508564

Lim, E., & Lo, D. (2013). R -energy for Evaluating Robustness of Dynamic Networks, (1).

Lim, E., Xie, W., Li, C., Zhu, F., & Gong, X. (2012). When a friend in Twitter is a friend in life. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 344–347. doi:10.1145/2380718.2380763

Limpens, F., Antipolis, F.-S., & Buffa, M. (2010). Helping online communities to semantically enrich folksonomies.

Lorince, J., & Todd, P. M. (2013). Can simple social copying heuristics explain tag popularity in a collaborative tagging system? *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 215–224. doi:10.1145/2464464.2464516

Lorince, J., Zorowitz, S., Murdock, J., & Todd, P. M. (2014). "Supertagger" behavior in building folksonomies. *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 129–138. doi:10.1145/2615569.2615686

Loyola, P., & Szederkenyi, F. (n.d.). Using the Web to Support Political Analysis : Identifying Legislative Bill Ideology in the Chilean Parliament, 190–199. doi:10.1145/2908131.2908166

Lu, D. (2015). Time to Introduce Myself ! Impact of Self-disclosure Timing of Newcomers in Online Discussion Forums.

Lu, X. (2012). The Configuration of Networked Publics on the Web : Evidence from the Greek Indignados Movement, 185–194.

Lu, Y., Zhang, L., Xiao, Y., & Li, Y. (2013). Simultaneously Detecting Fake Reviews and Review Spammers using Factor Graph Model, (ii).

Luczak-rösch, M. (2013). Beyond Positivism in Computer Science.

Lutz, C., & Tamò, A. (2015). RoboCode-Ethicists – Privacy-friendly robots , an ethical responsibility of engineers ?

Magdon-ismail, M. (2012). Guard Your Connections : Infiltration of a Trust / Reputation Based Network, 195–204.

Magdy, W., Darwish, K., Abokhodair, N., Rahimi, A., & Baldwin, T. (2016). #ISISisNotIslam or #DeportAllMuslims? *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 95–106. doi:10.1145/2908131.2908150

Mahata, D., Rock, L., Talburt, J. R., & Rock, L. (2011). From Chirps to Whistles Discovering Event-specific Informative Content from Twitter.

Margetts, H., John, P., Escher, T., & Reissfelder, S. (2009). Experiments for Web Science : Examining the Effect of the Internet on Collective Action.

Marshall, C. C., & Shipman, F. M. (2013). Experiences Surveying the Crowd : Reflections on Methods , Participation , and Reliability.

Martin, C. J. (n.d.). Socio-Technical Transitions Pathways for UK Open Government Data.

Matias, J. N. (n.d.). Measuring Social Media Quotations in Journalism.

Mazi, A. (2013). Toward Google Borders.

Mccabe, C., Watson, R. A., Prichard, J., & Hall, W. (2011). The Web as an Adaptive Network : Coevolution of Web Behavior and Web Structure.

Mccown, F., & Nelson, M. L. (2013). Resources for Teaching Web Science to Computer Science Undergraduates, 5–6.

Mccusker, J. P., Chastain, K., Stingone, J. A., & Mcguinness, D. L. (2017). Broad , Interdisciplinary Science In Tela : An Exposure and Child Health Ontology, 349–357.

Mejova, Y., & Srinivasan, P. (2012). Political Speech in Social Media Streams : YouTube Comments and Twitter Posts, 205–208.

Metaxas, P. T., & Kerry, S. J. (2010). From Obscurity to Prominence in Minutes : Political Speech and Real-Time Search.

Metaxas, P., & Mustafaraj, E. (2013). The rise and the fall of a citizen reporter. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 248–257. doi:10.1145/2464464.2464520

Mihail, M., Amanatidis, Y., & Young, S. (n.d.). Kernel Models for Complex Networks Random Graph Models based on Kernel Functions, 1–5.

Millard, D. E., Webber, C., & Hara, K. O. (2013). Unpicking the Privacy Paradox : Can Structuration Theory H elp to E xplain L ocation - B ased P rivacy D ecisions ?

Minder, P., & Bernstein, A. (2012). How to Translate a Book Within an Hour Towards General Purpose Programmable Human Computers with CrowdLang, 209–212.

Minguillón, J. (2002). Web and education , a successful open entanglement.

Mittelmeier, J., Hall, W., Rienties, B., Hall, W., Whitelock, D., & Hall, W. (2016). Using Social Network Analysis to predict online contributions : The impact of network diversity in cross- cultural collaboration. *ACM Web Science*, 269–273. doi:http://dx.doi.org/10.1145/2908131.2908169

Mitter, S. (n.d.). Understanding The Impact Of Socialbot Attacks In Online Social Networks.

Mitter, S., & Wagner, C. (n.d.). A Categorization Scheme for Socialbot Attacks In Online Social Networks.

Momeni, E., & Sageder, G. (2013). An empirical analysis of characteristics of useful comments in social media. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 258–261. doi:10.1145/2464464.2464490

Moore, R. (2011). Social Media on the Job : An exploration of the potential legal consequences of employees ' social media activities during the course of.

Moser, C. (n.d.). Preferential Attachment in Online Networks : Measurement and Explanations.

Mueller, J., & Stumme, G. (2017). Predicting Rising Follower Counts on Twitter Using Profile Information, *1017*, 121–130. doi:10.1145/3091478.3091490

Mulholland, P. (2015). Storyscope : Supporting the authoring and reading of museum stories using online data sources.

Murugadas, D. (2016). Do It Yourself Diagnosis A Study on Acquiring Health-Related Information Online, 276–285. doi:10.1145/2908131.2908147

Mustafaraj, E., & Metaxas, P. T. (2017). The Fake News Spreading Plague: Was it Preventable?, (June), 235–239. doi:10.1145/3091478.3091523

Nash, R. (2015). Considering a Wider Web ? Employing Multimodal Critical Discourse Analysis in Exploration of Multiple Online Spaces.

Naveed, N., & Staab, S. (2011). ATT : Analyzing Temporal Dynamics of Topics and Authors in Social Media.

Nguyen, N. P., Yan, G., Thai, M. T., Eidenbenz, S., & Science, I. (2012). Containment of misinformation spread in online social networks. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 213–222. doi:10.1145/2380718.2380746

Nie, D., Li, A., Hao, B., & Zhu, T. (2007). Personality Traits and Microblogging Behavior of Weibo Users : Onlies versus Others, 2–7.

Noy, N. F., Mortensen, J., Musen, M. A., & Alexander, P. R. (2013). Mechanical Turk as an Ontology Engineer ? Using Microtasks as a Component of an Ontology-Engineering Workflow.

Nuland, A. Van, Legˆ, S., & Aroyo, L. (2013). From Information Delivery to Interpretation Support : Evaluating Cultural Heritage Access on the Web.

Nurmikko, T. (1889). Web Science for Ancient History : Deciphering Proto-Elamite Online.

Nuzzolese, A. G., Presutti, V., Istc-cnr, S., Gangemi, A., Ciancarini, P., & Musetti, A. (2013). Aemoo : exploring knowledge on the Web, 4–7.

O'Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2013). Uncovering the wider structure of extreme right communities spanning popular online networks. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 276–285. doi:10.1145/2464464.2464495

Ogden, J., Halford, S., & Carr, L. (2017). Observing Web Archives, 299–308.

O'Hara, K., & Hall, W. (2012). Web Science. *Scientific American*, *51*(7), 60. Retrieved from http://eprints.ecs.soton.ac.uk/23097/

Oinas-kukkonen, H. (2010). Behavior Change Support Systems : The Next Frontier for Web Science.

Oinas-Kukkonen, H. (2010). The Discipline of Information Systems : A Natural Strategic Alliance for Web Science MOST PROMINENT SUB-DISCIPLINE OF. In *Web Science 2010*.

Orellana-rodriguez, C., Greene, D., & Keane, M. T. (2016). Spreading the News : How Can Journalists Gain More Engagement for their Tweets ? *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 107–116. doi:10.1145/2908131.2908154

Origgi, G., & Nicod, C. I. (2013). Democracy and Trust in the Age of the Social Web.

Pareti, P., Klein, E., & Barker, A. (2002). A Linked Data Scalability Challenge : Concept Reuse Leads to Semantic Decay, 3–7.

Park, J. Y. (2012). When Daily Deal Services Meet Twitter : Understanding Twitter as a Daily Deal Marketing Platform, 233–242.

Park, S., Kang, S., Lee, S., Chung, S., & Song, J. (2008). Mitigating media bias. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence - WebScience '08*, 47. doi:10.1145/1379157.1379169

Park, S., Ko, M., Lee, J., Choi, A., & Song, J. (2013). Challenges and Opportunities of Local Journalism : A Case Study of the 2012 Korean General Election.

Passant, A., Samwald, M., Breslin, J. G., & Decker, S. (n.d.). Federating Distributed Social Data to Build an Interlinked Online Information Society, 1–5.

Perez Vallejos, E., Koene, A., Portillo, V., Dowthwaite, L., & Cano, M. (2017). Young People's Policy Recommendations on Algorithm Fairness. *[WebSci2017]Proceedings of the 2017 ACM on Web Science Conference*, 247–251. doi:10.1145/3091478.3091512

Phethean, C. (n.d.). Assessing the Value of Social Media for Organisations : The Case for Charitable Use Uses of Social Media.

Phethean, C., & Harris, L. (2013). Rethinking Measurements Of Social Media Use By Charities : A Mixed Methods Approach.

Phethean, C., Simperl, E., Tiropanis, T., & Tinati, R. (2016). The Role of Data Science in Web Science.

Phillips, M. G. (2012). Designing for vigilance during intermittent use. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 243–246. doi:10.1145/2380718.2380749

Piccolo, L. S. G., Alani, H., De Liddo, A., Baranauskas, C. C., & Liddo, A. De. (2014). Motivating online engagement and debates on energy consumption. *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 109–118. doi:10.1145/2615569.2615695

Pickhardt, R., & Staab, S. (n.d.). Experiences from teaching Introduction to Web Science.

Pierre, J. (2013). Reverse Privacy Engineering.

Poirier, L. (2017). A Turn for the Scruffy: An Ethnographic Study of Semantic Web Architecture, 359–367. doi:10.1145/3091478.3091505

Powell, A. (2009). Lessons from the Net Neutrality lobby : Balancing openness and control in a networked society.

Preo, D., Cohn, T., & Kingdom, U. (2013). Mining User Behaviours : A Study of Check-in Patterns in Location Based Social Networks.

Preusse, J. (2012). Fairness on the Web : Alternatives to the Power Law, 175–184.

Proserpio, D., Counts, S., & Jain, A. (n.d.). The Psychology of Job Loss: Using Social Media Data to Characterize and Predict Unemployment, 223–232.

Qu, Y., Ge, W., Cheng, G., & Gao, Z. (n.d.). Class Association Structure Derived From Linked Objects Network Characteristics of CAG, 2–6.

Quercia, D. (2013). Don ' t Worry , Be Happy : The Geography of Happiness on Facebook.

Quercia, D., Askham, H., & Crowcroft, J. (2012). TweetLDA : Supervised Topic Classification and Link Prediction in Twitter, 247–250.

Quercia, D., Crowcroft, J., Bodaghi, M., & Crowcroft, J. (2012). Loosing "friends" on Facebook. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 251–254. doi:10.1145/2380718.2380751

Rahman, M., Recabarren, R., Carbunar, B., & Lee, D. (2017). Stateless Puzzles for Real Time Online Fraud Preemption. *Proceedings of the 2017 ACM on Web Science Conference*, 23–32. doi:10.1145/3091478.3091507

Ramchurn, S. D., Huynh, T. D., Venanzi, M., & Shi, B. (2013). Collabmap : Crowdsourcing Maps for Emergency Planning.

Rao, T. (2008). Modeling Movements in Oil , Gold , Forex and Market Indices using Search Volume Index and Twitter Sentiments, *I*.

Rheingold, H., & Mobs, S. (2009). i . e . Google Zero transaction costs “ Integrated user experience ” i . e . Blog comment rating “ user feedback forms ” I . e . eBay or Meetup Marginal transaction costs.

Richter, B. W., & Brown, I. (2008). Designing effective regulation for the “ Dark side ” of the Web, 4–8.

Rieder, B. (2013). Studying Facebook via Data Extraction : The Netvizz Application.

Riera, D. (2008). Science, the Web and Web Science. *Voc Papers*. Retrieved from http://uocpapers.uoc.edu

Rodríguez-ardura, I., Martínez-lópez, F. J., & Luna, P. (2008). The Consumer and the Web : a critical review of the contributions made by marketing and the consumer behaviour discipline to Web Science, 1–7.

Rogers, R. (2013). Debanalizing Twitter : The Transformation of an Object of Study.

Rogers, R. G., Hummer, R. A., & Nam, C. b. (2000). 2 - Data and Methods, 13–28. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780125931304500036\nhttp://www.sciencedirect.com.ezproxy.liv.ac.uk/science/article/pii/B9780125931304500036

Roure, D. De, & Goble, C. (2010). Anchors in Shifting Sand : the Primacy of Method in the Web of Data.

Roure, D. De, Hooper, C., Page, K., & Willcox, P. (2015). Observing Social Machines Part 2 : How to Observe ?

Rowe, M., & Alani, H. (2014). Mining and comparing engagement dynamics across multiple social media platforms. *Proceedings of the 2014 ACM Conference on Web Science - WebSci ’14*, 229–238. doi:10.1145/2615569.2615677

Rowe, M., & Ciravegna, F. (2010). Harnessing the Social Web : The Science of Identity Disambiguation.

Rowe, M., Keynes, M., & Hayes, C. (2012). Behaviour analysis across different types of Enterprise Online Communities, 255–264.

Rubart, J. (2008). Semantics through language sharing. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence - WebScience ’08*, 53. doi:10.1145/1379157.1379170

Sakurai, Y. (n.d.). How confident are you ? Worker classification in crowdsourcing.

Samory, M., & Peserico, E. (2016). Content attribution ignoring content. *Proceedings of the 8th ACM Conference on Web Science - WebSci ’16*, 233–243. doi:10.1145/2908131.2908156

Sanderson, R., Ciccarese, P., & Van de Sompel, H. (2013). Designing the W3C open annotation data model. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci ’13*, 366–375. doi:10.1145/2464464.2464474

Sariyuce, A. E., & Pinar, A. (2017). ϵ-WGX: Adaptive Edge Probing for Enhancing Incomplete Networks, 161–170. doi:10.1145/nnnnnnn.nnnnnnn

Sathanur, A. V. (2014). Annotation of Social Graphs, 187–191.

Schantl, J., Kaiser, R., Wagner, C., & Strohmaier, M. (2013). The utility of social and topical factors in anticipating repliers in Twitter conversations. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 376–385. doi:10.1145/2464464.2464481

Schi, D., Schneider, F., Schi, H., Schmid, S., & Uhlig, S. (2012). Tracing the Birth of an OSN : Social Graph and Profile Analysis in Google +, 265–274.

Schindler, M., V, W. D., & Vrandeˇ, D. (n.d.). Introducing new features to Wikipedia – Case studies for Web Science –.

Schümmer, T., & Haake, J. M. (2008). Supporting social practice learning in NGOs. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence - WebScience '08*, 17. doi:10.1145/1379157.1379163

Schwagereit, F., & Staab, S. (2011). Survey on Governance of User-generated Content in Web Communities.

Seitlinger, P., & Ley, T. (2016). Reconceptualizing imitation in social tagging: a reflective search model of human web interaction. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 146–155. doi:10.1145/2908131.2908157

Shadbolt, N. (n.d.). What Is Web Science? (Talk). Retrieved from http://webscience.org/professor-nigel-shadbolt-explains-web-science/

Shadbolt, N., Berners-Lee, T., & Shadbolt, N. (2008). Web Science Emerges. *Scientific American*, *299*(October), 76. Retrieved from http://eprints.ecs.soton.ac.uk/17143/

Shi, Y., Mast, K., Weber, I., Kellum, A., & Macy, M. W. (2017). Cultural Fault Lines and Political Polarization. *Proceedings of ACM Web Science Conference*, 8. doi:10.475/123 4

Shipman, F. M., & Marshall, C. C. (2013). Are User-contributed Reviews Community Property ? Exploring the Beliefs and Practices of Reviewers.

Shneiderman, B. (2007). Web Science : A Provocative Invitation to Computer Science. *Communications of the ACM*, *50*(6), 25–27.

Sillence, E., Hardy, C., & Briggs, P. (2013). Why don ' t we trust health websit es that help us help each other ? An analysis of online peer-to-peer healthcare.

Silva, C. S., Silva, I. S., Silva, T. S., & Mourão, F. (2017). Privacy for Children and Teenagers on Social Networks from a Usability Perspective : A Case Study on Facebook, 63–71.

Sizov, S., & Piller, S. (2016). Economic Value of Web Keyword Advertising Campaigns, 117–126. doi:10.1145/2908131.2908158

Smith, A., & Lewis, P. (2011). Visualising the Past : Annotating a Life with Linked Open Data.

Smith, G., & Ashman, H. (2008). Evaluating implicit judgements from Image search interactions.

Solé-ribalta, A., Domenico, M. De, & Gómez, S. (n.d.). Centrality Rankings in Multiplex Networks Categories and Subject Descriptors, 149–155.

Spaniol, M. (2017). ELEVATE : A Framework for Entity-level Event Diffusion Prediction into Foreign Language Communities, 111–120.

Spinelli, L., & Crovella, M. (2017). Closed-Loop Opinion Formation. *[WebSci2017]Proceedings of the 2017 ACM on Web Science Conference*, 73–82. doi:10.1145/3091478.3091483

Spiro, E. S., Fitzhugh, S., & Butts, C. T. (2012). Rumoring During Extreme Events : A Case Study of Deepwater Horizon 2010, 275–283.

Stalla-bourdillon, S., & Gilbert, L. (n.d.). A Content-Linking-Context Model for " Notice -and-take-down " Procedures, 161–165.

Study, C., Hus, E., & Diaz-aviles, E. (2011). Tracking Twitter for Epidemic Intelligence, 82–85.

Stuetzer, C. M., Carley, K. M., Koehler, T., & Thiem, G. (2011). The communication infrastructure during the learning process in web based collaborative learning systems.

Sugiura, L., Pope, C., & Webber, C. (2012). Buying unlicensed slimming drugs from the Web : a virtual ethnography, 284–287.

Syed, Z., Finin, T., Mulwad, V., & Joshi, A. (2010). Exploiting a Web of Semantic Data for Interpreting Tables, (April), 26–27.

Tang, M. (2015). Habits vs Environment : What Really Causes Asthma ?

Teng, C., & Arbor, A. (2012). Coevolution of Network Structure and Content, 288–297.

Thamm, M. (n.d.). When Politicians Tweet : A Study on the Members of the German Federal Diet.

Thomas, E., Pan, J. Z., Taylor, S., & Ren, Y. (2010). Lightweight Reasoning and the Web of Data for Web Science.

Thomas, L., Briggs, P., & Little, L. (2013). Location tracking via social networking sites. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 405–412. doi:10.1145/2464464.2501852

Thuermer, G., Roth, S., Luczak-Rösch, M., & O'Hara, K. (2016). Internet use, in- and exclusion in decision-making processes within political parties. *Proceedings of the 8th ACM Conference on Web Science*, 205–214. doi:10.1145/2908131.2908149

Tikhonov, A., & Bogatyy, I. (2015). What can be Found on the Web and How : A Characterization of Web Browsing Patterns.

Tinati, R., Madaan, A., & Hall, W. (2017). InstaCan: Examining deleted content on Instagram. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 267–272. doi:10.1145/3091478.3091516

Tinati, R., Halford, S., Carr, L., & Pope, C. (2012). Mixing methods and theory to explore web activity. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 308–316. doi:10.1145/2380718.2380758

Tinati, R., Luczak-roesch, M., Simperl, E., & Hall, W. (2016). "Because Science is Awesome": Studying Participation in a Citizen Science Game. *ACM Web Science 2016*, 45–54. doi:10.1145/2908131.2908151

Tinati, R., Luczak-roesch, M., Simperl, E., & Shadbolt, N. (2014). "/ Command " and Conquer : Analysing Discussion in a Citizen Science Game, (August).

Tinati, R., Philippe, O., Pope, C., Carr, L., & Halford, S. (2014). Challenging Social Media Analytics : Web Science Perspectives, 177–181.

Tiropanis, T., Davis, H., Millard, D., & Weal, M. (2009). Semantic Technologies for Learning and Teaching in the Web 2 . 0 Era : A survey of UK Higher Education Web Science 2009 Conference - Athens , March 2009, (March).

Tolksdorf, R. (n.d.). Art As A Source For Innovation In Knowledge Processing.

Torres, D., Molli, P., & Alicia, D. (2013). BlueFinder : Recommending Wikipedia Links Using DBpedia Properties.

Totti, L. C., Costa, F. A., Avila, S., Valle, E., Meira, W., & Almeida, V. (2014). The impact of visual attributes on online image diffusion. *Proceedings of the 2014 ACM Conference on Web Science - WebSci '14*, 42–51. doi:10.1145/2615569.2615700

Tran, T. (2012). Pay-less Entity Consolidation – Exploiting Entity Search User Feedbacks for Pay-as-you-go Entity Data Integration, 317–325.

Treem, J. W., & Thomas, K. Y. (2010). Beyond Convergence : User Evaluations of Identical Syndicated News Content Across Media.

Tsakalakis, N., O'Hara, K., & Stalla-Bourdillon, S. (2016). Identity assurance in the UK. *ACM Conference on Web Science*, 55–65. doi:10.1145/2908131.2908152

Tunkelang, D. (2014). Web Science : How is it Different ?, 4503.

Tvarožek, M., & Bieliková, M. (2008). Collaborative multi-paradigm exploratory search. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence*, 29–33.

Urrutia, M. L., & Blair, R. (n.d.). Web Science for the masses .

van Atteveldt, W., Sheafer, T., & Shenhav, S. (2013). Automatically extracting frames from media content using syntacting analysis. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 423–430. doi:10.1145/2464464.2464504

Varol, O., Ferrara, E., Ogan, C. L., Menczer, F., & Flammini, A. (2013). Evolution of Online User Behavior During a Social Upheaval, (i), 81–90.

Vedula, N., & Shalin, V. L. (2017). Predicting Trust Relations Within a Social Network : A Case Study on Emergency Response, 53–62. doi:10.1145/3091478.3091494

Veloso, A., & Ferraz, F. (2011). Dengue surveillance based on a computational model of spatio-temporal locality of Twitter.

Venkatanathan, J., Karapanos, E., Kostakos, V., & Gonçalves, J. (2012). Network, personality and social capital. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 326–329. doi:10.1145/2380718.2380760

Victor, P., Cornelis, C., & Teredesai, A. M. (n.d.). Trust- and Distrust-Based Recommendations for Controversial Reviews.

Vigil, J., Tait, A. S., Wienberg, C., & Gordon, A. S. (2014). Friends You Haven ' t Met Yet : A Documentary Short Film, 2617797.

Vigo, M. (2013). Considering People with Disabilities as Uberusers for Eliciting Generalisable Coping Strategies on the Web, 1–4.

Vo, K. D., Tran, T., Nguyen, T. N., Zhu, X., & Nejdl, W. (2016). Can we find documents in web archives without knowing their contents? *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 173–182. doi:10.1145/2908131.2908165

Voskresenskiy, V., Musabirov, I., & Alexandrov, D. (2016). Private and public online groups in apartment buildings of St. Petersburg. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 301–306. doi:10.1145/2908131.2908173

Wang, N., Giles, S., She, J., & Chen, J. (2012). How " Big Vs " dominate Chinese Microblog A comparison of verified and unverified users on Sina Weibo, (July), 182–186.

Wang, W. (2008). PowerMeeting on CommonGround: Web based synchronous groupware with rich user experience. *Hypertext 2008 Workshop on Collaboration and Collective Intelligence 2008, WebScience'08*, 35–39. doi:10.1145/1379157.1379166

Wang, W., Mokhtar, M., & Macaulay, L. (2008). C-index: Trust depth, trust breadth, and a collective trust measurement. *Hypertext 2008 Workshop on Collaboration and Collective Intelligence 2008, WebScience'08*, 13–16. doi:10.1145/1379157.1379161

Wang, W., & French, S. (2008). A multi-dimensional framework for facilitating wide participation and common understanding. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence - WebScience '08*, 23. doi:10.1145/1379157.1379164

Wang, Y., Chakrabarti, A., Sivakoff, D., & Parthasarathy, S. (2017). Fast Change Point Detection on Dynamic Social Networks, 171–179. Retrieved from http://arxiv.org/abs/1705.07325

Webb, H., Jirotka, M., Stahl, B. C., Housley, W., Edwards, A., Williams, M., … Burnap, P. (2017). The Ethical Challenges of Publishing Twitter Data for Research Dissemination. *Proceedings of the 2017 ACM on Web Science Conference - WebSci '17*, 339–348. doi:10.1145/3091478.3091489

Weber, I., Borra, E., & Garimella, V. R. K. (2012). Mining web query logs to analyze political issues. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, 330–334. doi:10.1145/2380718.2380761

Weber, M. S. (n.d.). Big Data ? Big Issues : Degradation in Longitudinal Data and Implications for Social Sciences.

Webster, J., Gibbins, N., Halford, S., & Hracs, B. J. (2016). Towards a Theoretical Approach for Analysing Music Recommender Systems As Sociotechnical Cultural Intermediaries. *Proceedings of the 8th ACM Conference on Web Science*, 137–145. doi:10.1145/2908131.2908148

Welch, J. R., & Halford, S. (n.d.). Information and Communication Technologies (ICTs) and Peacebuilding : a Conceptual Framework.

Weller, K., & Kinder-Kurlanda, K. E. (2016). A manifesto for data sharing in social media research. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, *0049*(0), 166–172. doi:10.1145/2908131.2908172

West, A. G. (2017). Analyzing the Keystroke Dynamics of Web Identifiers. *Proceedings of the 2017 ACM on Web Science Conference*, (3), 181–190. doi:10.1145/3091478.3091482

White, S., Coskun, E., & White, S. (2014). Web Science : learning from experience.

White, S., Coskun, E., & White, S. (2014). Web Science : learning from experience.

White, S., Croitoru, M., Bazan, S. S., Cerri, S., Davis, H. C., Jonquet, C., … Vafopoulos, M. Negotiating the Web Science Curriculum Development through Shared Educational Artefacts, ACM WebSci '11 11:1–11:8 (2011). Koblenz, Germany, NY, USA: ACM. doi:10.1145/2527031.2527046

White, S., Davis, H. C., & Dickens, K. (2014). Making a MOOC: an interdisciplinary approach to teaching Web Science. Retrieved from http://eprints.soton.ac.uk/363095/1/WhiteEtAl2014HEA-STEM_MOOCs.pdf

White, S., & Vafopoulos, M. N. (2012). Web Science: expanding the notion of Computer Science. In *43rd ACM Technical Symposium on Computer Science Education, 28th February-3rd March 2012*. Rayleigh, NC. Retrieved from http://eprints.ecs.soton.ac.uk/22710/

Wienberg, C., & Gordon, A. S. (2015). Insights on Privacy and Ethics from the Web ' s Most Prolific Storytellers.

Wienberg, C., Roemmele, M., & Gordon, A. S. (2013). Content-based similarity measures of weblog authors. *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 445–452. doi:10.1145/2464464.2464483

Willcox, P., Glaser, H., & Roure, D. De. (n.d.). Archetypal Narratives in Social Machines : Approaching Sociality through Prosopography.

Wu, J., Teregowda, P., Fern, J. P., Mitra, P., Zheng, S., & Giles, C. L. (2012). The Evolution of a Crawling Strategy for an Academic Document Search Engine : Whitelists and Blacklists, 340–343.

Wu, Z., Jiang, L., & Zheng, Q. (2010). A Peep at Pornography Web in China.

Xu, J., Compton, R., Lu, T., & Allen, D. (n.d.). Rolling through Tumblr : Characterizing Behavioral Patterns of the Microblogging Platform Categories and Subject Descriptors, 13–22.

Xu, J., Xie, D., Lu, T.-C., & Cafeo, J. (2017). EDSV: Emerging Defect Surveillance for Vehicles. *Proceedings of the 2017 ACM on Web Science Conference*, 219–222. doi:10.1145/3091478.3091521

Yang, C., & Srinivasan, P. (n.d.). Translating Surveys to Surveillance on Social Media : Methodological Challenges & Solutions Categories and Subject Descriptors, 4–12.

Yeung, C. A. (2010). Analysis of Strategies for Item Discovery in Social Sharing on the Web Delay in Item Adoption.

Yeung, C. A., Noll, M. G., Gibbins, N., Meinel, C., & Shadbolt, N. (n.d.). On Measuring Expertise in Collaborative Tagging Systems.

Yip, M., Shadbolt, N., & Webber, C. (2013). Why Forums ? An Empirical Analysis into the Facilitating Factors of Carding Forums.

Yip, M., & Webber, C. (2011). Hacktivism : a Theoretical and Empirical Exploration of China ' s Cyber Warriors.

Zagheni, E. (2012). You are where you E-mail : Using E-mail Data to Estimate International Migration Rates, 348–357.

Zagovora, O., Flöck, F., & Wagner, C. (2017). "(Weitergeleitet von journalistin)": The gendered presentation of professions on Wikipedia. *WebSci 2017 - Proceedings of the 2017 ACM Web Science Conference*, 83–92. doi:10.1145/3091478.3091488

Zawilska, A., Ox, U. K., & Albury, S. (2015). An Ethnomethodologically-Informed Approach to Interface Design to Support Collective Web Practice Around Video.

Zhang, H., & Setty, V. (2016). Finding diverse needles in a haystack of comments. *Proceedings of the 8th ACM Conference on Web Science - WebSci '16*, 286–290. doi:10.1145/2908131.2908168

Zhang, L., & Tu, W. (n.d.). Six Degrees of Separation in Online Society, 1–5.

Zhou, Y., Guan, X., Zhang, Z., & Zhang, B. (2008). Predicting the tendency of topic discussion on the online social networks using a dynamic probability model. *Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence - WebScience '08*, 7. doi:10.1145/1379157.1379160

Zhou, Y., & Cristea, A. I. (2016). Towards Detection of Influential Sentences Affecting Reputation in Wikipedia, 244–248. doi:10.1145/2908131.2908177

Zollinger, M. (2013). Market-based SPARQL Brokerage : Economic Incentives for Linked Data.

Web Science Institute A Year of Progress. (n.d.).

A celebration of 10 Years of the Science of the Web CELEBRATING 10 YEARS OF WEB SCIENCE. (2016).

Wilson, C. (2001), *191*, 2001.

13 Appendix

13.1 Web Science Subject Categorization – Full Subject Listing

Web Science Subject categorization subject listing (Vafopoulos 2010)
A. General A.1 Instructional exposition A.2 research exposition A.3 General and miscellaneous specific topics A.4 Conference proceedings and collections of papers
B. Web History and Methodology B.1 General Web History and Methodology B.2 Web History Web Forerunners; Biographies and related stories B.3 Web Science Theory and Epistemology Two Magics of Web Science; Actor Network Theory
C. Web Technologies C.1 General Web Technologies C.2 Web Milieux Document technologies; Hypertext technologies; Internet technologies; Mobile Web technologies; Grid and Cloud computing technologies C.3 Basic Web Architecture HTTP and related technologies; URIs; HTML; XML;CSS and related technologies; Interfaces and Browsers; Servers Web Services C.4 Web 2.0 technologiesC.5 Semantic Web/Linked Data Metadata; Knowledge Representation; Ontology Languages; Linked Data; Natural Language Processing; Provenance systems in the Web C.6 Internet/Web of Things
D. Web Analysis D.1 General Web Analysis D.2 Mathematical Methods of Web analysis Web data sampling and analytics; Logic and Inference in the Web; Statistical Inference in the Web; Statistical Analysis of the Web; Web as a Complex System; Graphs; Networks; Mathematical methods for describing Web services; Crawling; Indexing and Searching; Data Mining; Information Retrieval and Machine Learning; Other Algorithms for the Web
E. Web Society E.1a Economics Goods in the Web; The Web economy; Antitrust Issues and Policies in the Web; Intellectual property and digital rights management; Web-based economic development E.1b Business E-commerce Business models in the Web; Advertising in the Web; sponsored search E.2 Social Engagement and Social Science Social networks; Mass phenomena; Collective intelligence; Peer production; Globalization; Systems; Social structures and processes; Virtual communities, groups and identity; Social capital and power inequality in the Web; On-line lives, intergenerational differences; Mass media E.3 Personal Engagement and Psychology System Psychology and Behaviour; Child and adolescent psychiatry; Tele-working E.4 Philosophy Philosophy of information; Objects; Reference and Cognition in the Web; Ethics in the Web E.5 Law Intellectual Property in the Web; Digital Rights Management; Digital crime; Laws for Web access; Antitrust Law E.6 Politics and Governance Political science; E-Government; E-Politics; E-Democracy; Policy and Regulation; Web Governance; Privacy; Trust; Security; Network neutrality; E-Inclusion
F. Teaching the Web F.1 Teaching the Web – General F.2 Pre-college teaching F.3 Undergraduate teaching F.4 Graduate teaching

14 Overview of Learning Taxonomies

14.1 Overview of Learning Models

When designing a curriculum, it is not simply comprised by the subjects which are taught. The learning artefacts (e.g. the lectures, the assignments, the group work projects) have affordances and certain skills which are taught. The combination of learning styles which comprise a curriculum are arguably as important as the subjects that are taught. Traditional curriculums such as Computer Science and Information Science include a number of standard core modules and key skills which are required in order to meet the required learning outcomes.

Learning models provide a benchmark and point of comparison which may aid the identification of learning objectives and 'core modules' taught within the existing Computer Science and Information Science curricula. This may then be applied and used as a point of comparison to identify skills within the Web Science programmes identified.

Table 61. Includes an exert from the National curriculum for England guidelines for school key stages 1 to 4. While this is a very different level to that of university study, the principles for meeting learning objectives remain the same. The guidelines recognise that different learning approaches may be required for students with different needs. While this is referring to meeting the needs of students with disabilities, it recognises that different learning approaches may be required to meet the needs of different student. This is also true of the fact that different learning approaches are required to teach students different skills.

Table 61. National curriculum in England: framework for key stages 1 to 4

National Curriculum in England Guidelines [65]
"*A wide range of pupils have special educational needs, many of whom also have disabilities. Lessons should be planned to ensure that there are no barriers to every pupil achieving. In many cases, such planning will mean that these pupils will be able to study the full national curriculum. The special educational needs and disability code of practice includes advice on approaches to identification of need which can support this. A minority of pupils will need access to specialist equipment and different approaches*."

The style in which students have learnt will have an impact upon the skills that they demonstrate upon competition of their programme. For example the experiences of a student who has participated in a traditional taught programme featuring lectures and group work will be quite different to the experience gained by a student who has completed a largely online programme. The following section provides a brief overview of learning taxonomies, which serve as a measure of learning methods skills. Gaining an understanding of these will aid in the identification of learning methods within Web Science taught curricula.

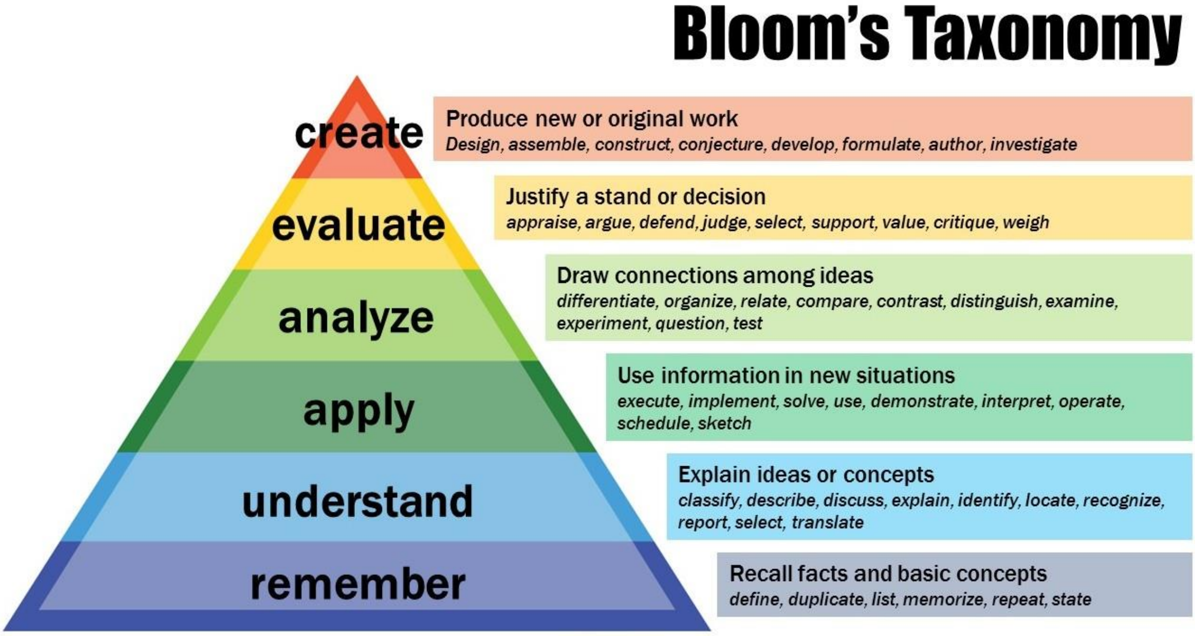
14.1.1 Blooms Taxonomy

Blooms Taxonomy is an effort to categorise types and levels of learning. It includes "*three "domains" of learning, each of which is organised as a series of levels or pre-requisites."* [66] The different tiers illustrate the level of expertise necessary to achieve each measurable learner objective. Figure 11 shows Blooms Taxonomy as revised by (KRATHWOHL & ANDERSON 2001).

[65] https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4/the-national-curriculum-in-england-framework-for-key-stages-1-to-4

[66] http://www.learningandteaching.info/learning/bloomtax.htm

Figure 11. Blooms Taxonomy [67]



The following Table 62 shows a breakdown of the original levels of Blooms Taxonomy in more detail.

Table 62. Blooms Taxonomy Structure (KRATHWOHL & ANDERSON 2001)

The Structure of Blooms Taxonomy
1.0 Knowledge
1.10 Knowledge of specifics
1.11 Knowledge of terminology
1.12 Knowledge of specific facts
1.20 Knowledge of ways and means of dealing with specifics
1.21 Knowledge of conventions
1.22 Knowledge of trends and sequences
1.23 Knowledge of classifications and categories
1.24 Knowledge of criteria
1.25 Knowledge of methodology
1.30 Knowledge of universals and abstractions in afield
1.31 Knowledge of principles and generalizations
1.32 Knowledge of theories and structures
2.0 Comprehension
2.1 Translation
2.2 Interpretation
2.3 Extrapolation
3.0 Application
4.0 Analysis
4.1 Analysis of elements
4.2 Analysis of relationships
4.3 Analysis of organizational principles
5.0 Synthesis
5.1 Production of a unique communication
5.2 Production of a plan, or proposed set of operations
5.3 Derivation of a set of abstract relations
6.0 Evaluation
6.1 Evaluation in terms of internal evidence
6.2 Judgments in terms of external criteria

[67] https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

14.1.2 Biggs Model of Constructive Alignment

The fundamental principle of Constructive alignment is 'outcomes-based teaching and learning'. The learning process is process 'Constructive' "*because it is based on the constructivist theory that learners use their own activity to construct their knowledge or other out-come*" (Biggs 2007). Who also explain that *"The idea of aligning assessment tasks with what it is intended that students should learn is very old – and very obvious. It's called 'criterion-referenced assessment'"*.

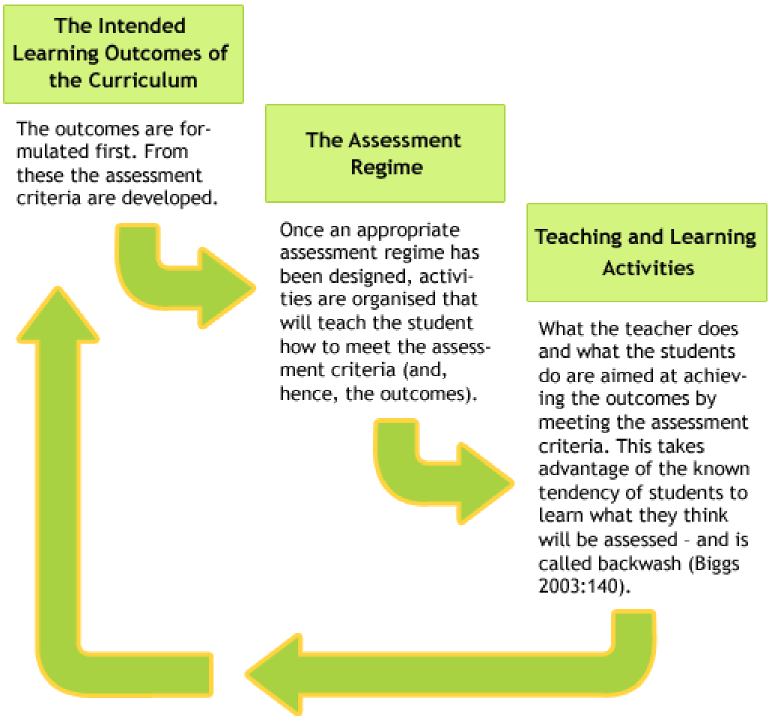


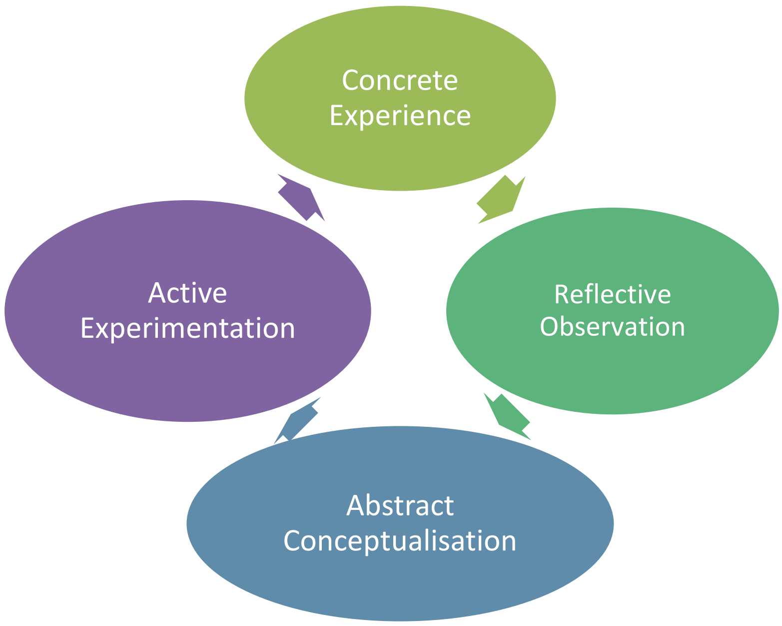
Figure 12. A Basic Model of an Aligned Curriculum [68]

[68] http://www.ucdoer.ie/index.php/Using_Biggs'_Model_of_Constructive_Alignment_in_Curriculum_Design/Introduction

14.1.3 Kolb's Experiential Learning Cycle

(Kolb 1984) describes the Experiential Learning Cycle as a "*holistic perspective on learning that combines experience, perception, cognition and behaviour*."

Figure 13. Kolb's Experiential Learning Cycle



The learner will first experience something, this will then be followed by personal reflection, which then results in the derivation of rules describing the experience, resulting in experimentation, which leads to the next 'cognitive experience' [69].

[69] http://www.learningandteaching.info/learning/experience.htm

15 Southampton Future Learn Web Science MOOC Experience

15.1 Web Science MOOC Experience

During the first semester of the author's Web Science PhD, the University of Southampton launched a Web Science MOOC[70] (Massive Online Open Course), which provided a whole new experience and style of learning. The MOOC consisted of six structured weeks of material on different areas of Web Science, providing a broad overview of the Web Science subject. The author found the first week particularly inspiring, particularly the opening discussion video entitled 'What is Web Science?', which explained this key question clearly and concisely, confirming some of the author's existing conceptions of Web Science, and also providing some new perspectives. The MOOC added a new dimension to learning; as well as being able to pause and playback the video 'lectures', at the viewer's convenience, there was also the added benefit of feedback and discussion from other MOOC users. This provided some useful insights into what others thought of various topics. The structure of the MOOC as well as an overview of the content is outlined in the table below. Additionally, a screenshot showing the MOOC interface is shown on the following page.

Table 63. An overview of the University of Southampton MOOC Content

Week 1 **What is web science?**	*Week2* **Networks**	*Week3* **Crime and Security**
• Personal Impact of the Web • The Pre-History and Invention of the Web • The Recent History of the Web • Image representing Web Science • Constructing the Web	• What are Networks? • Introduction to Network Properties • Towards a Science of Networks • Is 'Network Thinking' the best paradigm for Web Science? • Network image • Network Analytics • Reflection - Types of Power in a Social Network • Power, Influence and Action	• Crime, Trust and the Web • 'Drifting on and off-line: Humanising the cybercriminal' • Not just a 'cyber' crime? • Cryptocurrencies and Crime • Security Economics: Why Cybercrime is • Discussion: Botnets on the Web • How Criminals Use the Web • Buying (Prescription) Medicine from the Web • Discussion: Drugs on the Web
Week 4 **Democracy**	*Week 5* **Economy**	*Week 6* **What next for the Web?**
• Democracy and the Web • How does industrial action change with the Web? • Democracy image • Introduction to open government data • The role of open data in government • Technologies for open data • Police UK website - how do people interpret data visualisations? • discussion • e Web and activism • The Web is increasingly viewed as a platform for activism. This activity will look at the role of the Web in activism and also examine some of the	• Introduction to digital business • How digital data is changing the way we do business • Types of digital data • Introduction to 'big data' • How counting votes is a 'big data' problem • Tools for big data • Big data - 'hope' or 'hype'? • Personal data in the hands of individuals • How valuable is your Personal Data? • How much personal and professional data should you make freely available? • What's changed and what are employers now doing differently? • So what should job seekers be doing?	• The Future of Web Science • The future of knowledge online: the semantic web • The future of gadgets online: the Internet of Things • The next 5,000 days of the Web... • If I ruled the Web ...

[70] http://www.ecs.soton.ac.uk/news/4315

research methods that can be deployed in the study of this. • The Web and activism: an introduction • The Web and the 2009 Iranian election • The Arab Spring • The Occupy Movement • Investigating 'Stop G8!' and the mistrust of social media • How has the Web changed activism? • Researching activism on the Web • How easy is it to study peoples' activities on the Web?	• Can social networking help you get your dream job? • An introduction to the value of social media for businesses • What are the advantages for businesses of using social media for marketing? • Problems of measuring social media • Social media data values • Task: What can we learn from network visualisations? • Venture creation in the digital age	

Figure 14. The University of Southampton Web Science MOOC Interface

The study of this MOOC introduced the author to the key concept that Web Science includes a very clear emphasis on the social science of the Web, and how understanding the effect that the Web has on people is just as important as the technology behind it. While this was something that the author was previously aware of, the valuable and considerable proportion of input from social scientists such as Professor Susan Halford helped to explain it with a far clearer emphasis. Another example of a valuable lesson that the author learnt from her experience studying the MOOC, was that in order to understand the Web as it is today, it is important to understand the origins of the Web, and why it works the way that it does. This is a perspective that the author had not considered prior to the study of this MOOC, and was hugely valuable as a background better to understanding the Web. As mentioned previously, there are currently only three dedicated Web Science MOOCs, however, the comparison of these might provide some interesting results.

16 Web Science Teaching Institutions

List of Web Science Teaching Institutions

The following section includes an exert from the desk survey identifying Web Science teaching institutions across the globe. The rating explained bellow is simply an indication based upon the authors initial assessment, suggesting if a programme may be either heavily biased towards technical, social, or a so-called 'mixed approach' programme. It is not scientifically, proven, it is simply given as a guide, and is subject to revisions.

Location	Course Title	Level	Type	Comments\Observations
Aristotle University of Thessaloniki - Thessaloniki, Greece	Masters in Web Science	Masters	T	A Masters degree in Web Science is mentioned. This MSc course would seem to have a heavy bias towards the technical and mathematical elements of Web Science – almost a Web Technologies course; there is little evidence of the social impact of the Web being taught. Link to course content: http://www.webscience.auth.gr/?q=content/courses
British university in Egypt - EL SHEROUK CITY, Cairo, Egypt	Informatics and Computer Science MSc (ICS)	Masters	T	The content is described as follows: "Web Languages Technologies, Net-Centric Computing, Services Oriented Software Engineering, Introduction to Web Science" ICS currently offers a Master of Science (MSc.) degree in Web Science. The objective of the degree is to "understand how the Web evolves and to understand how we can make the Web a better tool for our individual needs, for society, and for our organisations. For this purpose, we must look at the Web from an interdisciplinary perspective. The merging of computer related disciplines and information technologies has characterised the information age and resulted in the emergence of huge global information markets, where the demand is to provide high volume and high speed information traffic for an ever more information conscious society through the Web. The Master of Web Science which is offered to graduate students is unique in the region. It combines four specialisation currently offered by the faculty of Informatics and Computer Science: Computer Science, Information Systems, Software Engineering, and Computer Networks. All of the four specialisations are blended to serve state of the art technologies and resarches targeted to Web sciences addressed from different prospects."
Cologne University, Germany	MSc Web Science	Masters	M	This institution offers a Web Science MSc course. The course website describes Web Science as the need for an "interdisciplinary approach and of focusing on the possibilities of the "Web" as the interconnection of people, services and systems." The modules offered include: "Design, Web Trust & Security, Decision & Management, Web & Cooperation, Web & Society. This would seem to suggest a fairly web rounded curriculum with slightly more emphasis upon the social impact of technology.
Eindhoven University of Technology	(BSc) Major program Web Science	Undergrad	M	This is an example of an undergraduate degree in Web Science. The course description explains that "Web Science combines computer science with other disciplines such as psychology and sociology. This enables us to understand how online communication is used." This would seem to suggest that this is a course with a 'multi-disciplinary' focus, combining both the social and the technical aspects of Web Science.
Georgia Tech University – Web Science Courses	WS Modules available		NA	The website explains that there is a forthcoming course planned. However, current students interested in Web Science can take current modules in Web Science which are available as part of other degree programmes.
Goldsmiths London	MA/MSc in Digital Sociology	Masters	S	Focuses upon providing a Social perspective on the use of digital technology, such as 'big data'.
Johannes Kepler University Linz	Master's Degree in Web Sciences	Masters	M	This intuition provides a Masters degree dedicated to Web Science, which is described as providing an "in-depth look at fields relevant to the web such as technology, business, law, society, art and culture".
Korea Advanced Institute of Science and Technology (KAIST)	Division of Web Science and Technology	Masters/ PhD	T	Masters and PhD Programmes. Description: IT industry, a core growth power for the national advancement in the future, is evolving from manufacturing hardware to developing software and applications. Thus, KAIST shifts its educational paradigm of cultivating leading IT experts by more focusing on software and embarks on the graduate school of Web Science & Technology under the umbrella of College of Information Science & Technology.
MIT - Massachusetts Institute of Technology	The MIT Human Dynamics Lab		?	This institution appears to have no defined Web Science programme, however it has links with a Web Science collaboration at the University of Southampton. :www.ecs.soton.ac.uk/news/1047 It also receives a mention on the Web Science Trust Website: http://webscience.org/wstnet-laboratories/wstnet-labs/mit-human-dynamics-lab/

Location	Course Title	Level	Type	Comments\Observations
Northwestern University School of Communication	'Doctor of Philosophy in Media, Technology, and Society'	PhD	S	This is another example of a PhD which studies the impact of technology on society.
Oxford Internet Institute	'*MSc in Social Science of the Internet'* and *'doctoral programme (DPhil) in Information, Communication and the Social Sciences'*	Masters and PhD	M	The Oxford Internet Institute includes both a masters and a PhD course which look at Web Science from a Social Science perspective. (Also includes a summer doctoral programme)
Rensselaer Polytechnic Institute	'Information Technology and Web Science'		T	This institution offers a BSc course with a Web Science component. This is an example of a course which approaches Web Science from an IT perspective. There are also opportunities to study further at PhD level.
RWTH Aachen University	Web Science - Informatik 5	Unknown	NA	Includes some modules dedicated to Web Science - level unknown
Saint-Joseph University of Beirut	Master in Web Science and Digital Economy	Masters	NA	This course allows Digital Economics MSc students to specialize in Web Science, providing students with the opportunity to study "The impact of the Web and its technologies in the economic and financial world".
The University of Edinburgh: School of Social and Political Science	'MSc in Science and Technology in Society'	Masters	S	The MSc course specifically looks at the impact of technology on society. Describes technology as a powerful 'agent of change', which needs to be studied in order to determine its impact.
UAH MediaLab, University of Alcalá (Spain)	The Digital Networked Media and Web Science Master of Science, and The Communication, Information and Technology in the Networked Society Ph.D.	Masters/ PhD		UAH Media Lab is an open space for research and creation in the intersection of technology, culture and social studies. The Lab emphasizes research across disciplines and disruptive ideas on how to people use and interact with technology. The Lab was created in 2013 as a vehicle to bring together scholars and projects from different disciplines in the University of Alcalá, including Computer Science, Communication Studies, Information Science and Humanities.
University College London	MSc Web Science and Big Data Analytics and MRes Web Science and Big Data Analytics	Masters	T	This is another example of a more technically biased Web Science course, which gives a good cross section of subjects from "information search and retrieval, data mining and analytics, large-scale distributed and cloud computing, to e-commerce". While these technologies are covered comprehensively, there is little suggestion that the social impact is considered.
University of Erlangen-Nürnberg	Innovation Technology (Master)	Masters	NA	This 'course' contains related modules – difficult to determine if this is a full masters course or just referring to one lecture\module.
University of Koblenz-Landau, Institute for Web Science and Technologies,	MSc Web Science	Masters	M	This MSc course would seem to offer a 'fully rounded' Web Science curriculum looking at both the technological and social aspects of Web Science. Topics mentioned include: computer science, social science, economics, and studies of the law, linguistics, semiotics, communication studies, and mathematics.
University of Liverpool	MSc in Web Sciences and Big Data	Masters	M	This is an example of an online MSc course. This online Masters programme adopts an interdisciplinary approach encompassing web engineering and more general management skills. Designed in consultation with industry experts, the course covers all the most sought-after skills and competencies needed for leadership in this field. You will explore web development technologies in depth, covering social computing, cloud computing and mobile and grid computing. The programme also covers key aspects of big data infrastructure, data analytics, data modelling and business intelligence.
University of Southampton, UK	'BSc Web Science' '*MSc Web Science'* and *'Web Science Doctoral Training Programme'*	Undergrad, Masters, PhD	M	Most detailed syllabus and widest range of courses discovered. Southampton is unusual in the fact that it offers Web Science courses right through from undergrad to PhD level. Undergrad level offers two pathway options, one with a bias toward social science, and the other with a bias toward computer science.

*Type: M=Mixed approach S= Social Science focused T=Technology focused

16.1 Current Desk Survey Results – Modules

Location	Resource Title	Topic(s)	Course Teaching\Assessment Format	Contributors(s)	Materials	Level	Duration	Language
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS.01 Web Science	Web Development, Epistemology and Didactics, Network Operation Centers. The AUTH Network, e-Services and semantic AUTH, Web Security. Attacks and Defence,Research Methodology,Project Management,Web Economics and Business,Web and the Law,Social Networks and Data Journalism, Social Networks and Semantic Web Challenges, Workshop Wikipedia and DBpedia Projects, Privacy and Trust in the Web	Classes\Workshops	Antoniou I. , Bratsas C., F. Loukos, Metakides G., Nouskalis G., Petridou S. , Tzounakis P., Varsakelis N., Veglis, Stampoulis	http://cosynet.auth.gr/node/475	Masters	Class dates: 10-01-2014, 17-01-2014, 24-01-2014, 07-02-2014, 14-02-2014, 28-02-2014, 07-03-2014, 14-03-2014, 21-03-2014, 28-03-2014, 04-04-2014,	Greek/English
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS.04 Networks and Discrete Mathematics	Fundamentals of Graphs and Networks, Random Graphs, Descriptive Statistics on Real-World Networks. Distances, Centrality, Correlation, Communities, Small Worlds and Scale-Free Networks, Workshop on Graphs and Descriptive Statistics on Social, Biological Networks and the Web, Information, Entropy, Probability, Statistics, Channels and Coding, Sources of Information, Coding and Cryptography	Classes\Workshops, Assessment: examinations: Networks 50% , Information 50%	Antoniou I., Karagiannis V., Moyssiadis P.	http://cosynet.auth.gr/node/476	Masters	Class dates: 01-11-2012, 08-11-2012, 08-11-2012, 22-11-2012, 29-11-2012, 06-12-2012, 13-12-2012, 20-12-2012, 24-01-2013, 31-01-2013	Greek/English
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS.13 Web Languages and Technologies	Web of Trust in Semantic Web, Web 2.0 Technologies (Mash up, Web analytics, Web Markup Languages, Client/Server-side Web Programming), Services and Architectures for the Web and Semantic Web, Web Information Retrieval, Search Engine Technologies, Web Mining, Cloud Computing - Internet of Services, Web Metadata standards, RDF, RDF-S, OWL, Ontology Design and Development, Graph pattern matching and SPARQL Query Language	Classes. Assessment: examination	Bratsas C., Kehagias D., Stamatiou I.	http://cosynet.auth.gr/node/477	Masters	Class dates: 30-10-2012, 06-11-2012, 13-11-2012, 20-11-2012, 27-11-2012, 04-12-2012, 11-12-2012, 18-12-2012	Greek/English
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS05 Statistical Analysis of Networks	Introduction to networks, correlation and time series, Measure of correlation, complexity and coupling for time-series, Multivariate time series and networks, Network Sampling, Samplingnd Web (Traffic Ranking, Distributions, Periodicities), Game Theory on Networks, Semi Markov Models with Rewards, Network Evolution,	Classes. Assessment: examination	Antoniou I., Farmakis N., Kugiumtzis D., Papadopoulou A., Spirakis P.	http://cosynet.auth.gr/node/506	Masters	Class dates: 05-03-2013, 12-03-2013, 19-03-2013, 26-03-2013, 02-04-2013, 09-04-2013, 16-04-2013, 23-04-2013, 14-05-2013, 28-05-2013, 04-06-2013	Greek/English
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS.11 Knowledge Processing in the Web	Logic and Programming in the Web, Logic and programming in the Web Workshop, Semantic Processing and Reasoning, Linked Data on the Web, Future Internet and Linked data, Linked Open Data Challenges, Big Data, Logics and Ontologies for the Semantic Web,	Classes\Workshops, Assessment: examination	Bratsas C., Kehagias D., Metakides G., Nicoletseas S.	http://cosynet.auth.gr/node/507	Masters	Class dates: 08-03-2013, 15-03-2013, 22-03-2013, 29-03-2013, 05-04-2013, 12-04-	Greek/English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
							2013, 19-04-2013, 26-04-2013, 31-05-2013	
Aristotle University of Thessaloniki - Thessaloniki, Greece	WS18 Biological Networks	Biological Networks Introduction, Biomedical Ontologies. Semantic Applications, Services and Repositories, Statistical Analysis of Biological Data, Brain Networks, Maps and Connectomics, Mutualistic Networks and Food Webs, Macroevolution- Projects, Landscape Connectivity-Metapopulation Theory Networks of Protected Areas	Classes. Assessment: examination	Antoniou I., Bratsas C., Kehagias D., Mazaris A., Bamidis C., Pantis J., Sgardelis S., Scouras Z.	http://cosynet.auth.gr/node/508	Masters	Class dates: 07-03-2013, 14-03-2013, 21-03-2013, 28-03-2013, 04-04-2013, 11-04-2013, 18-04-2013, 25-04-2013, 13-06-2013	Greek/English
British University in Egypt	Web Modeling (Core module)	Web Modeling. Web Statistics. Stochastic Processes on Networks. Web Traffic. Markov Networks. Diffusion on Networks. Semi-Markov Processes and Reward. Percolation Processes. Web Evolution	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Web Languages and Technologies (Optional module)	Basic Internet Protocols. Basic Web technologies (http, html, URI). XML-based languages. Client-side Web Programming (client-side JavaScript). Server-side Web Programming (Java Servlets, PHP). Asynchronous client-server model. Web Information Retrieval algorithms & Search engines Technologies. Web 2.0 Technologies and Applications Advanced Internet Protocols and applications (Multicasting Technologies). Multimedia Content Distribution Protocols over IP. Mobile Web Technologies - Social networking	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Software Agents (Optional module)	Reasoning. Agent Adaptability. Interface Agents. Belief Desire Intention. Agent Communication Language. Multi-agent Systems. Agent Programming Paradigms. Platforms for Software Agents. Agent Trust. Agent Patterns. Agent Frameworks and Architectures. Mobile Agents. Agent Applications	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
British University in Egypt	Parallel and Distributed Systems (Optional module)	Advanced control-structures with an emphasis on concurrency and writing concurrent programs at the programming-language level. Parallel/distributed programming models and interfaces - shared memory vs. message passing vs. remote procedure call (RPC) vs. global address space languages: e.g., p- threads, MPI, Open-MP, HPF, UPC, language-level threads (e.g., Java). Parallel machine architectures - shared and distributed memory machines, multi-core and multithreaded chips, interconnection networks Parallel program optimization techniques - synchronization granularity, dependences, scheduling, load balancing Synchronization - hardware primitives, clocks, mutual exclusion, transactions. Consistency and coherence - data-centric versus client-centric consistency models, cache coherence protocols. Distributed file systems - Coda, NFS, XFS. Fault tolerance and reliability - two- and three-phase commits, reliable group communication, check-pointing, message logging	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Computer Vision (Complementary Module)	Images, cameras, image formation , Image statistics, edges, and texture, regularization, diffusion, and Markov Random Fields, Optical flow (image motion): affine flow, regression, dense flow, Stereo, Tracking, Segmentation and grouping, Bayesian inference, Principal component analysis and Eigen-models of objects.	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Advanced Human Computer Interaction(Complementary Module)	Overview of human-computer interaction and role of contributing disciplines including psychology, sociology, etc. Some appropriate examples of detailed knowledge from contributing areas - e.g. 3D vision, user centred design focus and related design issues (e.g. user experience in consumer products). Techniques for managing and designing interactions: e.g. scenarios, task analysis, dialogue modelling, special issues for networked systems: timing issues, distributed interface architectures, impact of QoS on user experience,...etc. Technologies: ubiquitous computing, wearable computing, augmented reality, bio-sensing, visualisation…etc. Small group design project.	A- Computer Science Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Service Oriented Software Engineering (Core module)	Review the history of the evolution of Enterprise Architecture and Enterprise Application Integration (EAI). Technical aspects of SOA. Orchestration of multiple service invocations by the use of WS-BPEL (Web Services Business Process Execution Language). Workflow engines & Integration. Analysis of workflow systems based on BPMN (Business Process Modeling Notation). SOA quality attributes. SOA development phases, techniques, methodologies & governance framework	B- Software Engineering Specialisation:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
British University in Egypt	Social networks (Optional module)	Client-server paradigm versus Peer-to-peer paradigm .Ubiquitous computing & embedded systems. Web-based social networks (Architecture & Technologies).Peer-to-peer social networks (Architecture & Technologies).Network Protocols related to social networking. Requirements specification & design. Quality attributes for social networking	B- Software Engineering Specialisa tion:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Software Engineering for Web application s (Optional module)	Web engineering concept, requirements engineering for Web applications, Web applications architectures, Technology-aware Web Application Design, web Project management, The Web Application Development Process, using patterns for engineering, high quality web applications, Usability of Web Applications, testing methods for web applications, security for web applications	B- Software Engineering Specialisa tion:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Interacting in Virtual Environme nts (Optional module)	Virtual Environments Standards and Terminology , Structured Development of Virtual Environments , Virtual Environment Models , Principles for the Design of Performance - Oriented Interaction Techniques , Design Approaches and Implementations Strategies, The design of Multisensory Virtual Environments, Internet - Based Virtual Environments , Usability Engineering of Virtual Environments , Human Performance Measurement in Virtual Environments	B- Software Engineering Specialisa tion:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Formal Specificatio n (Compleme ntary Module)	ntroduction to Z schemas, Z-Schema cases studies, State-based approaches, Event based approaches, Algebraic specifications, Temporal logic, Language Of Temporal Ordering Specification (LOTOS), Petri nets, B-method, Action systems, Software Specification and Description Language (SDL), Hardware Description Language (VHDL), Properties of programs, Specification, Verification and Validation.	B- Software Engineering Specialisa tion:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Software Developme nt for Real Time System (Compleme ntary Module)	Real - time embedded systems, Hardware fundamentals, Microprocessor case study , Interrupts, Survey of software architectures, Rea l- Time Operating Systems (RTOS) , Basic design using a RTOS, Debugging techniques.	B- Software Engineering Specialisa tion:	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Introductio n to Web Science (Core module	Web history, epistemology and didactics. Research methodology. Research practice in the Web. Conceptual framework for Web Models. Web in society: Collective intelligence, Social Systems theory, Social structures & processes, Groups, identity, globalization, Social capital, power inequality, Virtual communities & politics. Privacy & trust in the Web. Web & the Law. Web & Linguistics. Web, Life Sciences, Evolution. Web and Psychology – Psychiatry. Web Governance	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
British University in Egypt	Information Theory and Processing (Optional module)	Historical background.Quantities of (Definitions of Probabilities and Entropy, Joint entropy, Conditional entropy, Mutual information, Information gain). Data compression I: Symbol codes, Data compression II: coding, Noisy channel coding I, Noisy channel coding II. Error Correcting Codes and Real Channels. Bayesian inference. Applications	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Web Mining and Information Retrieval (Optional module)	Principles of information retrieval. Dictionaries and tolerant retrieval. Index construction. Index compression. Scoring, term weighting, and the vector space model. Computing scores in a complete search system. Relevance feedback. Different classification: - Naive Bayes models. Spam filtering. - K nearest neighbors, decision boundaries, vector space classification using centroids.- Support vector machine classifiers. Kernel function. Evaluation of classification. Micro- and macro-averaging. Learning rankings. Clustering :-. Introduction to the problem. Partitioning methods: k-means clustering; Hierarchical clustering. -Latent semantic indexing (LSI). Applications to clustering and to information retrieval. Web categories :Web 1: Web search overview, web structure, the user, paid placement, search engine optimization/spam. Web size measurement. Web 2: Crawling and web indexes. Near-duplicate detection. Web 3: Link analysis	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Knowledge Processing in the Web (Optional module)	Logic & programming in the Web. Ontology & semantic processing. Semantic Web languages. Linked data. Semantic Web applications: FOAF, SWiki. Security procedures. Content creation and management. Distributed processes: Cloud computing, P2P. Ambient intelligence, Sensor Nets. Future Internet, Internet of Things	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Systems Analysis & Design (Complementary Module)	Conceptual material about both traditional approaches to systems development such as process oriented and data-oriented methodologies and evolving approaches such as object-oriented development methods. Key stages of the systems development life cycle including planning, analysis, and design. Models and procedures for understanding and modeling an organization's existing and planned information systems. Computer-aided software engineering tools are used to provide hands-on experience in designing information systems. A case-based approach is used to provide students an opportunity to apply the analytical and design techniques covered in the course. In addition, students are expected to do a real-life systems development project. Issues and challenges in managing systems development.	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
British University in Egypt	Data Mining & Warehousing (Complementary Module)	Depth understanding of the role of Data Mining within IT dependent enterprises. Students develop capabilities in the analysis, evaluation and implementation of Data Mining principles as well as the ability to employ the technologies for data mining and data warehousing in a variety of appropriate contexts. Web Mining and Web Search and Applied Data Mining Techniques. The course also aims to produce graduates who are equipped to both work in the software industry and pursue research in Data Mining.	C- Information Systems Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Web Topologies (Core module)	Graphs & Combinatorics. Graph Topology. Random Graphs. Small Worlds. Scale-free Graphs. Information, Entropy, Probability Statistics. Information and Networks. Social networks	D- Computer Networks Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Emerging Computer Networks (Optional module)	The course explores emerging trends we well as provides in-depth technical and business perspectives in computer networks area. Topics include backbone networks, last-mile access, wireless networks, internet/IP-based applications, application requirements/quality-of-service (QoS), and advanced security. Students get to apply knowledge from course towards future application scenarios for live businesses, including technology trade-offs and recommended architectures	D- Computer Networks Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Net-centric Computing (Optional module)	The course covers strong emphasis on the services provided by information and communication technologies to distributed applications with focuses on the impact of the QoS paradigm on global networks. focus on the issues related to providing broadband network services of high quality. By the end of the course students should be able to examine current trends in Web application Development, conduct an in depth review of key network security standards, examine key network and data technologies and emerging trends, an in depth review key network technologies, and effectively research current issues in net-centric computing.	D- Computer Networks Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		
British University in Egypt	Network monitoring, visualization, characterization, and analysis (optional module)	The aim of the module is to expose students to the theoretical concepts and standards associated with network management, troubleshooting, and provide knowledge with network management protocols, configuration, and network management tools. Students will also learn network performance and reliability; factors that affect these characteristics, ways in which network monitoring and analysis can improve performance and reliability.	D- Computer Networks Specialisation	Assoc. Prof. Osman Ibrahim, Assoc. Prof. Samy Ghoniemy, Dr. Ahmed Gawish, Dr. Mostafa Salama	http://www.bue.edu.eg/index.php/apply/14-sample-data-articles/180-ics-master-program	Masters		

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Colonge University of Applied Sciences	Introduction to Web Science: Part of the Foundations and Principles I Module	Engineering the Web, Analyzing the Web, Social Aspects of the Web, Web Economy, Design and The Web, Business Administration, Management and The Web	Lectures, Online Sessions. Assessment: a written examination of 45 minutes,	Kristian Fischer, Gerhard Hartmann	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Introduction_to_Web_Science	Masters	Term 1, Winter 2014	
Colonge University of Applied Sciences	Web Architectures: Part of the Foundations and Principles I Module	Evolution of the Web during the past Decades, Fundamental Concepts, Identification, Interaction, Data Formats, The Architects's View of the Web, Web 2.0, Architectural Patterns, Information Architectures, Extensible Markup Language (XML), XML Schema Languages, Java Script Object Notation (JSON), Resource Description Framework, Micro Formats, Synchronous Interaction, Web Services, REST Architectural Style, SOAP Architectural Style, Business Process Modeling, Asynchronous Interaction, Event Driven Architectures, Instant Messaging, (Micro)blogging, Peer-to-peer Interaction, File Sharing: Gnutella, Mass Distribution: Bittorrent, Conferencing: SIP,	Introductory lesson onsite, followed by online sessions	Kristian Fischer	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Web_Architectures	Masters	Term 1, Winter 2014	
Colonge University of Applied Sciences	Quality Management for the Web: Part of the Foundations and Principles I Module	Basic Principles and Tools, Statistics, Measurement and Metrics, Product Quality, Quality in use, External and internal Quality, Information Quality, Quality Management Methods, Techniques, and Tools, Organizational Measures, Preventive (Constructive) Measures, Analytical (Detective) Measures, Process and Organization Quality, Capability Maturity Model Integration (CMMI), ISO 15504 (SPICE), Didactic Concept, Schedule and Assignments	Introductory lesson onsite, followed by online sessions	Mario Winter	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Quality_Management_for_the_Web	Masters	Term 1, Winter 2014	
Colonge University of Applied Sciences	Human-Computer Interaction Basics: Part of the Foundations and Principles I Module	Fundamentals of Cognitive Psychology, Models of Human-Computer Interaction, Conceptualization of a Human-Centered Design Process, Conceptualization of Accessibility, Didactic Concept, Schedule and Assignments	Introductory lesson onsite, followed by online sessions	Gerhard Hartmann	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Human-Computer_Interaction_Basics	Masters	Term 1, Winter 2014	
Colonge University of Applied Sciences	Design Basics: Part of the Foundations and Principles I Module	Basic Laws of Visual Design, The Power of Type and Words, Pictures and Picture Language, Visual Seduction, Analysing Design	Introductory lesson onsite, followed by online sessions	Uwe Stoklossa	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Design_Basics	Masters	Term 1, Winter 2014	

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Colonge University of Applied Sciences	Controlling and Performance Management: Part of the Foundations and Principles II Module	The data landscape, clickstream data and uncertainty, contextual data, user/customer surveys, Metrics and Key Performance Indicators (KPI), Web Analytics Methodologies, Benchmarking and Segmentation	Introductory lesson onsite, followed by online workshops	Gerhard Hartmann	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Controlling_and_Performance_Management	Masters	Term 1 summer 2014	
Colonge University of Applied Sciences	Computer Ethics: Part of the Foundations and Principles II Module	Ethics of technology, Specific features of information, Professional responsibility	Assessment Information: 3 Optional assigments and one written exam (assigments can count for 50% of the exam mark at the student's request)	Wolter Pieters	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Computer_Ethics	Masters	Term 1, workload 300 hours	
Colonge University of Applied Sciences	Strategic Management: Part of the Foundations and Principles II Module	Mission and Objectives, Situation Analysis, Strategy Formulation, Strategy Implementation and Monitoring	Introductory lesson onsite, followed by online sessions and a wrap-up session onsite. Assessment Information: 1 exam (45 mins)	Jan Karpe	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Strategic_Management	Masters	Term 1 : Winter 2014 or summer 2014	
Colonge University of Applied Sciences	Eentrepreneurship: Part of the Foundations and Principles II Module	Web 1.0 versus Web 2.0, Formation of an Enterprise, Formation of an Information Economy Enterprise, Principles of Information and Information Economics, Pricing of Information Goods, Brand Naming, Copyright & Trademark, Profitability, Web Mining, Competitive Intelligence, Web Technologies,	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	William Sen	See referemces section of: https://webscience.fh-koeln.de/smwiki/index.php/Eentrepreneurship	Masters	Term 1: Winter 2014, or Summer 2014	
Colonge University of Applied Sciences	Web Security: Part of the 'Web Trust and Security' Module	Introduction: Basics of IT-Security and Web Security, Web Security and Web Trust, Perception of Security, Correlation between Web Security an Web Trust, Attacks on Web Applications and Design of Secure Web Architectures	The course concept contains online workshops, online discussions, milestone meetings and audits. In addition there is an introductory and final on site presence.	Stefan Karsch	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Web_Security	Masters	Term 2: Winter 2014	

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Colonge University of Applied Sciences	Web Trust: : Part of the 'Web Trust and Security' Module	General conditions of Web Trust, Threats in trust of web resources	The course concept contains online workshops, online discussions, milestone meetings and audits. In addition there is an introductory and final on site presence.	Hans Ludwig Stahl	See references secition: https://webscience.fh-koeln.de/smwiki/index.php/Web_Trust	Masters	Term 2: Winter 2014	
Colonge University of Applied Sciences	Organizational Behavior: Part of the ' Decision and Management' Module	Transaction Cost Theory as a Basis of Modern Organization Theory, Modern (hybrid) Forms of Organization Between Market & Company, Make or Buy, The Role of Transaction Costs, Virtual Organizations & Strategic Networks, Principal Agent Approach as a Modern Behavior Theory in Organizations, Organizational Behavior	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Jan Karpe	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Organizational_Behavior	Masters	Term 2: Summer 2014	
Colonge University of Applied Sciences	Risk Management: Part of the ' Decision and Management' Module	Introduction: Basics of IT-Security- and IT-Risk-Management, Security Policy, Concrete ISMS and Practical Methodology	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Stefan Karsch	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Risk_Management	Masters	Term 2: Summer 2014	
Colonge University of Applied Sciences	Media Psychology: Part of the 'Web and Society' Module	Introduction To Media Psychology, Media and Media Psychology, Psychological Foundations, Stereotypes, Use and effects of media in Advertisement, Qualitative and Quantitative Research Methods in Media Psychology	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Christiane Grünloh	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Media_Psychology	Masters	Term 3: Summer 2014	
Colonge University of Applied Sciences	Privacy: Part of the 'Web and Society' Module	Fundamentals – Origin and notions of Privacy, Everyday Life, Everyday Privacy, Privacy and the Web, Post-Privacy	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Tim Schneider	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Privacy	Masters	Term 3: Summer 2014	
Colonge University of Applied Sciences	Risks and Opportunities of Social Media Data: Part of the 'Web and Society' Module	Motivations and data characteristics, Data Mining and Profiling, Decentralisation of social media services, Internet infrastructure and tracking	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Benjamin Krumnow	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Risks_and_Opportunities_of_Social_Media_Data	Masters	Term 3: Summer 2014	

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Colonge University of Applied Sciences	Cooperation Systems: Part of the ' Web and Cooperation' Module	Cooperation Systems in Socio-Technical Systems, Social Concepts for Cooperation, Architectures, Systems and Platforms, Future Prospects, Resources, Journals, Conferences, Collections	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Kristian Fischer	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Cooperation_Systems	Masters	Term 3: Winter 2014	
Colonge University of Applied Sciences	Intercultural Teams: Part of the ' Web and Cooperation' Module	Basic studies and concepts of intercultural management, Elements of intercultural communication, Examples of intercultural communication differences, Recommendations for managing intercultural teams, Guidelines for communication in intercultural teams	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Siegfried Stumpf	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Intercultural_Teams	Masters	Term 2, Winter 2014	
Colonge University of Applied Sciences	Interaction Design: Part of the 'Design' module	Introduction To Interaction Design, Cognitive Psychology, Conceptual Design, Creativity Techniques, Design Principles, Design Qualities AND Design Rationale, Interaction Paradigms, -Styles and -Modes, Metaphor Engineering For The Web, Sketches, Scribbles, Mockups, Prototypes, Storyboards, Wireframes, Evaluation Methods And Techniques,	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Gerhard Hartmann	See references section: https://webscience.fh-koeln.de/smwiki/index.php/Interaction_Design	Masters	Term 3, Winter 2014	
Colonge University of Applied Sciences	Designing for the web: Part of the 'Design' module	Dive into HTML, Typography on the Web, HTML5, Grid Systems and Responsive Design	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Silvio Barta	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Designing_for_the_web	Masters	Term 3, Winter 2014	
Colonge University of Applied Sciences	Corporate Design: Part of the 'Design' module	Corporate Design Basics, Strategy, Values, Naming, Brandmarks, Look and Feel, Media, Brand Management	Introductory lesson onsite, followed by online sessions, Assessment Info: examination	Uwe Stoklossa	See refrences section: https://webscience.fh-koeln.de/smwiki/index.php/Corporate_Design	Masters	Term 3, Winter 2014 or Summer 2014	
Goldsmiths London	Sensory Sociology: Imagining Digital Social Research	This module investigates the transformation of sociology in the age of digital technologies, and offers an advanced introduction to medium-specific social research. The module introduces students to sociological theories relevant to researching digital social life, and examines problems and opportunities that digital technologies open up for empirical social research. The module places special emphasis on the opportunities that digitisation offers for the re-invention of sociology: its overarching aim is to enable examination of, and engagement with, empirical sociology as an 'inventive' research practice, oriented towards the creative deployment of research devices and methods. The module combines advanced training in sociological thinking with practice-based forms of methods teaching, taking an explorative approach to renewing the empirical commitments of sociological research in a digital context.	A report of collaborative exercise(s) (text/visual/audio/online) (30%) A 3,500 word essay (70%).		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 1	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Goldsmiths London	Digital Sandbox Sociology Seminar		1 x 3-5,000-word essay		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 1	English
Goldsmiths London	Digital Sandbox		(3 x technical mini-projects each term)		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 1 to 2	English
Goldsmiths London	Digital Sociology in Practice		1 x 3,500 word essay		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 2	English
Goldsmiths London	Digital Social Research Methods	This module introduces students to social research methods that have special salience in the digital context: ethnography, network analysis and online textual analysis and issue mapping. The module offers an advance introduction to computer-enabled sociological methods and then proceeds to examine specific methods and their digitisation on a case-to-case basis. The module provides an overview of the central principles of these sociological methods, and then offers a hands-on introduction to correlated online research tools and platforms. The module provides experience of a range of current searching and database technologies, and techniques and commands for the analysis of online social content. Finally, the module explores the sociological implications of the changing status of social research methods in the digital environment, as methods are materialised in search engines, data visualisation tools, and so on.	1 x 2,500-3,000 word essay		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 3	English
Goldsmiths London	Digital Research Methods		3 x assignments		http://www.gold.ac.uk/pg/ma-msc-digital-sociology/	Masters	Term 3	English
Korea Advanced Institute of Science and Technology (KAIST)	Introduction to Web Science and Technology	Web is a platform for online communication, trust, identity, and collective intelligence and also is a medium for social and political opinions, decision-making processes, and online reputation formulation. Web science and technology is an interdiscplinary program that combines law, social sciences, natural sciences and engineering and covers the role and functions of the web. This course offers an introduction to Web Science and Technology.	Lecture/Lab	Sue-B. Moon, Chin-Wan Chung	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Korea Advanced Institute of Science and Technology (KAIST)	Fundamentals of Searching Web-scale Datasets	In this course, we will study index structures and algorithms that are commonly applied in practice to support queries on large datasets. The course will be divided into three parts, which focus on data streams, multi-dimensional objects, and web-specific applications, respectively. Topics to be covered include sampling, hashing, sketch structures (e.g., count-min sketch, bloom filter, etc.), R-trees, nearest neighbor search, instance optimality, and so on. We will also discuss a series of foundamental techniques for analyzing the performance of algorithms.	Lecture/Lab	Yufei Tao	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Web Architecture	This course is on basic concept of web architecture and building a web service. It deals with web architecture design principles and its performance and security issues..	Lecture/Lab	Sue-B. Moon	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Software Engineering for the Web World	As the role of and functionality provided by the World Wide Web have evolved, the role of software engineering in creating Web-based systems, and the ability of software engineers to harness Web-based resources to perform engineering tasks, have also evolved. This has implications for the practice of software engineering by professional software engineers, and in this course we will study these implications. A further consequence of this evolution, however, is the rapid growth of web-based platforms supporting end-user programming. This in turn motivates concerns about "end-user software engineering", and in this course we will study this phenomenon as well.	Lecture/Lab	Gregg Rothermel	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Web Search and Text Analysis	This course will cover the modeling and algorithmic aspects of web search engines. Topics include tf-idf scoring, page rank, inverted index, trie, suffix tree, string B-tree, q-gram, error-tolerant keyword search, deep web, and so on. On completion, students are expected to have acquired solid understanding on fundamental text retrieval and link analysis.	Lecture/Lab	Yufei Tao	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Ontology Engineering	The objective of this course is to provide students with basic understanding of ontology engineering. This course consists of concepts and examples of ontology, construction of ontology, and how ontology used in philosophy, linguistic, artificial intelligence, and computer science.	Lecture/Lab	Tony Veale, Key-Sun Choi	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?
Korea Advanced Institute of Science and	Mobile Web Applications	This course deals with difference between the existing and mobile web environment and major issues due to the difference. Furthermore, applications of mobile web and their utilization are covered in this course.	Lecture/Lab	Sue-B. Moon	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Technology (KAIST)								
Korea Advanced Institute of Science and Technology (KAIST)	Research Methods for Information Behaviors on the Web	Web is considered as a rich repository for humans to gather, store, and process information. In order to develop efficient and effective web technologies and/or serivces for information behaviors, it is so necessary to understand what people need and what they desire for their lives. In turn, appropriate understanding of web users regarding what information they want to see on the web and how much they are satisfied with the present web technologies and/or services is the stepping stone to improve the existing web conditions and to compete for the pre-occupancy of the edges of future web technologies and/or services. To achieve such purposes, this course, from the behavioral perspectives, will inform class participants that how we can examine what web users want and hope, and that how we can evaluate the effectiveness of the potential technologies and services.	Lecture/Lab	Young-Min Baek	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/ PhD	Spring	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Web Audience Research	The rise of the web transforms our media landscape. Conventional media continuously lose their own audience who increasingly moves to the newer flatform, such as the web and/or mobile flatforms. Especially, in the era of Web 2.0, the audience's active participation in the production, remix, distribution, and exhibition of media content is getting more attention. This course convey knowledge in audience research in conventional media, examines the limitation of old audience research, and suggests novel ways how to measure the size of audience and how to evaluate the benefits or concerns - cognitive, social, as well as industrial - observed in the era of new media. With convergence of conventional media and new media, media competes and/or co-exists with the exploded cultural contents produced by both professionals and amateurs. Media consumption and people's reactions in non-conventional media should be developed, which serves the main challenging topic in the course.	Lecture/Lab	Young-Min Baek	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/ PhD	Fall	Korean/English?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Korea Advanced Institute of Science and Technology (KAIST)	Web Software Dependability	In this course we will investigate recent research on techniques for assessing and improving the dependability of various classes of Web software. We will begin by studying basic techniques for analyzing such software and for verifying its dependability; these include testing, static analysis, and dynamic analysis techniques. Of particular interest are Web applications built by non-professional (end-user) programmers, such as mashups and Web macros, and we will consider current research aimed at helping these users build more dependable programs. We will also spend time studying empirical methodologies for assessing techniques, including, in particular, empirical studies of programmers. There will be no exams in the course; instead, students will be asked to complete a course project which may involve research on new techniques, implementation of techniques, or empirical study of issues related to dependability. Prerequisite course-WST520	Lecture/Lab	Gregg Rothermel	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	External Memory Data Structures	The external memory (EM) model was introduced by Alok Aggarwal and Jeffrey S. Vitter in 1988, observing that I/O time is often the performance bottleneck in algorithms on massive datasets that do not fit in memory. This fact has become increasingly evident as the data volume continues to outgrow a computer's memory capacity, rendering the EM model the predominant computation model in studying I/O-efficient algorithms. In this model, a computer has an internal memory of M words and a disk of unbounded size. An I/O operation transfers a block of B consecutive words between the memory and disk. Space complexity is measured in number of disk blocks occupied, whereas time complexity is in number of I/Os performed. CPU calculation is free. In this course, we will discuss EM structures with good tradeoff between (worst-case) space and query time. A well-known example is the B-tree, which indexes a set of N real numbers in O(N/B) space, and supports retrieval of any number in O(logBN) I/Os. We will cover other structures settling problems such as stabbing-query on 1-d intervals, 2-d orthogonal range reporting, and 2-d orthogonal range count/max, all of which are rudimentary in a large number of applications. We will study numerous fundamental techniques for developing I/O-efficient structures, such as persistency, boot- strapping, fractional cascading, compression, weight-balancing, logarithmic method, etc.	Lecture/Lab	Yufei Tao	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Korea Advanced Institute of Science and	Multimedia Contents Protection	In this course, the technology related with the multimedia content protection is studied. Particularly, it covers the multimedia protection issues for the personal private information security on the web.	Lecture/Lab	Heung-Kyu Lee	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Technology (KAIST)								
Korea Advanced Institute of Science and Technology (KAIST)	High Performance Computing	This course will focus on high-end architectures and computing including recent supercomputers, and architectures used in modern datacenters. High-performance will be focus but with the increasing importance of energy, we will also investigate energy-efficient computing as well. We will also discuss recent trend in computing, which included heterogeneous computing and accelerator architecture.	Lecture/Lab	John Kim	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Web Search Technology	This course covers issues related to retrieval and extraction of information from a variety of information types on the Web. Topics to be covered are: Web crawling, text analysis and indexing, link analysis and social search, Web user query log analysis, distributed processing of large amount of data, automatic classification, and retrieval models. As the Web evolves, new contemporary topics will be also covered.	Lecture/Lab	Chin-Wan Chung,Sung-Hyon Myaeng	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Linguistic Web Science	This course will examine the linguistic make-up of the Web, and ask: what should a language-oriented Web Science look like? The Web has evolved into the largest corpus of language use in existence, and continues to grow on a daily basis. We will look at how computational linguists use the Web as a corpus, to extract common-sense or definitional knowledge from its texts, and to test linguistic hypotheses on a large scale. Because the Web captures the bleeding-edge of language change, we will look at how the essential creativity of language can be analyzed and modeled using Web content. More specifically, we will look at two different ways of using Web content: static databases of Web n-grams, and dynamic Web queries. A case-study will look at language use on the Web in a particular domain, such as online newspapers.	Lecture/Lab	Tony Veale	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Advances in Mobile and Wireless Networks	The popularity and the variety of usages of mobile systems continue to increase. In this course the students will learn the fundamentals as well as the practical advances in the areas of mobile communication systems, especially the advances in smartphones and mobile social networks. Various critical aspects of such systems including performance, reliability, security, and privacy issues will be addressed through discussions of recent technical publications and a comprehensive collaborative course project.	Lecture/Lab	Prasant Mohapatra	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Korea Advanced Institute of Science and Technology (KAIST)	Web-scale Image and Video Retrieval	In this class we will discuss various techniques related to image/video retrieval. Especially, we will go over image/video features (e.g., SIFTs and GISTs), their indexing data structures, and runtime query algorithms. We will also study scalable techniques that can handle web-scale image/video databases..	Lecture/Lab	Sung-Eui Yoon	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Online Social Network	Online social networks are the new leading technology in the IT industry after Microsoft, Google, and Apple. While existing online social networks offered a single feature of social networking, Facebook and Twitter are changing the definition of OSNs by presenting themselves as a platform for new services. The emergence of OSNs as a service platform poses new challenges in interdisciplinary research with cognitive science, sociology, polticial science, and communications. In this course we cover the basic features of online social networks, analytical basis for interdisciplinary research, and social network data analysis.	Lecture/Lab	Mee-Young Cha, Won-jae Lee	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Fall	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Human-Centered Data Analysis	This course will focus on analyzing user behavior through large-scale data analysis. This includes discovering and leveraging search and navigation patterns, understanding how elements of interaction impact behavior, and how we can use controlled experiments in combination with user studies and other techniques to gain insights into human behavior with a particular emphasis on multimedia, particularly in the context of social media.	Lecture/Lab	Alejandro Jaimes	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring	Korean/English?
Korea Advanced Institute of Science and Technology (KAIST)	Special Topics in Web Science and Technology	This course is offered to meet the ad hoc demand of students in special areas of Web Science and Technology which is not covered by regular courses.	Lecture/Lab	Staff	http://eng.webst.kaist.ac.kr/content.php?db=_subject_page#C1	Masters/PhD	Spring/Fall	Korean/English?
Oxford Internet Institute	Social Dynamics of the Internet	Theories of Society and the Internet, Globalization and Domestication, Social Networking Sites, Mobile Phones, the Internet, and Perpetual Contact, Search and Access to Knowledge and Information, Microblogging among New and Old Media, Media Systems in India and China, The Internet and Democracy,	The course will be taught during Michaelmas term in eight weekly classes, each consisting of a lecture followed by a one hour seminar with student presentations and discussion. There will also be smaller bi-weekly seminars of one hour which will be led by a teaching	Professor Ralph Schroeder	Social Dynamics of the Internet	Compulsory for OII MSc and DPhil students	Michaelmas Term (Weeks 1-8). Tuesdays 9:00-10:00.	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
			assistant, each student will be required to give a short presentation on a specific aspect of the session topic. Assessment: 3 Hour exam					
Oxford Internet Institute	Internet Technologies and Regulation	Internet architecture, History and development of the Internet, Internet governance and regulation, Regulatory Responses to Public Debates on Emerging ICTs, Digital copyright, Digital identity and authentication, Content regulation and filtering, Privacy and security,	The course is taught in eight weekly classes, each consisting of a lecture followed by student presentations and seminar discussion. All classes will begin with a lecture at 10:00 in LR23 in Balliol College.	Professor Ian Brown, Dr Joss Wright	Internet Technologies and Regulation	OII MSc students, optional for OII DPhil students	Michaelmas Term (Weeks 1-8). Wednesdays 10:00-13:00.	English
Oxford Internet Institute	Digital Social Research	Accessing Research Data from the Social Sciences, Advanced Multiple Regression for Internet Research, Experimental Approaches, Analysing Data, Digital Ethnography, Digital Interviewing, Logistic Regression and Principal Components Analysis for Internet Research, Simulating Society, CAQDAS and the Internet, Metrics, Measurement, and Online Visibility, Information Visualisation, Big Data Analytics,	There are six components to the Digital Social Research element of the degree. The combined total of these six components accounts for 25% of the final mark for the year. All students must take Digital Social Research: Methods Core and Digital Social Research: Statistics Core in Michaelmas term.	Dr Jonathan Bright, Dr Grant Blank, Dr Bernie Hogan, Eric T. Meyer, Rebecca Eynon, Dr Andrew Przybylski, Professor Helen Margetts, Dr Stefano De Sabbata, Mike Thelwall, Professor, Dr Taha Yasseri,	See PDF links: http://www.oii.ox.ac.uk/graduatestudy/msc/courses.cfm?id=20		At the beginning of the academic year, students will select four 4-week methods option courses taught in Hilary term, each focusing on advanced methods topics. There are currently twelve methods option courses divided into two groups of six - Group A and Group B. Student will choose two methods options from each group.	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Rensselaer Polytechnic Institute RPI	Web Science 2013	This course attempts to provide the foundations of that understanding, exploring the fundamentals of the World Wide Web's function including the HTTP protocol, key algorithms that make the Web function, future trends, and social issues with respect to Web use and effect.	Taught classes, homework assignments, project assignments	Prof Jim Hendler	**Class Syllabus (pdf) [Download],** Web Science Article from Science Magazine (pdf) [Download], Tim Berners-Lee: Web Roadmap 2008 (jpg) [Download], How Google Works (jpg) [Download], Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine", Kristine's Research Methods Slideshow [Download] (From class Mon 2/11), Prof. Hendler's Search Engine slides [Download] (From class Thurs 2/21) , Prof. Hendler's Some Web 2.0 background slides [Download] [Download] , Prof H's Final Class Lecture [Download]		January 22, 2013 - May 17, 2013, Monday and Thursday, 4pm - 5:5opm	English
Rensselaer Polytechnic Institute RPI	Advanced Web Science 2014	Semantic Web, Semantic Foundations, Social Web, Web Science. This course attempts to provide the foundations of that understanding, exploring the fundamentals of the World Wide Web's function including the HTTP protocol, key algorithms that make the Web function, future trends, and social issues with respect to Web use and effect.	Taught classes, homework	Prof Jim Hendler	http://tw.rpi.edu/web/taxonomy/term/128		Spring 2014	English
RWTH Aachen University	Web Science	More than twenty years after the birth of the World Wide Web, Web Science has been becoming a new study field in Computer Science. This course introduces fundamental concepts (web centralities & algorithms, network models and web engineering principles) of Web Science. We then give an overview on regular and random network models. We will learn fundamental algorithms for web page ranking like PageRank and HITS as well as advanced community detection algorithms. The anatomy of recommender systems and dynamic processes on complex networks will finish this part. In the engineering part we dig into emerging cloud computing approaches and Post-HTTP protocols like the XMPP and WebRTC. We will learn about Web Services and their RESTful implementation. With the knowledge gained in the preceding chapter we can analyze and engineer advanced Web applications like video annotation environments, personal learning environments and storytelling environments.	Weekly lecture, WS 15/16. Assessment: Exam on 05. Feb 2016	Prof. Dr. Matthias Jarke, PD Dr. Ralf Klamma, Mohsen Shahriari	http://dbis.rwth-aachen.de/cms/teaching/WS1516/web-science	Unknown	Weekly lectures starting on: 19. October 2015, Monday, 10:15 - 11:45	English/German?

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Saint-Joseph university of Beirut	Semerster 1	Digital Economy, Public Economy, Companies Assesment, Monetary Economy, International Finance, Financial Planning, Modelization 1	Seminars, classes	Loubna Oueidat, Stéphane Bazan, Chantal Chelala, Irma Majdalani, Pierre Abi Nader, Joseph Gemayel, Suzy Semerdjian, Marwan Mikhael, Najwa Yaacoub	http://fse-wsen.blogs.usj.edu.lb/files/2013/09/programmeWSEN-EN.pdf	Masters		French
Saint-Joseph university of Beirut	Semerster 2	Economy of innovation, Advanced Digital Economy, IT Project Management, Information Systems, Network Economy, Modelization 2 , Social Media, Mobile technologies, Webeconomics & Business	Seminars, classes	Loubna Oueidat, Chantal Chelala, Nassim Mouchantaf, Antoine Yazigi, Lina Koleilat, Najwa Yakoub, Marilyn Zakhour, Patrick Chemaly, Michalis Vafopoulos	http://fse-wsen.blogs.usj.edu.lb/files/2013/09/programmeWSEN-EN.pdf	Masters		French
Saint-Joseph university of Beirut	Semerster 3	Web Business Models, IT and economicmutations, Web project design, Digital Law, E-Banking Strategy, Web Science Epistemology & Research, E-Government,	Seminars, classes	Chantal Chelala, Loubna Oueidat, Stéphane Bazan, Elias Chédid	http://fse-wsen.blogs.usj.edu.lb/files/2013/09/programmeWSEN-EN.pdf	Masters		French
Saint-Joseph university of Beirut	Semerster 4	E-CRM & ERP	Seminars	Chantal Chelala/ Loubna Oueidat	http://fse-wsen.blogs.usj.edu.lb/files/2013/09/programmeWSEN-EN.pdf			
The University of Edinburgh: School of Social and Political Science	Science, Knowledge and Expertise	What is science?, Politics of knowledge and expertise, Science as practice and the places of science, Communicating science, Science in public, Gender and situated knowledges, Science in everyday life, Science and/in policy, Science in a global context, The future and relevance of science, technology and innovation studies	Lectures	Dr Steve Sturdy	KEY TEXTS LISTED AT: http://www.sps.ed.ac.uk/gradschool/prospective/taught_masters/course_Information/taught_courses_a-z/science,_knowledge_and_expertise	Masters	Lecture, 16th Sep 2014 1:30pm-4:30pm. Every week on Tuesday, until 25th Nov 2014.	English
The University of Edinburgh: School of Social and Political Science	Understanding Technology	What is technology? The relationship between technology and society, The politics of technology: foundational debates, The politics of technological knowledge: how do we know the properties of technology, Practices of technology production, Technological Systems and Entrenchment, Technology embedding & sociotechnical transitions: purposive efforts to embed novel or sustain entrenched technologies, Technology in everyday life, Technology exclusion and inclusion : Lessons from feminist technology studies, New challenges for technology policy and governance, The future and relevance of STS	Lectures. Assessment information: 1 book review 1000 words 25%, 2 paper 3500 words 75%	Prof Robin Williams	Key Texts Listed at: http://www.sps.ed.ac.uk/gradschool/prospective/taught_masters/course_Information/taught_courses_a-z/understanding_technology	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
The University of Edinburgh: School of Social and Political Science	Innovation Systems: Theory and Practice	Discovery, invention, innovation, research, development and commercialisation, Innovation as a systems phenomenon Â¿ general theories and models, Geography of production and geography of innovation Â¿ nations and regions, Entrepreneurship and knowledge spillovers, The role of institutions and organisation and their changing dynamics	Lectures.	Dr Alessandro Rosiello	Key Texts listed at: http://www.sps.ed.ac.uk/gradschool/prospective/taught_masters/course_Information/taught_courses_a-z/innovation_systems_theory_and_practice_i	Masters	Lecture, 19th Sep 2014 9:00am-10:50am. Every 2 weeks on Friday, until 26th Nov 2014.	English
The University of Edinburgh: School of Social and Political Science	Introduction to Risk, Regulation and Governance	Week 1 Introduction to the Concepts of Risk, Risk-Assessment and Uncertainty, Week 2 Understanding Different Models of Risk-Governance and Regulation, Week 3 The Roles of Evidence and Expertise in Managing the Science-Policy,Interface	Seminars	Dr James Mittra	Key Texts Listed at: http://www.sps.ed.ac.uk/gradschool/prospective/taught_masters/course_Information/taught_courses_a-z/risk,_regulation_and_governance_i	Masters	Seminar, 26th Sep 2014 9:00am-10:50am. Every 2 weeks on Friday, until 21st Nov 2014.	English
University College London	Information Retrieval & Data Mining	Indexing, Retrieval Methods, Measurements, Data Mining, Emerging Areas	Method of Instruction: Lecture presentations, Practical exercises. Assessment: Written Examination (2.5 hours, 60%), Coursework (40%)	Jun Wang, Emine Yilmaz	Introduction to Information Retrieval, Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Cambridge University Press. 2008., Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach and Vipin Kumar, Addison-Wesley, 2006, Gigabytes (2nd Ed.) Ian H. Witten, Alistair Moffat and Timothy C. Bell. (1999), Morgan Kaufmann, San Francisco, California., Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer (2006).	Masters	Term 2	English
UAH MediaLab, University of Alcalá (Spain)	Web technology, standards and applications	The World Wide Web was conceived only around 1990, but it has now a quite exciting history of development and innovation. Knowing the underlying base protocols that are the fabric of the Web is essential to understand how it evolves and how we can get Web data using different methods. In this course, we approach the history of the Web and its main developments, also aiming at providing a solid understanding of the insitutions that drive the evolution of the Web as the W3C consortium. This is complemented with providing an understanding of the protocols used in the Web and the main markup languages (HTML, CSS and others) and other languages used to add interactivity to Web pages as Javascript.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
UAH MediaLab, University of Alcalá (Spain)	Social networks and analysis of on-line communities	On-line social networks and social media platforms as Facebook, Twitter or LinkedIn nowadays provide an opportunity to better understand the behaviour and motives of Web users. In addition to these platforms, most sites today integrate some form of social data, be it comments, likes or other forms of endorsement or assessment for people or contents.The techniques, methods and theories of Social Network Analysis (SNA) that developed in the field of sociology provide the Web scientist with the workbench for inquiry and research in on-line behaviour, informed by the specifics of our knowledge on Web users.This course aims at providing an introduction to SNA and skills to use social network analysis tools for the analysis of Web data. Further, the course elaborates on the interpretation of that data in the light of our knowledge about on-line communities and on-line interaction behaviour.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Web infrastructures and data mining	What is the shape of the Web? How do search engines exploit content and links for ranking? How can we get analyze Web data? These are critical questions to any data science approach to the Web.In this course, main topics about understanding, getting and analyzing data are presented through examples. Students are expected to gain insights on the Web and its interconnected infrastructure and acquire some basic skills to analyzing Web data.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	The Semantic Web and the Web of Data	The idea of the Semantic Web was originally that of the application of Artificial Intelligence methods to a richer user experience in the Web. This was based on the key technology of shared, community-curated conceptializations called ontologies and the use of a standardized Web Ontology Language.The developments in the field of the Semantic Web have resulted in a way of exposing any kind of data openly in the Web using RDF and reusing Linked Open Vocabularies and ontologies. This has resulted in a Web of Data available for reuse for analysis and applications.In this course, students will understand ontologies as the fabric of the Semantic Web and will gain basic skillls in ontology editing. They will also be exposed to RDF and the SPARQL query language from a practical perspective.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
UAH MediaLab, University of Alcalá (Spain)	Digital culture	We approach digital culture through its multiple manifestations. We will go deep into the tools that help transform people's everyday practices. Some particularly relevant points in its history will be explored, especially popular culture generated by new media. These are some of the topics that will be addressed: Rethinking new media and digital culture, Cultural hierarchies, Is there anything new digital culture?, Digital culture as institutionalized practice, Digital media in their historical contexts, Digital media and collective intelligence, The language of new media, What is new media?, Interface and operations	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Image, media and virtual worlds	People living in the 21st Century come in contact with hundreds of images every day presented in multiple formats. Through digital architecture, the old dream of transforming the world acquires new possibilities. Images appear and disappear interactively or autonomously embracing multiple forms of audiovisual representation. Images travel through space and time in 2 or 3 dimensions and change the universe in our everyday life. Approaching virtual images today requires an interdisciplinary approach. In this course, we aim to know and understand the fundamentals of communication through text, static or moving images and sound.The images phenomenon in the 21st Century, New digital formats, Displaying change through computer images, How images think, Simulation, viewing and immersion, Interactive worlds: Video games, New tools for creating and analyzing images, Gaze's practices, Media as art and society	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Digital representations, games and mixed reality	The course focuses on two fundamental concepts requiring an interdisciplinary approach: "MIXED REALITY" and "PERFORMANCE". By mixed reality, we understand a combination of real and virtual worlds in real time. In 1994, Paul Milgram and Fumio Kishino approached the concept and understood virtual reality as a continuum, at the heart of what is currently known as augmented reality, which allows for the interaction of the physical and virtual world mediated by the action of the user. In its traditional context, performance can be understood as an act accompanying script and staging, performed by actors in an environment of mixed reality. In this case, the viewers become active participants. Global vs virtual / Local vs real: Global hybrid spaces, Temporal structure and "mixed reality", Cultural events through interactive interfaces, Experience with "mixed reality" viewers, actors and orchestra	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
UAH MediaLab, University of Alcalá (Spain)	Participatory culture, communication media and social practice	When art and media reach people in a digital universe, they can become a powerful tool for personal and social transformation. These tools reshape new environments, both virtual and physical, transforming culture and society. New transmedia spaces (dynamic, fluid and open) are born. Media ecologies are configured in networks, through interactive and mobile platforms, games and other digital experiments. Considering this context, the course explores new and emerging models of cultural practice involving social change and public participation. By examining debates and case studies, we will explore how closely new narrative, interactive relationships between the creative industries, communities and organizations interact with each other. Main topics are as follows: The Roots of participatory culture, Participation and power, The place of fans, Transmedia phenomena	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Design of mobile experiences	Tools allowing "mobile communication experiences" transform the way people relate to each other and, no doubt, their routines when accessing information. We'll see these new tools that enable the mobility of information and how they transform social practices. Mobile communications and new media ecologies, The design of user-centered mobile applications, Usability and user-centered development.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Web economics		Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
UAH MediaLab, University of Alcalá (Spain)	Innovation, creativity and learning	This course aims to understand innovation in close relationship with the learning processes occurring through life. Some of the topics to be addressed are the following: Innovation – what it is and why it is important, Development of innovation strategies, Sources of innovation, Innovation networks, Decision-making and uncertainty, Creation of new products and services, Exploitation of new businesses, The benefits of innovation, Learning and innovation	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Entrepreneurship and regulation in the Web		Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				
UAH MediaLab, University of Alcalá (Spain)	Fundamental rights and equality	Fundamental rights are understood as human strong aspirations enshrined in the laws. These rights are linked to the achievement of human dignity so that without their recognition and guarantee the individual would not be autonomous to carry out life plans, what is especially noteworthy for those subjects belonging to vulnerable groups.Thus, progress is needed in order to achieve freedom and a real and effective equality, for it is essential that legal and political measures of equal opportunities between members of a democratic society are implemented.	Unkown	Alcalde, Rafael, Arroyo, Sinuhé, Colomo-Palacios, Ricardo, Cortés, Sara, Fuente-Prieto, Julián, García-Pernía, María-Ruth, Lacasa, Pilar, Martínez-Borda, Rut, Monjelat, Natalia, Sánchez-Alonso, Salvador, Sicilia, Miguel-Angel				

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University College London	Distributed Systems and Security	Course introduction; OS concepts, Design: Worse is Better; Concurrent IO; RPC & Transparency, Ivy: Distributed Shared Memory, Bayou: Weak Connectivity and Update Conflicts; GFS: The Google File System, RouteBricks: Cluster-Based IP Router; Introduction to Security; User Authentification, Cryptographic Primitives I; Cryptographic Primitives II;, Secure Sockets Layer (SSL); Reasoning Formally about Authentification : TAOS, Software Vulnerabilities and Expoits; Preventing Exploits, Containing Buggy Code: Software-based Fault Isolation; OKWS: Approximating Least Privilege in a Real-World Web Server,	Method of Instruction: Lectures, case-studies, Assessment: Written Examination (2.5 hours, 70%), Coursework Section (30%)	Prof Brad Karp	http://www0.cs.ucl.ac.uk/staff/b.karp/	Masters	Term 1	English
University College London	Multimedia Systems	Introduction and overview, Audio Coding, Video, System Streams, Signalling, OS Issues, Describing Network Traffic, Congestion control and Resource Management, Enhanced Quality of Service, IP Multicast, Digital rights management	Method of Instruction: Lecture presentations, Assessment: Written Examination (2.5 hours, 85%), Coursework Section (1 piece, 15%)	Mark Handley	See page for list of taught subjects: http://www.cs.ucl.ac.uk/students/syllabus/mscncs/gz05_multimedia_systems/	Masters	Term 2	English
University College London	Complex Networks and Web	Network science: Complex networks, Network graphic metrics, Random networks, Small-world networks, Scale-free networks, Network mathematical models, Network structural constraints, Network centrality measures, Temporal networks, Spatial networks, Network visualisation/ Communication and information networks: Internet core structure – evolution and modelling, Structure of the Web – PageRank and document networks, Online social media networks - Twitter, Facebook, Amazon,/Network functions and behaviours: “Rich gets richer” phenomenon, Link, neighbourhood and community, Cascades and epidemics, Network structure balance, Sentimental, temporal and spatial analysis of social media networks	Lectures. A Moodle webpage is created for the course. All course materials, such as lecture notes and online resources will be shared. By using the Moodle, students will also be able to discuss ideas and questions with the lecturer and other students., In the second half of the term, there will be a weekly one-hour lab/tutorial session, where the lecturer and/or a teaching assistant will discuss questions with students. Assessment, Unseen 2.5 hour written examination (70%), Coursework I (15%): essay writing (2000-3000 words);	Shi Zhou	D. Easley and J. Kleinberg. Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press, 2010., M. E. J. Newman. Networks: An Introduction, Oxford University Press, 2010., S. N. Dorogovtsev. Lectures on Complex Networks, Oxford University Press, 2010., Other books for interest: D. J. Watts. Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University Press, 1999, Dodge and R. Kitchin. Atlas of Cyberspace, Pearson Education, 2001., S. N. Dorogovtsev and J. F. F. Mendes. Evolution of Networks: From Biological Nets to the Internet and WWW, Oxford University Press, 2003., M. Mitchell. Complexity: A Guided Tour, Oxford University Press, 2009.	Masters	Term 1	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
			due in the last week of Term-1., Coursework II (15%): individual project on network data analysis (programming is usually required); due in the first week of Term-2., To pass the module students must achieve a pass mark of 50% when all elements are combined.					
University College London	Web Economics	System design: Web basics: HTTP, HTML5 referrer, Link and Click-through analysis, etc, Basic Economic Principles and Economic analysis: Micro vs. Macro economics, Basic elements of Supply and Demand, Equilibrium, Incentives: Game theory, and Auction theory, Business Models in the Internet: auction and bidding (the Ebay Model, swoopo, and b2c and b2b auctions (alibaba), Subscription (Compulsory license, dropbox premier model, spotify, apple icloud, pay per use). Online retailing (Amazon, Apple Apps)., digital goods & bundling, Computational advertising, Vickrey auction and the second price auction, Search-based advertising, Contextual advertising and Behaviour targeting, Demand-side platform and Real-time bidding, Ad exchange and futures and options, Digital Right Management, Spam/fraud control and Internet radio, Computing as a service/utility, Social media mining, Management and optimization, Dynamical pricing models (air-tickets) and Yield management and scheduling (online advertising), Search engine optimization, People : Attention economics and Personalization and Long tail, Prediction market and its accuracy, Human computing and Social computing systems, Crowdsourcing and Amazon Mechanical Turk (MTurk) and Collective intelligence, System design (ESP game, reCAPTCHA etc), Bittorrent and Peer-to-peer file sharing	Lectures. A website or/and moodle webpage will be created for the course and the course materials such as lecture notes, sample codes, will be shared. By using moodle, students will also be able to discuss relevant ideas and have questions answered by the lecturer.	Emine Yilmaz, Jun Wang	[1] Noam Nisan (Editor), Tim Roughgarden (Editor), Eva Tardos (Editor), Vijay V. Vazirani (Editor), Algorithmic Game Theory, Cambridge University, 2007., [2] David Easley and Jon Kleinberg, Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press, 2010, [3] R. Preston McAfee, Introduction to Economic Analysis www.mcafee.cc/Introecon/IEA.pdf, [4] Nir Vulkan, The Economics of e-Commerce, Princeton University Press, 2003, [5] Carl Shapiro, Hal R. Varian, Information rules: a strategic guide to the network economy, 1999	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University of Koblenz-Landau	Introduction to Web Science	Basic understanding of computer science as is taught in a type-2 bachelor programme. Expected knowledge will include basic capabilities of programming in a language like Java or C, algorithmic understanding, knowledge about basic data structures and basic internet networking.	Lectures, Written or oral exam (depending on class size)	Ansgar Scherp, Sergej Sizov, Steffen Staab,York Sure	Brügger, Niels (2010). Web History. Peter Lang.,Tim Berners-Lee and Mark Fischetti, Weaving the Web, 1999., Lawrence Lessigund Jonathan Zittrain. The Future of the Internet - And How to Stop It. Yale University Press, 2008/2009,Tim Berners-Lee, Wendy Hall, James A. Hendler, Kieron O'Hara, Nigel Shadbolt, Daniel J. Weitzner. A Framework for Web Science. Foundations and Trends in Web Science, Now Publishers, 1(1), 2006; DOI: 10.1561/1800000001.	Masters	Summer Semester	English
University of Koblenz-Landau	Network Theory and Dynamic Systems	Graph Theory and Social Networks, a.Paths, b.Core network measures, c.Strong and weak ties, d.Homophily and link prediction, e.Taxonomy of network types,, II.Game Theory, a.Definition of a game, b.Best responses and Nash equilibrium, c. Mixed strategies, d.Pareto and Social Optimality, e.Modeling network traffic using game theory, f.Tragedy of the commons,III.Information Networks and the World Wide Web, a. Structure of the Web, b.Link analysis, c. Sponsored search markets,IV. Network dynamics: Population models, a.Information cascades, b. Economy with/without network effects, c.Stability, Instability and Tipping points, d. Power Laws and rich-get-richer phenomena, e. Long tail,V. Network dynamics: Structural models, a. Diffusion, b. Small-world, c. Epidemics, 6. Group decision making, a. Different voting schemes,	Lectures, Exercise, Written or oral exam (depending on class size)	Steffen Staab and multiple lecturers	David Easley and Jon Kleinberg: Networks, Crowds, and Markets - Reasoning About a Highly Connected World, Cambridge University Press 2010	Masters	Summer Semester	English
University of Koblenz-Landau	Web Retrieval	The lecture will give an introduction in established retrieval models for text based documents, models that exploit the graph structure of the WWW, the topic of evaluating the performance of retrieval systems and related tasks like classification and clustering of web documents. The concepts communicated in the lecture will be applied in practical exercises and tutorials., More specifically the lecture will cover the topics: Information seeking behaviour on the web and user models, Evaluation of retrieval systems, Boolean retrieval and essential data structures, Vector space retrieval model, Probabilistic information retrieval models, Language models, Cross language retrieval, Topic models, Web crawling, Authority	Lectures, Exercise, Written or oral exam (depending on class size)	Sergej Sizov, Steffen Staab	R. Baeza-Yates, B. Ribeiro-Neto. Modern Information Retrieval. Addison-Wesley. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press, 2008.	Masters	Variable Teaching times	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
		ranking, Document clustering and classification, Information extraction						
University of Koblenz-Landau	Semantic Web	The student should be enabled to design and develop semantic Web applications. The student should be enabled to progress Semantic Web technologies in order to broaden and facilitate their use. The student should be enabled to understand the interaction of different standards, their technical implications as well as the social processes that underly various Semantic Web applications.,Foundations, Problems, Basic approach, Languages, RDF, OWL, Rule Languages, Ontologies, Applications, Modeling trade-offs, Foundational Ontologies, A Core Ontology for Software, Semantic Annotation, Re-using Databases, Information Extraction, Multimedia Annotation, Ontology Alignment, Information Flow, Learning Alignments	Oral or written exam depending on class size. Participation in the tutorial is a prerequisite for admission to the examination.	Ansgar Scherp, Sergej Sizov, Steffen Staab	S. Staab, R. Studer, Handbook on Ontologies, International Handbooks on Information Systems, Springer Verlag, 2004, S. Handschuh, S. Staab, Annotation ort he Semantic Web, IOS Press, 2003,P. Hitzler, S. Rudolph, M. Krötzsch. Foundations of Semantic Web Technologies, Chapman & Hall, 2010, A. Dengel (Hrsg). Semantische Technologien. Spektrum, 2012.,J. Domingue, D. Fensel, J. Hendler (Eds) Handbook of Semantic Web Technologies, Springer 2011.	Masters		English
University of Koblenz-Landau	Web Engineering	Participants should be able to master the main activities needed for the development of large software systems. They should be able to use the main languages and to apply the main methods of software engineering. They should be able to describe different views on software using UML, and they should know the most important software process models., The students understand the particuliarities of web engineering compared to classical software engineering. They have fundamental knowledge of the languages involved in web-based systems, and they are able to classify the most important technologies and tools used. They have deepened knowledge of software processes with respect to the area of web-based systems., Introduction, web applications, requirements, characteristics and quality goals, World Wide Web, hypermedia, languages (html, xml), protocols and layers, application protocols, terminology, languages (http), Server-side Components, application servers, frameworks, components, languages (php), Client-side Components, browsers, plugins, languages (javascript),Web Development Process, requirements, modeling, architecture, quality assurance	Lectures, Assessment: Exam	Jürgen Ebert	Emilia Mendes, Nike Mosley: Web Engineering. Springer, 2006, ISBN 978-3-540-28196-2, Gerti Kappel, Birgit Pröll, Siegfried Reich, and Werner Retschitzegger: Web Engineering - The Discipline of Systematic Development of Web Applications. John Wiley & Sons, 2006	Masters	Winter semester	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University of Koblenz-Landau	E-Participation	The course will provide a deeper understanding of the field e-participation. It will deal with the following contents:, Overview of e-participation as a research field, Research methods, theoretical grounds as well as tools and technologies for e-participation, Methods for analyzing, conceptualizing and managing e-participation projects, Introduction to selected cases of e-participation. The cases will be investigated along theoretical grounds, methodological approaches used in the case implementation, technical and organizational settings, as well as impact of e-participation endeavor and use of solution., Monitoring and evaluating e-participation projects, Impact of e-participation to society and assessment of the public value-add	Lectures and Seminars;, Assessment: Examination	Maria Wimmer	Proceedings of annual eParticipation conferences (Springer LNCS), R. Davis (ed.), The Web of Politics. Oxford: Oxford University Press, Deliverables of DEMO-net. See www.demo-net.org, Project cases and deliverables of MOMENTUM and respective projects. See www.ep-momentum.eu, Proceedings of Online Deliberatin 2010, Leeds University Business School, UK, 2010, OECD, 2003, Promise and problems of e-democracy. OECD study, Macintosh, A. & Coleman, S. (2004). Promise and Problems of E-Democracy, Challenges of online citizen engagement, OECD,Initiative eParticipation & Stiftung MITARBEIT (Hrsg.): E-Partizipation - Beteiligungsprojekte im Internet. Verlag Stiftung MITARBEIT: Bonn, 2007	Masters	Summer Semester	English
University of Koblenz-Landau	Online Consumer Behaviour	Students understand the characteristics of consumer and purchasing behavior on the Internet and can thereby design the marketing process of Internet companies. In particular, students can analyze consumer behavior in different contexts. Referring to the traditional marketing, students master the tools of online marketing and are able to distinguish those from the classical marketing mix instruments., Online Consumer Behaviour -- Introduction to the course; E-commerce market mechanism; Retailing in e-commerce; The culture of the Internet, Marketing strategy and the web basic context; Consumer behaviour (Decision making process, 1to1 marketing , market research, customer services etc), Consumers as Decision-Makers; Group Influence, Opinion Leadership, seeking/giving word of mouth, Pricing and Marketing models strategy and decision, Promotion and IMC online; Web advertising, methods, topics; Multichannel communication campaigns, Place: Online geographies of Virtual Technologies; Convergence or not; Networked places; The digital divide, Legal and Ethical issues of	Lectures and seminars, Assessment: Exam	wechselnde Dozenten, Matthias Gouthier	D. Chaffey, F. Ellis-Chadwick, K. Johnston, R. Mayer, Internet Marketing, 4th edition, Prentice Hall 2009, J. Strauss, R. Frost: E-Marketing, 2nd Edition, Prentice Hall, 2000, G. Zinkhan, Advertising Research - The Internet, Consumer Behavior, and Strategy, South-Western Educational Pub 2000	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
		online buyer behaviour and E-business and Marketing; Computer networks civil society and the state; Privacy liberty in cyberspace; Security						
University of Koblenz-Landau	New Product Developme nt	The lecture is devoted to the development of new products (including new services) with a strong focus on the new product development process and technology marketing. In particular, the aim of the course is to understand the new product development process, to learn how to integrate the customer and knowledge of the customer into this process, to learn and apply concepts and tools appropriate for new product development analysis, and to develop specific recommendations and action plans for companies facing difficult decisions about bringing new products to market.,Introduction to new product development: Invention vs. Innovation, New products in consumer markets, business to business markets, The new product development process: Idea phase, conceptual phase, development phase, launch, Intellectual Property, Outsourcing and Out-Licensing, Strategic Product Devlopment: Blue Ocean Strategy, Niche Markets, Two sided platforms, Product Development with Communities, Product Portfolio Management, Technology Marketing,	Lecture: Test (60 min) and Case work, Seminar: Assignment for a particular topic (approx. 5000 words) and presentation (15 min) plus discussion	Mario Schaarschmidt	Crawford, M. and di Benedetto, A. (2011). New Products Management, 10th ed. Irwin, MacGraw Hill., Trott, P. (2008). Innovation Management and New Product Development. 4th ed. Wiley.,Ulrich, K.T. and Eppinger, S.D. (2007). Product Design and Development, 4th ed, McGraw-Hill/Irwin	Masters	Every third semester	English
University of Koblenz-Landau	Computatio nal Social Science	The seminar introduces students into the research area of Computational Social Science., The students will be trained to select, test, apply and provisionally evaluate methods from the computer science field in order to answer social science questions., A basic understanding of the possible application of Data Mining methods will be developed for this purpose, as well as an understanding of the significance and possibilities of operationalization of issues and problems established in the social sciences.,The content of the module include at least following areas : Social Science Research Questions and Methods, The Small World Phenomenon, Search in Social Networks, Python and Octave tutorials, Social Network Analysis, Affiliation Networks, Community Detection, Social network models and Generation.,	Lecture, Seminar / exercise, written exam: 50%, home assignment (code submissions): 50%	Markus Strohmaier	Stanley Wasserman and Katherine Faust, Social Network Analysis - Methods and Applications, 1995, Monge, Peter R., and Noshir S. Contractor. Theories of communication networks. New York: Oxford University Press, 2003.,David Easley and Jon Kleinberg, Networks, Crowds, and Markets: Reasoning About a Highly Connected World, 2010 (free online book)	Masters	Summer Semester	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Univeristy of Liverpool	Computer Structures	This module covers everything from computer architecture to databases, algorithms, languages, operating systems, communications, computer networks, artificial intelligence and the theoretical foundations of computation. By studying this module, students will gain a sound theoretical and practical grounding on which to build their understanding of future technical developments., Integrate and outline the fundamental terminology and paradigms of CS., Compare the basic hardware and software components of a computer system., Appreciate the purpose of the operating system and of computer networks., Construct and outline (or design) algorithms in a clear and unambiguous manner., Appreciate what is meant by the complexity of an algorithm, and by computability., Produce and structure data organisation, from its initial structure up to databases., Organise and critique current and potential future trends in CS, including Artificial Intelligence, Software Engineering, modern high level languages.	Online		Glenn Brookshear, Computer Science: An Overview (Tenth Edition-2008), Addison Wesley/Pearson, World: ISBN: 0-321-54428-5, ISBN 13: 978-0-321-54428-5/US: ISBN: 0321524039, ISBN 13: 9780321524034	Masters		English
Univeristy of Liverpool	Programming the Internet	This module provides students with a critical understanding of markup languages and related advanced technologies. On completion of the module students will be able to design and create advanced website applications and will have the knowledge to undertake complex internet projects., An ability to critically appraise and employ successfully Internet Programming tools, techniques and current standards., An ability to critically assess the merits of alternative solutions to Web programming problems., A critical understanding of the process required to construct www applications that access a back-end databases., A full understanding of the process required to build scalable web database systems., An ability to design and create an effective and maintainable Web site in accordance with established Web standards.	Online		Deitel, H.M., Deitel, P.J., Goldberg, A.B, Internet & World Wide Web How to Program, Prentice Hall, 4th edition, 2008, ISBN 0-13-175242-1	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Univeristy of Liverpool	Big Data	The Big Data module provides students with an insight into this currently developing area of interest in the field of Computer Science. Students who complete the module will acquire a critical understanding of the concept of Big Data and its relevance in the modern commercial world. The module will also allow students to acquire understanding and skills for using analytic techniques that can be used to gain insight and predictive value with respect to Big Data. An in-depth and critical understanding of the concept of 'Big Data', the analytic techniques that can be used with respect to Big Data, and how these techniques can be used to gain competitive advantages., An awareness of major cases of Big Data usage in science and industry and the associated Big Data challenges., A critical understanding of the Big Data Architecture Framework, the main components and their inter-relation., A critical understanding of Big Data infrastructures., A critical understanding of data structures used in the context of Big Data., An understanding of Big Data Analytics tools and platforms and methods., An understanding of Big Data security and protection issues., A critical awareness of the commercial relevance of Big Data.	Online		Cathy O'Neil and Rachel Schutt (2014). Doing Data Science: Straight Talk from the Frontline. O'Reilly, 2014.	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Univeristy of Liverpool	Professional Issues in Computing	To function effectively, professional IT practitioners need not only the appropriate technical knowledge, skills and experience, but also a broad understanding of the context in which they operate. IT professionals must understand the relationship between technological change, society and the law, and the powerful role that computers and computer professionals play in a technological society. This module helps IT professionals to be capable in their chosen professional areas, operate effectively in work and community situations and to be aware of their environments., Differentiate between ethics, and ethical behaviour, and approach to ethical decision making in an IT context., Explain the roles and responsibilities of IT professionals in organisations and society, and evaluate relevant work-related behaviours and issues and how professionals should respond., Identify the legislative and regulatory obligations and standards as they relate to IT environments in organisations, including freedom of speech and illectual property, and appreciate their impact on the work of the IT professional., Assess the impact of business compliance imperatives and how these impact on the work of the IT professional., Analyse the issues involved in consulting and contracting., Demonstrate a sound understanding of the full range of impacts of IT on human societies & organisations., Analyse trends that have increased the risk of misusing information technology., Identify and describe the elements of Risk Management: planning, identification, analysis, response, monitoring and control as the relate to IT.	Online		Herman Tavani (2013) Ethics and Technology: Controversies, Questions and Strategies for Ethical Computing, 4th ed., John Wiley & Sons Ltd. (Electronic version also available)	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
Univeristy of Liverpool	Professional Issues in Computing	To function effectively, professional IT practitioners need not only the appropriate technical knowledge, skills and experience, but also a broad understanding of the context in which they operate. IT professionals must understand the relationship between technological change, society and the law, and the powerful role that computers and computer professionals play in a technological society. This module helps IT professionals to be capable in their chosen professional areas, operate effectively in work and community situations and to be aware of their environments. Differentiate between ethics, and ethical behaviour, and approach to ethical decision making in an IT context., Explain the roles and responsibilities of IT professionals in organisations and society, and evaluate relevant work-related behaviours and issues and how professionals should respond., Identify the legislative and regulatory obligations and standards as they relate to IT environments in organisations, including freedom of speech and intellectual property, and appreciate their impact on the work of the IT professional., Assess the impact of business compliance imperatives and how these impact on the work of the IT professional., Analyse the issues involved in consulting and contracting., Demonstrate a sound understanding of the full range of impacts of IT on human societies & organisations., Analyse trends that have increased the risk of misusing information technology., Identify and describe the elements of Risk Management: planning, identification, analysis, response, monitoring and control as they relate to IT.	Online		Herman Tavani (2013) Ethics and Technology: Controversies, Questions and Strategies for Ethical Computing, 4th ed., John Wiley & Sons Ltd. (Electronic version also available)	Masters		English
Univeristy of Liverpool	Cloud Computing	Cloud Computing is an important new paradigm in the area of Information Technology (IT) that currently provides a (common) basis for a new wave of technology development such as Big Data and in pushing further the mobility-based pervasive computing concept. Cloud Computing embodies advances in computing, networking, and storage technologies which individually and collectively include major hardware and software breakthroughs. These include computer virtualisation, distributed and replicated storage, and software based networking. In its own turn, Cloud Computing motivates further technology advancement and changes how the modern IT infrastructure is built and evolves. There is a growing demand for specialists with strong technical background and deep knowledge of the Cloud Computing technologies.,Explain concepts and major application areas., Critically compare similar concepts (and concepts inter-relation) and alternatives as well as application-specific areas., Classify technologies and	Online		No Core text	Masters		English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
		their relation to concept.						
Univeristy of Liverpool	Social Computing	The last decade has seen an exponential growth of the concept of social computing as exemplified by platforms such as Facebook, Twitter and LinkedIn. Social computing is a phenomenon that has great potential with respect to both commercial and non-commercial enterprises, especially with respect to marketing and public relations. The aims of this module are to provide students with both an in-depth understanding of the social computing phenomenon and of how the analysis of such social media can be beneficially undertaken. More specifically the module is directed at providing students with the know-how and capabilities to analyse social media while at the same time understanding the social computing context and the legal framework in which it resides., Summarise the history of social computing and networking and critically review predictions for the future., Analyse social computing and networks using nodes and ties in a graphical representation., Deploy technologies to harvest, analyse and visualise social data., Identify the technology requirement for a business looking to add a social computing and networked methodology., Analyse and apply the use of social learning services for eLearning and Virtual workgroups., Analyse existing social computing and networked services and recognise their strengths and weaknesses., Develop recommendations for improvement of existing social network products or systems., Develop a professional and ethical approach for creating, maintaining and utilising new and existing social network services., Recognise the concerns relating to social networking and develop informed responses and strategies with respect to these issues.	Online		Shelly, G. and Frydenberg, M. (2011) Web 2.0, Concepts and Application Cengage (ISBN-10: 1439048029; ISBN-13: 9781439048023) (available in Paperback, eBook, Rental or eChapter Editions)	Masters		English
University of Southampton	Foundations of Web Science	Introduction to Web Science, Web Science and the Law, Statistical models for use with Web Science, Methods and techniques for social study, Game theory and mathematical models for Web Science, Biological underpinnings for Web Science, The economics of Web Science,	Lecture - 36 hours per semester, Tutorial - 12 hours per semester, Assesment: 10% - Book review. 90% - Wiki contribution.	Prof Leslie Carr, Dr Mark J Weal, Prof Nigel R Shadbolt, Professor Dame Wendy Hall, Claire Wyatt	http://www.edshare.soton.ac.uk/1647/	Masters	Term 1, Oct 2014 to Jan 2015	English
University of Southampton	Independent Interdisciplinary Review	the concepts innon-computing disciplines that are applicable to studies of the Web, the differences in disciplinary approaches to Web analysis, relate methodologies and techniques to a range of practical applications, techniques and issues in designing, building and modelling Web systems and Web users, relevant technologies and tools for modelling the Web and its use,	Lecture - 36 hours per semester, Tutorial - 12 hours per semester, Assessment: 70% - Literature Search, Interim and Final report., 20% - Poster and Presentation.,	Prof Leslie Carr, Dr Su White	http://www.ecs.soton.ac.uk/module/COMP6044#syllabus	Masters	Term 1, Oct 2014 to Jan 2015	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
			10% - Peer review.					
University of Southampton	Hypertext and Web Technologies for Masters	Principles of Web Architecture, Principles of Web Information Design, Searching the Web Graph, the factors that have led to the development of the web in its present form and current trends in the development of hypertext in the web.	Lecture - 36 hours per semester, Tutorial - 12 hours per semester, Assessment: 30% - Research Issues in Hypertext Paper, 70% - Exam, 0 hour(s)	Dr Nicholas Gibbins, Prof Leslie Carr	Lowe D and Hall W, Hypermedia and the Web-An Engineering Approach, Wiley 1999, Rosenfeld L, Information Architecture for the World Wide Web, 2nd Edition, O'Reilly 2002, Deitel, Deitel and Goldberg, Internet and World Wide Web - How To Program, 3rd Edition, Prentice Hall 2004;, LINK: http://www.ecs.soton.ac.uk/module/COMP6045#syllabus	Masters	Term 1, Oct 2014 to Jan 2015	English
University of Southampton	Computational Thinking	Operating systems (1960s, resource management, UNIX/Linux, Windows, Mac, thin clients, cloud computing), Databases (SQL, third normal form, Hadoop, data centres), Devices (Mainframes, PCs, iPhones, sensor networks), Programming Languages (binary, assembler, C, Object orientation, Java, LISP, Prolog, functional, scripting), Algorithms (sorting, complexity, tractability, IP), Artificial Intelligence (Lisa to Machine Learning, Neural Networks), Graphics (OpenGL, PS3! GPUs), Software Engineering (methodologies, projects, mythical man year), Networks (Ethernet, X25, TCP/IP, routers, IPv6, Wifi, 3G, Wimax) visit ORC.,	Lecture - 36 hours per semester, Tutorial - 12 hours per semester, Assessment: 30% - Journalistic Article, 20% - Secondary school or 6th form college teaching activity, 50% - Public engagement lecture.	Prof Hugh Davis, Prof Leslie Carr	Wing, J. (2006) Computational thinking, Communications of the ACM, v.49 n.3	Masters	Term 1, Oct 2014 to Jan 2015	English
University of Southampton	Further Web Science	the nature and history of Web Science as an emerging research area;, the breadth of disciplines that contribute to Web Science research;, a detailed understanding of a number of areas of Web Science;, approaches to interdisciplinary research,	Lecture - 36 hours per semester, Tutorial - 12 hours per semester, Assessment: 100% - Exam, 0 hour(s)	Prof Leslie Carr	EDShare	Masters	Term 2, Jan 2015 to June 2015	English
University of Southampton	Interdisciplinary Thinking	The issues surrounding navigating the languages of different disciplines, Case studies in the application of interdisciplinary approaches to real-world problems, Methods for constructing arguments from multi-disciplinary perspectives, Critical analysis in an interdisciplinary setting	Lecture - 36 hours per semester, Tutorial - 12 hours per semester;, Assessment: 10% - Group article 1., 10% - Group article 2., 10% - Group article 3., 70% - Individual article.,	Prof Leslie Carr	EDShare, Module Wiki, Repko A. F. (2008) Interdisciplinary Research: Process and Theory. Sage Publications.	Masters	Term 2, Jan 2015 to June 2015	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University of Southampton	Semantic Web for Web Scientists	Overview and Introduction, Knowledge Representation, Ontologies and Description Logic, Semantic Web in Depth: RDF and RDF Schema, Semantic Web in Depth: OWL, Writing OWL ontologies with Protege, Semantic Web Methodologies and Design Patterns, Semantic Web in Depth: SPARQL, Semantic Web in Depth: Rules, Linked Data and Publishing on the Semantic Web, Semantic Web Vocabularies and Applications, Semantic Web vs Web2.0, Trust and Community,	Lecture - 24 hours per semester, Tutorial - 12 hours per semester, 40% - Ontology Design. 60% - Exam, 0 hour(s)	Dr Nicholas Gibbins	EDShare	Masters	Term 2, Jan 2015 to June 2015	English
University of Southampton	Social Networking Technologies for Web Science	Service architectures for social networks, Common APIs for popular architectures (Facebook, OpenSocial, etc), OpenID and Shibboleth, Linked Data for social networks (FOAF, SKOS, etc), Social network properties and analysis methodologies, Social network interoperability, Social network topologies and ecosystems, Social networks in e-learning, enterprise and media, Identity, privacy and ownership in social networks, Aspects of recommendation engines and information retrieval in social networks,	Lecture - 12 hours per semester, Tutorial - 12 hours per semester;, Assessment: 20% - Reflective Summary., 80% - Exam, 2 hour(s)	Dr Thanassis Tiropanis, Dr Dave Millard	EDShare	Masters	Term 2, Jan 2015 to June 2015	English
University of Southampton	PAIR1001 Introduction to International Relations	This module provides an introduction to the study of international relations and focuses on the three classic approaches to international relations: realism, liberalism and Marxism. You will learn about the political theories underlying each of these approaches and their evolution to the present day. Seminars examine the links between these approaches and contemporary world politics.	Teaching methods include mainly lectures, seminars and class discussion. Key learning methods include reading, instructor-led discussions and student-led seminars, and research in preparation for classes and assessments.	Dr Ana Margheritis	Baylis, John, Steve Smith, & Patricia Owens, eds., The Globalization of World Politics (Oxford University Press, 2008)., Robert J. Art and Robert Jervis, eds., International Politics: Enduring Concepts and Contemporary Issues, 8th edition (New York: Longman, 2007)., David Held and Anthony McGrew (eds), The Global Transformations Reader - Second Edition, (Cambridge: Polity, 2003).	Undergraduate	YEAR1: Term 1, Oct 2014 to Jan 2015	English
University of Southampton	SOCI1001 Understanding Everyday Life	Why do young people so often seem to get a bad deal in the press? Why do working class kids get working class jobs? Why do women earn less than men? Does race still matter in British society? How does our family background influence our chances in life? These and other questions are explored in The Sociology of Everyday Life- a module which brings you some of today's most cutting edge sociological research by a team of leading sociological researchers.	Learning activities include: seminar discussions, small group work. Contact hours: 29, Private study hours: 121, Assessment: 1500 word essay 30%, 2-hour examination 70%,	Professor Pauline Leonard	Hunt, S (2005)The Life Course: A Sociological Introduction, Basingstoke: Palgrave, Hockey, J and Allison J (2003) Social Identities across the Life Course Basingstoke, Palgrave., Ortner, S.B. (2006) Anthropology and social theory : culture, power, and the acting subject Durham/London: Duke University Press,Payne, Geoff (ed) (2006) Social divisions Basingstoke, Palgrave,	Undergraduate	YEAR1: Term 1, Oct 2014 to Jan 2015	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University of Southampton	WEBS1001 Information , Technology and Social Change	The web, historical perspectives, Information Society, conceptualisation and approaches, The architecture of the Web, Web and society, Analysing the Web (introducing a range of scientific approaches to studying the Web including mathematical, statistical and social science approaches), Socio-technical perspectives on the Web,	Lectures (2 per week) will be used to present the material listed above, Seminars (alternate weeks) will discuss directed learning, Tutorials (alternate weeks);, Assessment: Examination 2 hours 60%, Coursework 1500-word essay 40%	Dr Mark Weal	http://www.southampton.ac.uk/webscience/undergraduate/modules/webs1001_information_technology_and_social_change.page#aims	Undergraduate	YEAR1: Term 1, Oct 2014 to Jan 2015	English
University of Southampton	Web Design	the technical principles of the World Wide Web;, the technical issues in creating and publishing content on the World Wide Web;, the legal issues when publishing content on the World Wide Web.	1- hour lecture, three times each week, 3-hour computer lab, once each week, Assessment:, 40% - Labs, 20% - Report, 40% - Website	Dr Thanassis Tiropanis, Prof. Alex Rogers	Sylabus: http://www.southampton.ac.uk/webscience/undergraduate/modules/comp1056_web_design.page#syllabus;, TEXTBOOKS: Castro, E. (2007). HTML for the World Wide Web [6th Edition]. Peachpit Press., Gosselin, D. (2008). JavaScript [4th Edition]. Thomson Course Technology.	Undergraduate	YEAR1: Term 2, Jan 2015 to June 2015	English
University of Southampton	Social Sciences: Sociology, Social Policy & Criminology	This unit looks at the emergence of 'modern' society and provides key concepts to understand the continuing social transformations of modernity. Looking at topics such as the tyranny of 'experts', the modern city, fashion, and global tourism, the unit examines the key processes of modernisation and how these have shaped the everyday lives of individuals. By examining the rise (and fall) of mass urban, industrial society, the unit shows how consumption and culture, space and nature, and individual self-identity have all been transformed by modernity. Attention is given to class, gender, race and ethnicity.	seminars and small group discussions;, Assessment: Examination and Essay	Dr Silke Roth	Robin Cohen & Paul Kennedy (2013) Global Sociology. Palgrave, Stuart Hall & Bram Gieben (1992) Formations of Modernity. Polity/Open University	Undergraduate	YEAR1: Term 2, Jan 2015 to June 2015	English
University of Southampton	STAT1003 Introduction to Quantitative Methods	Types of Data and Making/Using Tables, Graphical Methods, Averages and Percentiles, Measures of Spread and Inequality, Transforming Data, Correlation, The Normal Distribution, Samples and Sampling Distributions, Confidence Intervals, Simple Hypothesis Tests,	Weekly online exercises, assignment and examination.	Dr Andrew 'Amos' Channon		Undergraduate	YEAR1: Term 2, Jan 2015 to June 2015	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
University of Southampton	SOCI2031 Social Theory	• Identify and outline the principal theoretical approaches and key theoretical concepts used in the social theoretical examination of social life. • Analyse the contributions made by these approaches to debates concerning the nature of social worlds and the social and cultural phenomena that occur in them. • Critically evaluate the contributions made by selected social theories to our understanding of social life • Develop understanding of the theoretical frameworks that inform work in the social science disciplines (Sociology, Anthropology, Criminology, Psychological Studies) • Develop skills in theoretical discussion, analysis, critique and language.	Lectures and Seminars;, Assessment; Examination and Essay.	Dr Bindi Shah	Crow, G (2005) The Art of Sociological Argument, Basingstoke, Palgrave, May, Y & Powell, J (2008) Situating Social Theory, 2nd edition, Buckingham, OU Press	Undergraduate	YEAR2: Term 1, Oct 2014 to Jan 2015	English
University of Southampton	WEBS2002 Interdisciplinary Group Project	This module is offered in the context of a multi-disciplinary programme. The purpose of the module is to allow students to understand the challenges and problems that come from trying to reconcile multiple disciplinary perspectives and value systems on a single problem. This module draws together all the multidisciplinary content and methodologies that they have engaged with, and helps them to understand how to marshall them in a practical, commercial or political context.	Lectures (1 per week);, Assessment: Group report 60%, Individual report 40%	Dr Mark Weal	Repko, A. (2008). Interdisciplinary Research: Process and Theory. Sage Publications.	Undergraduate	YEAR2: Term 1, Oct 2014 to Jan 2015	English
University of Southampton	COMP2213 Interaction Design	User Psychology, Hardware (input/output) devices, Models and Metaphors, Interaction styles, Graphical User Interface (GUI) and windowing systems, Design methodology, Accessibility, Guidelines, standards and metrics, Evaluation, Advanced Interfaces,	1-hour lecture, three times each week, 1-hour tutorial, once each week;, Assessment: 75% - Exam, frequency: 1, exam duration: 2, 25% - Coursework, exam duration: 37.5	Dr Mark Weal, Dr Enrico Costanza	Preece, J., Rogers, Y., Sharp, H. (2007). Interaction design: beyond human-computer interaction [2nd Edition]. Wiley., Dix, A., Finlay, J., Abowd, G. and Beale, R. (2003). Human-computer interaction [3rd Edition]. Prentice Hall., Shneiderman, B. (2004). Designing the user interface [4th Edition]. Addison Wesley., Norman, D. (2002). The design of everyday things. Basic Books., Preece, J. and Keller, L. (1990). Human-computer interaction. Prentice Hall., Barfield L. (1993). The user interface: concepts and design. Addison Wesley., Cox, K. and Walker, D. (1993). User interface design [2nd Edition]. Prentice Hall., Preece, J., Rogers, Y., and Sharp, H.	Undergraduate	YEAR2: Term 2, Jan 2015 to June 2015	English

Location	Resource	Topic(s)	Course	Contributors(s)	Materials	Level	Duration	Language
					(2002). Interaction design: beyond human-computer interaction. Wiley., Sharp, H. (2007). Interaction design [2nd Edition]. Wiley., Smith-Atakan, S. (2006). The fast track to human-computer interaction. Thomson Learning., Jones, M. (2006). Mobile interaction design. Wiley.			
University of Southampton	SOCI2020 Research Skills	The unit will be delivered through a programme of lectures and seminars. The lectures will examine some of the key issues in the design and conduct of social science research including what counts as good research, quantitative and qualitative research, documentary research, access and ethical issues, how to review relevant literature, and organisation and time management. The seminars will be based around group activities to develop a hypothetical research design and guidance for the production of individual dissertation proposals. There will also be a library skills information session.	Lectures, Seminars, Practical exercises, Academic guidance on the production of a dissertation proposal, Independent study;, Assessment: Qualitative data collection and analysis exercise (2000 words) 30%, Dissertation proposal, 70%	Dr Katy Sindall	Seale, C. (2012) Researching Society and Culture (3rd edition), London: Sage.	Undergraduate	YEAR2: Term 2, Jan 2015 to June 2015	English
University of Southampton	COMP3016 Hypertext and Web Technologies	Principles of Web Architecture, Principles of Web Information Design, Searching the Web Graph, History and Current Research Issues in Hypertext,	1-hour lecture, three times each week, 1-hour tutorial, once each week;, Assessment: 70% - Examination, 30% - Research Issues in Hypertext Paper,	Dr Nicholas Gibbins	http://www.southampton.ac.uk/webscience/undergraduate/modules/comp3016_hypertext_and_web_technologies.page#syllabus	Undergraduate	YEAR3: Term 1, Oct 2015 to Jan 2016	English

16.2 Feedback from Web Science 2015 Workshop Presentation

Written notes from the 'Sound Recording script'

Differentiation was a key theme of this discussion – how do you tell if a programme is Web Science or not!?

Adriana Wilde Feedback:

- The list of institutions is not comprehensive
- Institutions don't offer info on their web sites – this is to be expected as it is a young discipline – part of the nature of such a young and fast developing subject.

Stefan Bazan Feedback

Stefan spoke a lot about the 'Web Science research and education community' – strong emphasis on this!

- "We are trying to better organise the Web Science community from inside the Web Science research community"
- Every year we try and list all the existing programmes
- Can look at this from two different points of view. You can start a Web Science programme and that's labelled as 'Web Science' and you're studying 'official Web Science' – the point of the Web Science education/research community was not to label programmes as Web Science.
- In Pakistan there is a Web Science course
- Making the map of web science on the education side is very difficult, because some people create a web science curriculum that corresponds to web science labelling and keywords because they know the community and are part of it, whereas other's may create a programme that looks like web science and call it something different, or something that looks like Web Science when it isn't!? – ***Lesson*** *- Differentiation is needed, but is very difficult!*
- Using the keywords 'web and science' to search produces a very small sample – because a lot of programmes look like Web Science, but are not called 'Web Science' – ***Lesson****: need to expand keyword searches to look for related topics?*
- If you are not part of the Web Science community – it is hard to call a programme 'Web Science'
- ***External input:*** – what if it isn't actually web science!? Example – Digital sociology overlaps quite a lot, but it's not the same subject…Person argued that it might be good to exchange with the digital sociology community, but that it ultimately not be the same programme. ***Lesson*** *– good idea to still represent these related programmes when they are found, but maybe differentiation is required – however, where is the line to be drawn between what is just related to Web Science and what actually is?!*
- Argument that Web Science is interdisciplinary, and should not be a single community, but a community of communities?! (A melting pot!?) Each person belongs to several communities – e.g. Web Science and Computer Science

Manuel Leon

- In Spain there are at least two programmes – one in Madrid and one in Barcelona – (similar to Digital Sociology) they are labelled as 'Digital Citizenship' – overlap with Web Science Masters at Southampton is quite big. They are more similar to the Web Science programme at Southampton than the (Digital Sociology?) Programme at Goldsmiths.

  When investigated, the following English speaking programmes were found at Madrid, however there was no reference to a digital citizenship programme.

    - *Bachelor in Architecture, Bachelor in Business Administration, Bachelor in Communication, Bachelor in International Relations, Bachelor of Laws, Bachelor in Psychology*

Su White

- Part of the knowledge is knowing what people call it – also understanding the overlaps between the subjects – almost like sets of overlapping bubbles. (Like the Web Science cluster diagram!?) Where do the fringe subjects fit in; how are they relevant and what is relevant to Web Science?
- The method for finding this out, is a bit like 'chain sampling' – when you're trying to find expertise, but you don't know who you've got to ask, you have to go by personal recommendations. Some of these recommendations may be dead ends, but you have to cast your net and extend your net in all ways that you possibly can….

More *Feedback from Stefan Bazan*

- I would suggest that, in your survey, if you can, look for the intention behind the creation of a programme. Was it the intention to create a Web Science programme, or was the intention to merge existing courses…and label them Web Science. Or is it just something that looks like Web Science, or is it even Web Science at all within the large scope, borders and boundaries of Web Science.
- The Web Science Masters in Beirut was designed according to the curriculum – we took what Vafopolos did with the Web Science Subject Categorisation, and extracted one section and stuck to it. The intention was really to what was outlined by the Web Science Trust at the time.

Lady 1's Feedback –

- The curriculum is very detailed - can provide search terms – look for the topics themselves, this may yield more results for related programmes. (Such as Digital sociology)

More Feedback from Adriana Wilde:

- Even in the .ac.uk domain you could still try searching for those keywords.

Gentleman 1's Feedback:

- MIT has a programme – it's called 'HASTS' 'History, anthropology, technology and science' – run for 20 years?! (Historic Web Science!?)
- Link: http://web.mit.edu/hasts/
- List of classes: http://web.mit.edu/hasts/classes/all-classes.html

Creating a Web Science textbook

Subsequent to the presentations, there was a period of discussion between delegates. One of the key conclusions of this discussion was the fact that it would be desirable to compile a Web Science textbook, based upon the Wikiversity Web Science MOOC. The consensus was that although Web Science is rapidly evolving, it would still be a worthwhile endeavour to attempt the creation of a Web Science textbook.

16.3 Trip Reports/Event Notes

16.3.1 Web Science 2015

This year's annual ACM Web Science was held in Oxford's beautiful Keble college.

I was presenting a workshop paper summarizing my research on the Web Science curriculum so far, including attempts to create a web crawler identifying web science teaching institutions.

I was also hoping to gain a better insight into some of the areas within the Web Science curriculum which I am not familiar with, albeit I was aware\expecting that this would be biased towards specific topics/areas.

As was the case with last year's conference content, I felt that there was (perhaps unsurprisingly) a lot of emphasis on social media based research. This year there was also a lot of emphasis on ethics, including a number of sessions which looked at robot ethics. Big data with also a recurring theme. I also heard numerous mentions of a quote by Sir Tim Berners-Lee "the Web we want" - sentiments of which were echoed throughout the conference. This involved an emphasis on the fact that we need to better understand the development of the Web in order to influence the development. It also suggests that the Web Science community should ask what we want of the Web if we want to shape it and influence it's development in a positive way.

Internet of Things was a key theme, including how robotics could lead to automated devices which tell us what to do - including intelligent smart meters which won't allow us to use more than a certain amount of electricity. Also networked devices which could potentially record and share all sorts of personal data about us with the world - This is where huge ethical questions were raised. Lead me to ask questions such as - should machines be allowed to tell us what to do? (A resounding no from my personal perspective!) If we have devices such a smart watches recording our heartbeat and life-signs, who should this data be shared with? Is gathering such data even ethical in the first place outside a strict medical context? - (These are some of the issues that were raised.)

I also had a number of useful conversations, including with Clare Hooper regarding her earlier work into topic representation within the WS subject. (She also agreed to complete my online survey)

16.3.2 University of Southampton Web Science and Dig Data Summer School

As suggested by the title, this summer school revolved around Big Data, and the uses of data, methods of data analysis, as well as ethical implications of Big Data.

Dame Wendy Hall gave a talk which explained the need for Web Science and a better understanding of how the Web works. Such an important worldwide resource, but without better understanding and management it could all collapse just as quickly as it was created. She cited examples such as the USA and China managing their own address space - could this be the start of the segmentation and decentralization of the Web? The entire concept is based upon the Web being 'worldwide' - what happens if countries begin to split off.....an interesting and somewhat worrying suggestion.

I was a little disappointed that there was a lack of context within the broader Web Science curriculum, - Web Science was not explained\promoted. (Considering it was part of the course title I expected some explanation\clarification of the Web Science. (Last year's WS3 'Age of Data' Southampton based Web Science summer school had far more emphasis upon the use of data within Web Science as a whole.) Many people I spoke to at the school had little idea of what Web Science was.....they attendees were heterogeneous in their backgrounds. A lot of people I spoke to seemed to come from multiple disciplines within Nottingham University, who do not teach a specific Web Science curriculum.

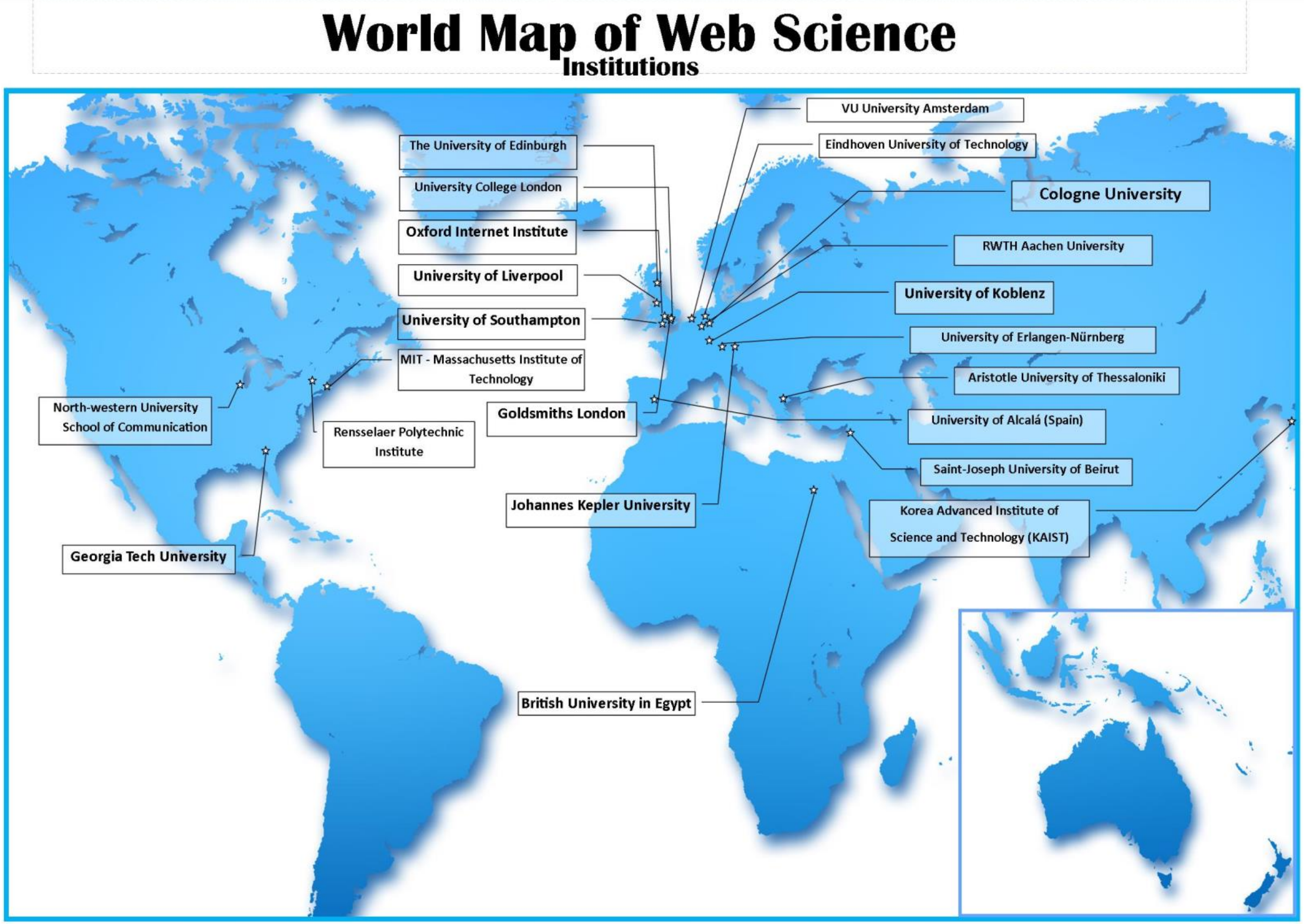
One of the main things I gained form this summer school was ideas and inspiration as to potential new data sources to aid in my research, (e.g. Web Science related tweets.) Another thing that the school brought to my attention was what I can do with my data - I was introduced to new data analysis tools, such as Pulsar and COSMOS - these are tools which facilitate analysis of social media data (e.g. Tweets) for people with little technical or programming knowledge. (Geared for social scientists). These tools also both facilitate data visualization. Pulsar is only usable with existing data sets, however COSMOS may be an option for free academic use, and is something which I will investigate further.

WEKA - is another tool for data analysis, however it is more biased towards qualitative statistical analysis, and is less suitable for the purposes of my own research, It's outputs are suitable for analysis with tools such as the programming language 'r' which I learnt is a programming language used by statisticians for data analysis including 'cleaning' data sets of erroneous and unnecessary information.

16.3.3 Note - Reflection on Tim's 2006 paper in relation to Web Science Teaching

Tim's paper (Berners-Lee, Weitzner, et al. 2006) puts the whole Web and the way that it works and why we need each protocol and layer into perspective. I was taught about many of these technologies as an Information Technology in Organisations undergraduate, but I never thoroughly gained an appreciation of where they fitted in in the bigger picture of the Web and why they were important. Had I seen the bigger picture, I have a feeling I would have found the study of what seemed like an obscure technology more important and relevant. It could be argued that introducing the Web Science perspective in teaching, as a way of putting a wider programme into context would be hugely beneficial, even for, for example, a Computer Science programme. E.g. "this is the Web, this is how huge it it….it has all these properties. We are going to be studying these aspects – this is where they fit in the bigger picture, and this is what you are going to be a part of, learning about and ultimately contributing to." – Looking at the scale and importance of this could be greatly motivational?!! (In my opinion as an ITO graduate!) I don't think the answer is even necessarily have more dedicated Web Science courses – but I think Web Science can be utilised within other subjects.. This includes non-technology based subjects, such as banking and finance, health and education into perspective – I think it is positive to see the bigger picture; all these disciplines are now significantly impacted by the Web in some way, and often use the Web as a platform through which to operate effectively. Therefore, it can surely be nothing but beneficial to understand the impact of the Web and how it works and impacts any given discipline.

16.4 World Map of Web Science



World map found at: http://www.psdmore.com/psd-files/02/blank-world-map-download-psd-file.html

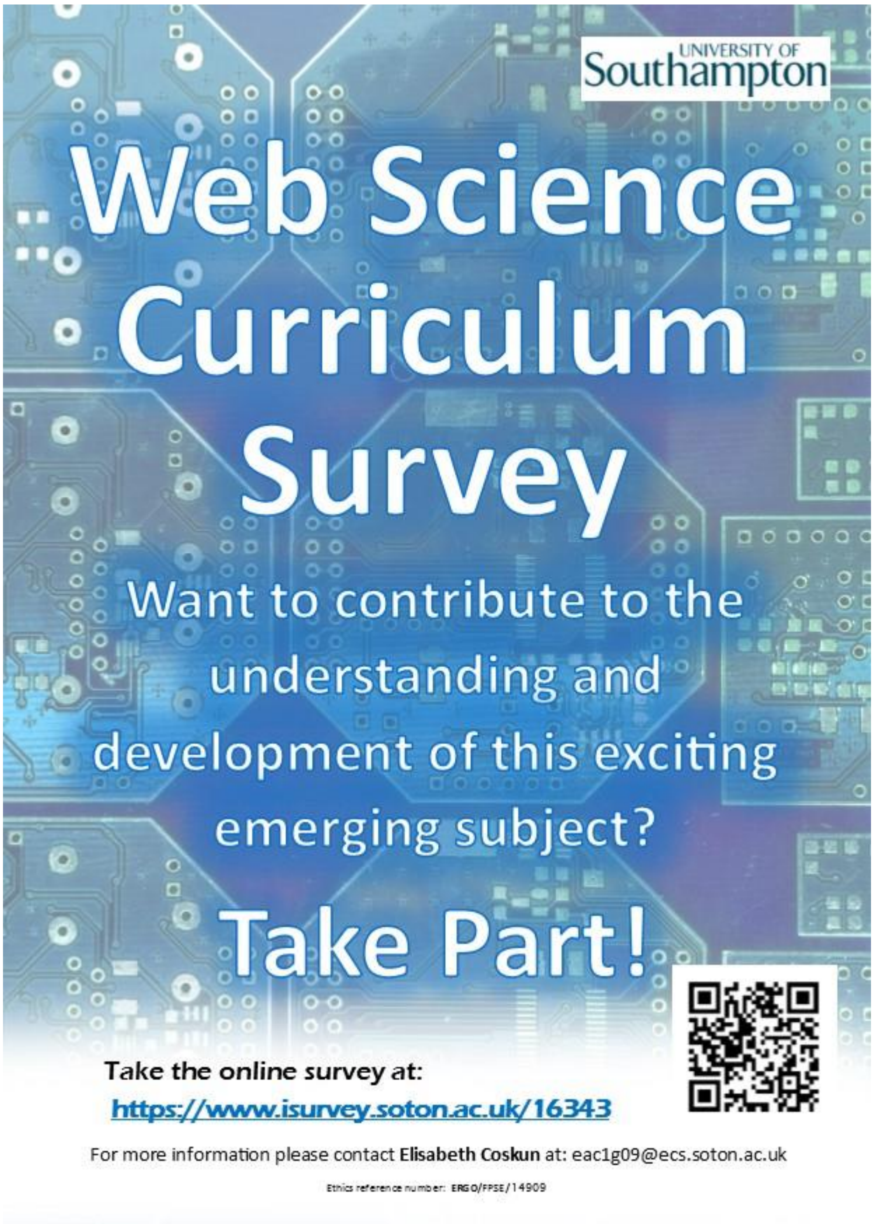
16.5 Top 100 Module Keywords

Word	Length	Count	Weighted Percentage (%)
web	3	248	3.89
social	6	94	1.47
data	4	69	1.08
networks	8	60	0.94
course	6	50	0.78
information	11	47	0.74
science	7	43	0.67
network	7	42	0.66
research	8	42	0.66
design	6	41	0.64
systems	7	41	0.64
analysis	8	38	0.60
media	5	38	0.60
new	3	38	0.60
understanding	13	38	0.60
technologies	12	34	0.53
models	6	33	0.52
applications	12	31	0.49
computing	9	31	0.49
students	8	31	0.49
issues	6	30	0.47
techniques	10	30	0.47
development	11	29	0.45
methods	7	29	0.45
semantic	8	28	0.44
technology	10	28	0.44
digital	7	27	0.42
internet	8	27	0.42
module	6	27	0.42
software	8	26	0.41
introduction	12	24	0.38
based	5	22	0.34
knowledge	9	22	0.34
language	8	22	0.34
management	10	22	0.34
computer	8	21	0.33
engineering	11	20	0.31
languages	9	20	0.31
services	8	20	0.31
society	7	20	0.31
world	5	20	0.31
approaches	10	19	0.30
retrieval	9	19	0.30
security	8	19	0.30
tools	5	19	0.30
user	4	19	0.30
online	6	18	0.28
process	7	18	0.28
search	6	18	0.28
concepts	8	17	0.27

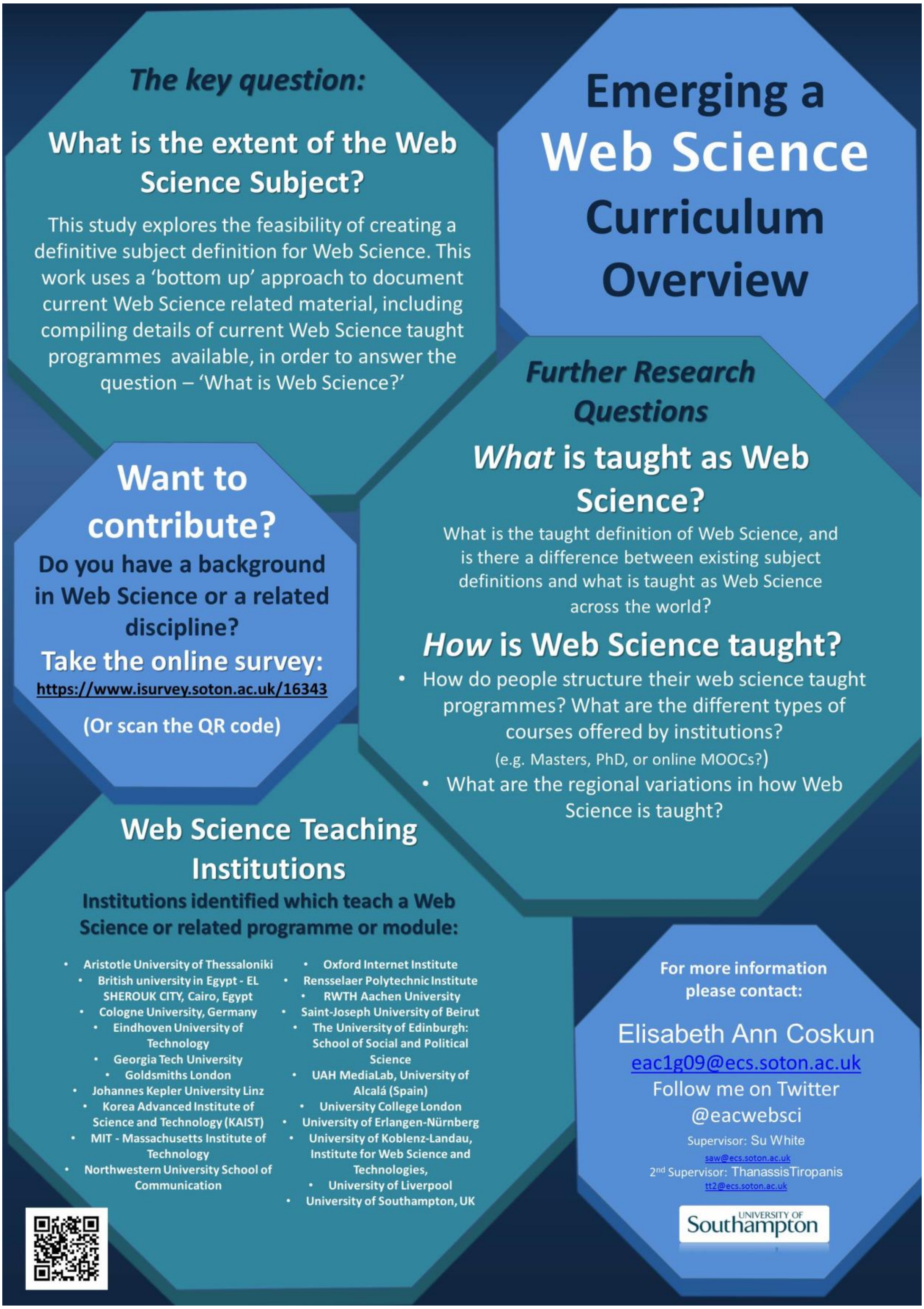
Word	Length	Count	Weighted Percentage (%)
virtual	7	17	0.27
also	4	16	0.25
big	3	16	0.25
understand	10	16	0.25
architectures	13	15	0.23
business	8	15	0.23
context	7	15	0.23
principles	10	15	0.23
algorithms	10	14	0.22
architecture	12	14	0.22
environments	12	14	0.22
including	9	14	0.22
innovation	10	14	0.22
interaction	11	14	0.22
mobile	6	14	0.22
programming	11	14	0.22
quality	7	14	0.22
time	4	14	0.22
application	11	13	0.20
content	7	13	0.20
performance	11	13	0.20
technical	9	13	0.20
used	4	13	0.20
using	5	13	0.20
well	4	13	0.20
approach	8	12	0.19
communication	13	12	0.19
future	6	12	0.19
mining	6	12	0.19
professional	12	12	0.19
theory	6	12	0.19
topics	6	12	0.19
areas	5	11	0.17
critical	8	11	0.17
impact	6	11	0.17
intelligence	12	11	0.17
marketing	9	11	0.17
participation	13	11	0.17
processes	9	11	0.17
professionals	13	11	0.17
protocols	9	11	0.17
related	7	11	0.17
structures	10	11	0.17
study	5	11	0.17
analyse	7	10	0.16
behaviour	9	10	0.16
concept	7	10	0.16
develop	7	10	0.16
human	5	10	0.16
images	6	10	0.16

16.6 Promotional Flyer

This flyer was designed by the author to promote the survey.



16.7 Promotional project poster



16.8 Ethics Documentation

Table of Contents

Refer to the *Instructions* and to the *Guide* documents for a glossary of the key phrases in **bold** and for an explanation of the information required in each section. The *Templates* document provides some text that may be helpful in presenting some of the required information.

Replace the highlighted text with the appropriate information.

Note that the size of the text entry boxes provided on this form does **not** indicate the expected amount of information; instead, refer to the *Instructions* and to the *Guide* documents in providing the complete information required in each section. Do **not** duplicate information from one text box to another.

ETHICS APPLICATION FORM

<table>
<tr><td>Reference number: ERGO/Error! Unknown document property name./14909</td><td>Version: 2</td><td>Date: 2015-06-02</td></tr>
<tr><td colspan="3">Name of investigator(s):
Elisabeth Coskun</td></tr>
<tr><td colspan="3">Name of supervisor(s) (if student investigator(s)):
Su White</td></tr>
<tr><td colspan="3">Title of study: Web Science Curriculum Survey</td></tr>
<tr><td>Expected study start date:
15/06/15</td><td colspan="2">Expected study end date: 30/09/17</td></tr>
<tr><td colspan="3">Note that the dates requested on the “IRGA” form refer to the start and end of data collection. These are not the same as the start and end dates of the study for which approval is sought.
Note that approval must be obtained before the study commences; retrospective approval cannot be given.</td></tr>
</table>

The investigator(s) undertake to:

- Ensure the study Reference number ERGO/**Error! Unknown document property name.**/14909 is prominently displayed on all advertising and study materials, and is reported on all media and in all publications;
- Conduct the study in accordance with the information provided in the application, its appendices, and any other documents submitted;
- Conduct the study in accordance with University policy governing research involving human

participants (http://www.southampton.ac.uk/ris/policies/ethics.html);

- Conduct the study in accordance with University policy on data retention (http://www.southampton.ac.uk/library/research/researchdata/);
- Submit the study for re-review (as an amendment through ERGO) or seek **Error! Unknown document property name.** EC advice if any changes, circumstances, or outcomes materially affect the study or the information given;
- Promptly advise an appropriate authority (Research Governance Office) of any adverse study outcomes, changes, or circumstances (via an adverse event notification through ERGO);
- Submit an end-of-study form as may be required by the Research Governance Office upon completion of the study.

REFER TO THE <u>INSTRUCTIONS</u> DOCUMENT WHEN COMPLETING THIS FORM.

PRE-STUDY

Characterise the proposed **participants**
Participants are largely students and academics within the international Web Science community. A desk survey identified nineteen institutions worldwide which teach either a programme or module relating to Web Science; the academics and students involved in these international curricula are they key target audience. A sample of these academics and students are expected to attend the upcoming 2015 Web Science conference. Attendees of this event will form a part of the sample. Members of social media groups dedicated to Web Science will also form a part of the sample. Another section of the sample will include the WAIS department within the University of Southampton, which may include some of the author's colleagues. The study may also be widened to include other disciplines related to Web Science, including the international Digital Sociology community.

Describe how **participants** will be approached
Participants will include international delegates of the 2015 Web Science conference, who will be largely unknown to and have no personal connection to the researcher. They will be approached by means of leaflets and word of mouth. The survey will also be promoted via social media, using a number of social media groups. A desk survey also identified number of institutions and some academics who might be contacted via email or phone.

Describe how inclusion and/or exclusion criteria will be applied (if any)
All participants who choose to answer the questionnaire will be included in the study.

Describe how **participants** will decide whether to take part
Before they begin the survey, participants will be asked to tick a consent box, stating that they agree to participate, and the length of time that the survey is likely to take. (No longer than 20 mins) Survey participants will give informed consent before taking part in the survey. Participants will be provided with a participant information sheet which will outline details of the study, including addressing potential concerns. It will explain that and will have the right to withdraw their information from the study at any time.

Participant Information

Provide the **Participant Information** in the form that it will be given to **participants** as an appendix. All studies must provide **participant information**.

Consent Form

Provide the **Consent Form** (or the request for consent) in the form that it will be given to **participants** as an appendix. All studies must obtain **participant** consent. Some studies may obtain verbal consent, other studies will require written consent, as explained in the *Instructions* and *Guide* documents.

DURING THE STUDY

Describe the study procedures as they will be experienced by the **participant**
Learners will be invited to take part in the study, either by word of mouth, leaflet or email. The invitation will provide participants with a link to the online survey, which includes 3 sections of questions. After answering questions in each section, participants may revisit a previous section and revise their answers if they wish, before submitting their final answers. A paper version of the same survey may also be printed for ease of distribution at events such as the Web Science conference.

Identify how, when, where, and what kind of data will be recorded (not just the formal research data, but including all other study data such as e-mail addresses and signed consent forms)
Data will collected by the iSurvey. All answers will be anonymised and only answers will be recorded for the study. The survey will include optional fields for participants to provide their names and email addresses. All other fields in the survey relate to gathering information the participant's academic institution's Web Science programme, as well as the participant's opinions on the development of the Web Science subject. Should a paper version of the survey be created, the completed papers will be stored in a secure location, and destroyed once the study is complete.

Participant questionnaire

As an appendix, if using a questionnaire, reproduce any and all **participant** questionnaires or data gathering instruments in the exact forms that they will be given to or experienced by **participants**. If conducting less formal data collection, provide specific information concerning the methods that will be used to obtain the required data.

POST-STUDY

Identify how, when, and where data will be stored, processed, and destroyed
Data will be stored on the investigator's personal page on the iSurvey website. Data will be retained until the end of the investigator's PhD (2017), after which time the data will be destroyed.

STUDY CHARACTERISTICS

(L.1) The study is funded by a commercial organisation: **No** (delete or highlight one) If 'Yes', provide details of the funder or funding agency here

(L.2) There are **restrictions** upon the study: **No** (delete or highlight one) If 'Yes', explain the nature and necessity of the **restrictions** here

(L.3) Access to **participants** is through a third party: **No** (delete or highlight one) If 'Yes', provide evidence of your permission to contact them as a separate appendix. Do not provide explanation or information on this matter here

(M.1) **Personal data** is collected or processed: **Yes** (delete or highlight one) Data will be processed outside the UK: **No** (delete or highlight one) If 'Yes' to either question, provide the **DPA Plan** as a separate appendix. Do not provide information or explanation on this matter here. Note that using or retaining e-mail addresses, signed consent forms, or similar study-related **personal data** requires M.1 to be "Yes"

(M.2) There is **inducement** to **participants**: **No** (delete or highlight one) If 'Yes', explain the nature and necessity of the inducement here

(M.3) The study is **intrusive**: **No** (delete or highlight one) If 'Yes', provide the **Risk Management Plan** and the **Debrief Plan** as appendices, and explain here the nature and necessity of the intrusion(s)

(M.4) There is **risk of harm** during the study: **No** (delete or highlight one) If 'Yes', provide the **Risk Management Plan**, the **Contact Information**, and the **Debrief Plan** as appendices, and explain here the necessity of the risks

(M.5) The true purpose of the study will be hidden from **participants**: **No** (doho) The study involves **deception** of **participants**: **No** (delete or highlight one) If 'Yes' to either question, provide the **Debrief Plan** as an appendix, and explain here the necessity of the deception

(M.6) **Participants** may be minors or otherwise have **diminished capacity**: **No** (doho) If 'Yes', AND if one or more Study Characteristics in categories M or H applies, provide the **Risk Management Plan** and the **Contact Information**, as appendices, and explain here the special arrangements that will be put in place that will ensure informed consent

(M.7) **Sensitive data** is collected or processed **No** (delete or highlight one) If 'Yes', provide the **DPA Plan** as a separate appendix. Do not provide explanation or information on this matter here

(H.1) The study involves: **invasive** equipment, material(s), or process(es); or **participants** who are not able to withdraw at any time and for any reason; or animals; or human tissue; or biological samples: **No** (delete or highlight one)

If 'Yes', provide further details and justifications as one or more separate appendices. Do not provide explanation or information on these matters here. Note that the study will require separate approval by the Research Governance Office

Technical details

If one or more Study Characteristics in categories M.3 to M.7 or H applies, provide the description of the technical details of the experimental or study design, the power calculation(s) which yield the required sample size(s), and how the data will be analysed, as separate appendices. Do not provide explanation or information on these matters here.

APPENDICES (AS REQUIRED)

While it is preferred that this information is included here in the Study Protocol document, it may be provided as separate documents.

If provided separately, be sure to name the files precisely as "Participant Information", "Questionnaire", "Consent Form", "DPA Plan", "Permission to contact", "Risk Management Plan", "Debrief Plan", "Contact Information", and/or "Technical details" as appropriate.

If provided separately, each document must specify the reference number in the form ERGO/**Error! Unknown document property name.**/xxxx, its version number, and its date of last edit.

Appendix (i): **Participant Information** in the form that it will be given to **participants.**

Appendix (ii): Data collection plan / Questionnaire in the form that it will be given to **participants.**

Appendix (iii): **Consent Form** in the form that it will be given to **participants.**

Appendix (iv): **DPA Plan.**

Appendix (v): Evidence of permission to contact **participants** or prospective **participants** through any third party.

Appendix (vi): **Risk Management Plan**.

Appendix (vii): **Debrief Plan**.

Appendix (viii): **Contact Information**.

Appendix (ix): Technical details of the experimental or study design, the power calculation(s) for the required sample size(s), and how the data will be analysed.

Appendix (x): Further details and justifications in the case of **invasive** equipment, material(s), or process(es); **participants** who are not able to withdraw at any time and for any reason; animals; human tissue; or biological samples.

Appendix (i): Participant Information

Participant Information Sheet

Study Title: Web Science Curriculum Survey

Researcher: Elisabeth Coskun

Ethics number: 14909

Version: 2

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

By contributing to this study, you are aiding in the PhD research project of a student studying the Web Science curriculum. The ultimate aim of the project is to gain a clearer understanding of this emerging subject, in order to provide resource(s) which help communicate the scope of Web Science to students and educators interested in the subject, and also to potentially provide a means of tracking the development of the subject. Key information required by the study, includes identifying details relating to what is currently taught as part of active Web Science taught curriculums, and also how Web Science is currently taught, and which institutions are teaching it. This will provide valuable insight into the development of the Web Science subject, as well as a point of reference for tracking the continued development of the discipline.

Why have I been chosen?

You have been approached for your contribution to this study because you are either teaching, studying or working in a field relevant to Web Science. Your input will provide valuable insight into the development of Web Science subject.

What will happen to me if I take part?

You will be agreeing to contribute around 15 minutes of your time to answer a series of survey questions relating to your involvement in Web Science. For the questions which allow text field responses, please provide as much detail as possible, as all feedback will provide the study with valuable insight. If you agree to provide your email address, you may be contacted by the author of the study for additional information about your involvement with Web Science, and your opinions relating to the development of the subject.

Are there any benefits in my taking part?

By contributing to this study, you are helping us to better understand this exciting new discipline, and also potentially provide a better definition which may help future students and educators in their understanding and teaching of the subject.

Will my participation be confidential?

This study complies with the Data Protection Act and University policy. Personal information such as names and e-mail addresses will be stored securely and will remain confidential to the author and her supervisor.

What happens if I change my mind?

You retain the right to withdraw from this study at any time without your legal rights being affected.

What happens if something goes wrong?

Should you have any concerns or complaints, please contact: Research Governance Manager (02380 595058, rgoinfo@soton.ac.uk)

Where can I get more information?

Should you wish to find out more about this study, please contact the author at: eac1g09@ecs.soton.ac.uk

APPENDIX (IA): PARTICIPANT INFORMATION AND CONSENT (ONLINE SURVEY VERSION)

PARTICIPANT INFORMATION SHEET

Please read this information carefully before deciding to take part in this research. If you are happy to participate, please tick the consent box.

What is the research about?

By contributing to this study, you are aiding in the PhD research project of a student studying the Web Science curriculum. The ultimate aim of the project is to gain a clearer understanding of this emerging subject, in order to provide resource(s) which help communicate the scope of Web Science to students and educators interested in the subject, and also to potentially provide a means of tracking the development of the subject. Key information required by the study, includes identifying details relating to what is currently taught as part of active Web Science taught curriculums, and also how Web Science is currently taught, and which institutions are teaching it. This will provide valuable insight into the development of the Web Science subject, as well as a point of reference for tracking the continued development of the discipline.

Why have I been chosen?

You have been approached for your contribution to this study because you are either teaching, studying or working in a field relevant to Web Science. Your input will provide valuable insight into the development of Web Science subject.

What will happen to me if I take part?

You will be agreeing to contribute around 15 minutes of your time to answer a series of survey questions relating to your involvement in Web Science. For the questions which allow text field responses, please provide as much detail as possible, as all feedback will provide the study with valuable insight. If you agree to provide your email address, you may be contacted by the author of the study for additional information about your involvement with Web Science, and your opinions relating to the development of the subject.

Are there any benefits in my taking part?

By contributing to this study, you are helping us to better understand this exciting new discipline, and also potentially provide a better definition which may help future students and educators in their understanding and teaching of the subject.

Will my participation be confidential?

This study complies with the Data Protection Act and University policy. Personal information such as names and e-mail addresses will be stored securely and will remain confidential to the author and her supervisor.

What happens if I change my mind?

You retain the right to withdraw from this study at any time without your legal rights being affected.

What happens if something goes wrong?

Should you have any concerns or complaints, please contact: Research Governance Manager (02380 595058, rgoinfo@soton.ac.uk)

Where can I get more information?

Should you wish to find out more about this study, please contact the author at: eac1g09@ecs.soton.ac.uk

Data Protection

I understand that information collected during my participation in this study is completely anonymous / will be stored on a password protected computer/secure University server and that this information will only be used in accordance with the Data Protection Act (1998). The DPA (1998) requires data to be processed fairly and lawfully in accordance with the rights of participants and protected by appropriate security.

I agree to take part in this study.

I understand my participation is voluntary and I may withdraw at any time and for any reason.

Please tick (check) this box to indicate that you consent to taking part in this survey ☐

16.9 Questionnaire

Appendix (II): Questionnaire

Web Science Curriculum Survey 2015/16

1. Section 1: Your Educational Background

Question 1.

Please state whether you are a student, academic or other:

If 'Other' please give more details:

Question 2.

Your name (Optional)

Question 3.

Would you be happy for us to contact you for further information relating to this research? If yes, please provide us with your email address: (Optional)

Question 4.

At what level do you teach or study?

- Undergraduate
- Masters
- PhD
- Other

Question 5.

Which field(s) best represent your educational or work background? Please choose all that apply:

- ☐ Arts
- ☐ Artificial Intelligence
- ☐ Computer Science
- ☐ Ecology
- ☐ Economics
- ☐ Education
- ☐ Electronic Engineering
- ☐ Geography
- ☐ Health
- ☐ Information Technology
- ☐ Languages
- ☐ Law
- ☐ Mathematics
- ☐ Media
- ☐ Physics
- ☐ Political Science
- ☐ Psychology

- [ ] Socio-cultural
- [ ] Sociology (Digital Sociology)
- [ ] Web Engineering
- [ ] Web Science
- [ ] Other

Question 6.

Please specify your degree title, research field or job title:

Question 7.

Are you aware of the term "Web Science"?

Yes No

Question 8.

If you have a background in Web Science, please explain any significant contributions you have made to the discipline:

2. Section: Your Curriculum Details

This section of the survey aims to gather data relating to teaching institutions and resources teaching Web Science (or related subjects, e.g. Digital Sociology). The aim of this is to gain an overview of the most commonly taught subjects within Web Science, as well as how the subject is taught. (If you are not associated with an academic institution, please answer questions 2.1 and 2.2, before skipping to section 3 of the survey.)

Question 1.

What is the name of your academic institution or company?

Question 2.

In which part of the world is your institution located?

Question 3.

Please give the name(s) of the Web Science (or related) taught programmes offered by your institution:

Question 4.

What level(s) of Web Science (or related) qualification are taught by your academic institution? (Please tick all that apply)

Undergraduate Masters PhD Other

If other, please describe:

Question 5.

Please provide a summary of the key topics covered by your institution's Web Science (or related) programme(s): (If this is a full degree programme, please provide module titles if possible)

Question 6.

Please list the language(s) in which your programme is taught:

Question 7.

Please provide an estimate of the number of teaching hours within your Web Science (or related) taught programme:

Question 8.

What is the duration of the programme? (Please give dates if possible)

Question 9.

What date was your programme first created? (Optional)

Question 10.

In what way(s) is your programme taught? (Lectures, online material, etc.)

Question 11.

Please provide a link to the programme website: (Optional)

3. Section 3: Curriculum Subjects

The following sections list some key topics, as described by the Web Science Subject Categorisation. (WSSC) How would you rate these subjects in terms of relevance to your institution's Web Science (or related subject) curriculum? (If you are not based within academia, how relevant are these topics to your company's area of expertise?)

Question 1.

Please rate the following categories relating to: WEB TECHNOLOGIES

	not at all relevant	not very relevant	unsure	relevant	very relevant
General Web Technologies	◯	◯	◯	◯	◯
Web Milieux: Document technologies; Hypertext technologies; Internet technologies; Mobile Web technologies; Grid and Cloud computing technologies	◯	◯	◯	◯	◯
Basic Web Architecture: HTTP and related technologies; URIs; HTML; XML; CSS and related technologies; Interfaces and	◯	◯	◯	◯	◯

Browsers; Servers Web Services					
Web 2.0 technologies					
Semantic Web/Linked Data: Metadata; Knowledge Representation; Ontology Languages; Linked Data; Natural Language Processing; Provenance systems in the Web; Other in Semantic Web/Linked Data					
Internet/Web of Things					

Question 2.

Please rate the following categories relating to: WEB ANALYSIS

	not at all relevant	not very relevant	unsure	relevant	very relevant
General Web Analysis	○	○	○	○	○
Mathematical Methods of Web analysis: Web data sampling and analytics; Logic and Inference in the Web; Statistical Inference in the Web; Statistical Analysis of the Web; Web as a Complex System; Graphs; Networks; Mathematical methods for describing Web services; Crawling; Indexing and Searching; Data Mining; Information Retrieval and Machine Learning; Other	○	○	○	○	○

Algorithms for the Web

Question 3.

Please rate the following categories relating to: WEB SOCIETY

	not at all relevant	not very relevant	unsure	relevant	very relevant
Economics and Business: Goods in the Web; The Web economy; Antitrust Issues and Policies in the Web; Intellectual property and digital rights management; Web-based economic development	○	○	○	○	○
Economics and Business (Business) E-commerce Business models in the Web; Advertising in the Web; sponsored search	○	○	○	○	○
Social Engagement and Social Science: Social networks; Mass phenomena; Collective	○	○	○	○	○

intelligence; Peer production; Globalization; Systems; Social structures and processes; Virtual communities, groups and identity; Social capital and power inequality in the Web; On-line lives, intergenerational differences; Mass media					
Personal Engagement and Psychology System Psychology and Behaviour; Child and adolescent psychiatry; Tele-working	○	○	○	○	○
Philosophy Philosophy of information; Objects; Reference and Cognition in the Web; Ethics in the Web	○	○	○	○	○
Law:Intellectual Property in the Web; Digital Rights Management;	○	○	○	○	○

Digital crime; Laws for Web access; Antitrust Law

Politics and Governance: Political science; E-Government; E-Politics; E-Democracy; Policy and Regulation; Web Governance; Privacy; Trust; Security ;Network neutrality; E-Inclusion

Question 4.

Do you feel that the above framework of the WSSC omits any key subjects which relate to Web Science or your individual curriculum? If yes, please list subjects and why you feel they are relevant:

Question 5.

Are you aware of any differences between the terms ‘Web Science’ and ‘Internet Science’? If yes, please elaborate:

Appendix (III): Consent Form

Consent Form

Ethics reference number: **ERGO/Error! Unknown document property name./**14909	Version: 2	Date: 2015-6-2
Study Title: Web Science Curriculum Survey 2015/16		
Investigator: Elisabeth Coskun		

Please initial the box(es) if you agree with the statement(s):

I have read and understood the Participant Information (version 1 dated 2015-6-2) and have had the opportunity to ask questions about the study.	☐
I agree to take part in this study.	☐
I understand my participation is voluntary and I may withdraw at any time and for any reason.	☐

Data Protection

I understand that information collected during my participation in this study is completely anonymous / will be stored on a password protected computer/secure University server and that this information will only be used in accordance with the Data Protection Act (1998). The DPA (1998) requires data to be processed fairly and lawfully in accordance with the rights of participants and protected by appropriate security.

Name of participant (print name)..

Signature of participant..

Date...

APPENDIX (IV): DPA PLAN

16.10 DPA Plan

Ethics reference number: **ERGO/Error! Unknown document property name./**14909	Version: 2	Date: 2015-06-17
Study Title: Web Science Curriculum Survey		
Investigator: Elisabeth Coskun Name of supervisor(s) (if student **investigator**(s)): Su White		

The questionnaire shown in appendix II of this document provides the complete list of all the data that will be collected.

Data Relevance: The data is relevant to the study purposes because it will gather data relating to the Web Science curriculum development. The data is adequate because it adequately identifies participants without gathering unnecessary personal data, and it also asks a range of carefully chosen questions relating to the curriculum and the participants views of the curriculum. The data is not excessive because it includes no unnecessary personal data to identify the participant. The only personal data requested by the study are the names and e-mail addresses of participants. These fields are both optional.

Data Accuracy: The majority of the information required by the study is subjective and relates to the knowledge and opinion of the participant and their level of involvement within Web Science. Therefore the accuracy of the information recorded depends on the knowledge of each participant.

Data Retention: The data will be archived on ePrints by the investigators at the end of the research period (end of September 2017) in accordance to university policy http://library.soton.ac.uk/researchdata Physical data will be shredded after it has been scanned and added to the collection of data on the researchers' computers. Personal data such as names and emails will be separated from the rest of the data and anonymized for data analysis purposes, and replaced with placeholders such as 'participant 1'.

Data Processing: The data will be processed fairly, and the participants will have given explicit consent to take part in the study. The data will be processed in accordance with the rights of the participants. Rights: Participants will have the right to access, correct, and/or withdraw their data at any time and for any reason. Participants will be able to exercise their rights by contacting the investigators (e-mail: eac1g09@ecs.soton.ac.uk) or the project supervisor (e-mail: saw@ecs.soton.ac.uk).

Consent forms will be attached to the survey, and participants will be required to fill these out before they complete the attached survey.

The following data will not be transferred outside the European Economic Area (EEA).

17 WAIS FEST 2016 ebook Interviews: Ethics Application

Table of Contents

Refer to the *Instructions* and to the *Guide* documents for a glossary of the key phrases in **bold** and for an explanation of the information required in each section. The *Templates* document provides some text that may be helpful in presenting some of the required information.

Replace the highlighted text with the appropriate information.

Note that the size of the text entry boxes provided on this form does **not** indicate the expected amount of information; instead, refer to the *Instructions* and to the *Guide* documents in providing the complete information required in each section. Do **not** duplicate information from one text box to another.

ETHICS APPLICATION FORM

Reference number: **ERGO/Error! Unknown document property name./22953**	Version: 3	Date: 23rd Aug 2016
Name of **investigator**(s): Elisabeth Coskun		
Name of supervisor(s) (if student **investigator**(s)): Su White		
Title of study: Crowdsourcing a Web Science Primer		
Expected study start date: 24/08/16	Expected study end date: 31/12/17	
Note that the dates requested on the "IRGA" form refer to the start and end of *data collection*. These are not the same as the start and end dates of the study for which approval is sought. Note that approval must be obtained before the study commences; retrospective approval cannot be given.		

The investigator(s) undertake to:

- Ensure the study Reference number ERGO/**Error! Unknown document property name./22953** is prominently displayed on all advertising and study materials, and is reported on all media and in all publications;
- Conduct the study in accordance with the information provided in the application, its appendices, and any other documents submitted;
- Conduct the study in accordance with University policy governing research involving human

participants (http://www.southampton.ac.uk/ris/policies/ethics.html);

- Conduct the study in accordance with University policy on data retention (http://www.southampton.ac.uk/library/research/researchdata/);
- Submit the study for re-review (as an amendment through ERGO) or seek **Error! Unknown document property name.** EC advice if any changes, circumstances, or outcomes materially affect the study or the information given;
- Promptly advise an appropriate authority (Research Governance Office) of any adverse study outcomes, changes, or circumstances (via an adverse event notification through ERGO);
- Submit an end-of-study form as may be required by the Research Governance Office upon completion of the study.

Refer to the Instructions document when completing this form.

Pre-study

Characterise the proposed **participants**
Participants are students and academics within the (WAIS) Web and Internet Science department at the University of Southampton. These individuals will have completed the Web Science MSc Programme, and will now be either studying for a Web Science PhD, or will be a lecturer or research fellow within the department. Some of these individuals may include the author's colleagues. This study seeks to identify the current research topics of the individuals, and gain an insight into how their prior experiences of the Web Science MSc programme (or other MSc background) have impacted and influenced the current PhD studies of the participants involved. The study also seeks to gain an understanding of student's opinions of the MSc Programme, including the knowledge, skills and understanding that they gained from studying the MSc programme, as well as areas for improvement.

Describe how **participants** will be approached
Participants will all be members of the WAIS department within the University of Southampton. The initial format of the study will be an online survey, which may be distributed to the department via email and social media groups. There will also be a paper version of the survey, which may be distributed during the department's annual research week (known as 'WAIS Fest'). The survey may also be used as a template for one-to-one interviews with participants, which may be arranged either by e-mail or in person.

Describe how inclusion and/or exclusion criteria will be applied (if any)
It is explained to participants, that by agreeing to participate in the study, they are agreeing to contribute their information to be published within a Web Science ebook. If they do not consent to this, their information will not be used in a case study. Some questions in sections 2 and 3 are listed as 'optional;'. This allows participants to not answer if they are not comfortable with providing the information. For example, some people may not be happy to say what they didn't like about the Web Science MSc and have that information published. Questions 1 of section 3 (Participant's name) is optional, but it is stated that on the survey that it is required if participants wish to contribute to the Web Science ebook, as the information provided will be attributed to them. Respondents who participate in face to face interviews will be asked if they consent to the interview being recorded for transcription purposes. If they do not consent to this, the interview will not be recorded. However, if this leads to missing or incomplete information, then participants who do not allow recorded interviews may not be included in the final case studies published.

Describe how **participants** will decide whether to take part
Before they begin the survey or interview, participants will be asked to tick a consent box, stating that they agree to participate, and the length of time that the survey or interview is likely to take. (No longer than 20 mins) Participants will give informed consent before taking part in the survey. Participants will be provided with a participant information sheet which will outline details of the study, including addressing potential concerns. It is explained to participants, that by agreeing to participate in the study, they are agreeing to contribute their information to be published within a Web Science ebook. If they do not consent to this, their information will not be used in a case study. It will also explain that and will have the right to withdraw their information from the study at any time.

Participant Information

Provide the **Participant Information** in the form that it will be given to **participants** as an appendix. All studies must provide **participant information**.

Consent Form

Provide the **Consent Form** (or the request for consent) in the form that it will be given to **participants** as an appendix. All studies must obtain **participant** consent. Some studies may obtain verbal consent, other studies will require written consent, as explained in the *Instructions* and *Guide* documents.

During the study

Describe the study procedures as they will be experienced by the **participant**
Learners will be invited to take part in the study, either by word of mouth or email. Participants will be provided with a consent form, and invited to answer the questions, which should take no longer than 20 mins. The majority of the surveys will be conducted in interview format, in which the investigator would be the interviewer and would record the answers in the relevant survey sections. Interviews will be sound recorded for transcription purposes only, and each participant will be asked if they consent to this. Should any participants not be reachable for an interview, they will be sent an email invitation to complete an online version of the survey.

Identify how, when, where, and what kind of data will be recorded (not just the formal research data, but including all other study data such as e-mail addresses and signed consent forms)
Survey and interview data will collected within iSurvey. Should a paper version of the survey be created, the completed papers will be stored in a secure location, and destroyed once the study is

complete. The survey will include optional fields for participants to provide their names and email addresses, and personal website addresses. Other fields in the survey relate to gathering information the participant's research topic and their experiences of the Web Science MSc. Interviews will be recorded using a dedicated diaphone device, having obtained consent from each individual participant. The sole purpose of audio recordings will be to aid transcription of the interviews, should any information be missed during the actual interview.

Participant questionnaire

As an appendix, if using a questionnaire, reproduce any and all **participant** questionnaires or data gathering instruments in the exact forms that they will be given to or experienced by **participants**. If conducting less formal data collection, provide specific information concerning the methods that will be used to obtain the required data.

Post-study

Identify how, when, and where data will be stored, processed, and destroyed
Data will be stored on the investigator's personal page on the iSurvey website. Sound recordings of the interviews will be made using a dedicated Dictaphone device and will be stored on this device, which will be kept in a secure location, and on a university desktop computer. Data will be retained until the end of the investigator's PhD (Dec 2017), after which time the data will be destroyed.

Study characteristics

(L.1) The study is funded by a commercial organisation: **No** (delete or highlight one) If 'Yes', provide details of the funder or funding agency here

(L.2) There are **restrictions** upon the study: **No** (delete or highlight one) If 'Yes', explain the nature and necessity of the **restrictions** here

(L.3) Access to **participants** is through a third party: **No** (delete or highlight one)

If 'Yes', provide evidence of your permission to contact them as a separate appendix. Do not provide explanation or information on this matter here

(M.1) **Personal data** is collected or processed: **Yes** (delete or highlight one)
Data will be processed outside the UK: **No** (delete or highlight one)

If 'Yes' to either question, provide the **DPA Plan** as a separate appendix. Do not provide information or explanation on this matter here. Note that using or retaining e-mail addresses, signed consent forms, or similar study-related **personal data** requires M.1 to be "Yes"

(M.2) There is **inducement** to **participants**: **No** (delete or highlight one)

If 'Yes', explain the nature and necessity of the inducement here

(M.3) The study is **intrusive**: **No** (delete or highlight one)

If 'Yes', provide the **Risk Management Plan** and the **Debrief Plan** as appendices, and explain here the nature and necessity of the intrusion(s)

(M.4) There is **risk of harm** during the study: **No** (delete or highlight one)

If 'Yes', provide the **Risk Management Plan**, the **Contact Information**, and the **Debrief Plan** as appendices, and explain here the necessity of the risks

(M.5) The true purpose of the study will be hidden from **participants**: **No** (doho)
The study involves **deception** of **participants**: **No** (delete or highlight one)

If 'Yes' to either question, provide the **Debrief Plan** as an appendix, and explain here the necessity of the deception

(M.6) **Participants** may be minors or otherwise have **diminished capacity**: **No** (doho) If 'Yes', AND if one or more Study Characteristics in categories M or H applies, provide the **Risk Management Plan** and the **Contact Information**, as appendices, and explain here the special arrangements that will be put in place that will ensure informed consent

(M.7) **Sensitive data** is collected or processed **No** (delete or highlight one) If 'Yes', provide the **DPA Plan** as a separate appendix. Do not provide explanation or information on this matter here

(H.1) The study involves: **invasive** equipment, material(s), or process(es); or **participants** who are not able to withdraw at any time and for any reason; or animals; or human tissue; or biological samples: **No** (delete or highlight one) If 'Yes', provide further details and justifications as one or more separate appendices. Do not provide explanation or information on these matters here. Note that the study will require separate approval by the Research Governance Office

Technical details

If one or more Study Characteristics in categories M.3 to M.7 or H applies, provide the description of the technical details of the experimental or study design, the power calculation(s) which yield the required sample size(s), and how the data will be analysed, as separate appendices. Do not provide explanation or information on these matters here.

Appendices (as required)

While it is preferred that this information is included here in the Study Protocol document, it may be provided as separate documents.

If provided separately, be sure to name the files precisely as "Participant Information", "Questionnaire", "Consent Form", "DPA Plan", "Permission to contact", "Risk Management Plan", "Debrief Plan", "Contact Information", and/or "Technical details" as appropriate.

If provided separately, each document must specify the reference number in the form ERGO/**Error! Unknown document property name.**/xxxx, its version number, and its date of last edit.

Appendix (i): **Participant Information** in the form that it will be given to **participants.**

Appendix (ii): Data collection plan / Questionnaire in the form that it will be given to **participants.**

Appendix (iii): **Consent Form** in the form that it will be given to **participants.**

Appendix (iv): **DPA Plan.**

Appendix (v): Evidence of permission to contact **participants** or prospective **participants** through any third party.

Appendix (vi): **Risk Management Plan**.

Appendix (vii): **Debrief Plan**.

Appendix (viii): **Contact Information**.

Appendix (ix): Technical details of the experimental or study design, the power calculation(s) for the required sample size(s), and how the data will be analysed.

Appendix (x): Further details and justifications in the case of **invasive** equipment, material(s), or process(es); **participants** who are not able to withdraw at any time and for any reason; animals; human tissue; or biological samples.

APPENDIX (I): PARTICIPANT INFORMATION

Participant Information Sheet

Study Title: Crowdsourcing a Web Science Primer

Researcher: Elisabeth Coskun

Ethics number: **22953**

Version: 3

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

This study involves the production of a case study for each participant which will contribute to a Web Science ebook. Your name and email will be included in the case study in order to attribute your work to you, and is an opportunity for you to promote your work in the Web Science field.

What is the research about?

The aim of this project is to obtain an overview and insight into of the diverse range of PhD research projects currently being studied by students within the WAIS research group, in order to gain an understanding of the scope of Web Science, as studied by WAIS.

The study also seeks to obtain an insight into the knowledge, skills and understanding gained by students who have completed the University of Southampton Web Science MSc programme (or alternative MSc). In order to understand how student's MSc experiences have impacted their PhD studies.

The information gained during this study will aid in the compilation of a case study for each participant, which will contribute to a Web Science ebook, a resource outlining the scope of the Web Science discipline.

Why have I been chosen?

You have been approached for your contribution to this study because you are completing a PhD within the WAIS research group. You are also likely to have completed the Web Science MSc programme at the University of Southampton. The skills knowledge and understanding that you gained during this (or alternative MSc) and how it has contributed to you PhD research is the subject of this study.

What will happen to me if I take part?

You will be agreeing to contribute no more than 20 minutes of your time to answer a series of questions relating to your PhD research topic and your MSc experiences. Some of the information gathered during this study will be used to compile student cases studies, which will be included as part of a Web Science eBook.

You will be asked if you consent to the interview sound being recorded – this is for transcription purposes ONLY.

Are there any benefits in my taking part?

This is also an opportunity for you to promote your own work within the wider Web Science community, as a case study based summary of your work will be published as part of the proposed Web Science ebook. By contributing to this study, you are contributing to a Web Science eBook. This will help to inform students and educators about the scope and content of Web Science.

You are also providing valuable feedback which could help to improve the Southampton Web Science MSc for future students.

Will my participation be confidential?

This study complies with the Data Protection Act and University policy. By agreeing to take part you are agreeing that your feedback may be used to contribute to the eBook.

What happens if I change my mind?

You retain the right to withdraw from this study at any time without your legal rights being affected.

What happens if something goes wrong?

Should you have any concerns or complaints, please contact: Research Governance Manager (02380 595058, rgoinfo@soton.ac.uk)

Where can I get more information?

Should you wish to find out more about this study, please contact the author at: eac1g09@ecs.soton.ac.uk

APPENDIX (IA): PARTICIPANT INFORMATION AND CONSENT (ONLINE SURVEY VERSION)

Participant Information Sheet

Ethics number: **22953**

Version: 3

Please read this information carefully before deciding to take part in this research. If you are happy to participate, please tick the consent box.

This study involves the production of a case study for each participant which will contribute to a Web Science ebook. Your name and email will be included in the case study in order to attribute your work to you, and is an opportunity for you to promote your work in the Web Science field.

What is the research about?

The aim of this project is to obtain an overview and insight into of the diverse range of PhD research projects currently being studied by students within the WAIS research group, in order to gain an understanding of the scope of Web Science, as studied by WAIS.

The study also seeks to obtain an insight into the knowledge, skills and understanding gained by students who have completed the University of Southampton Web Science MSc programme (or alternative MSc). In order to understand how student's MSc experiences have impacted their PhD studies.

The information gained during this study will facilitate the compilation of a case study for each participant, which will contribute to a Web Science ebook, a resource outlining the scope of the Web Science discipline.

Why have I been chosen?

You have been approached for your contribution to this study because you are completing a PhD within the WAIS research group. You are also likely to have completed the Web Science MSc programme at the University of Southampton. The skills knowledge and understanding that you gained during this (or alternative MSc) and how it has contributed to you PhD research is the subject of this study.

What will happen to me if I take part?

You will be agreeing to contribute no more than 20 minutes of your time to answer a series of questions relating to your PhD research topic and your MSc experiences. Some of the information gathered during this study will be used to compile student cases studies, which will be included as part of a Web Science eBook.

Are there any benefits in my taking part?

By contributing to this study, you are contributing to a Web Science eBook. This will help to inform students and educators about the scope and content of Web Science. This is also an opportunity for you to promote your own work within the wider Web Science community, as a case study based summary of your work will be published as part of the proposed Web Science ebook.

You are also providing valuable feedback which could help to improve the Southampton Web Science MSc for future students.

Will my participation be confidential?

You will be agreeing to contribute no more than 20 minutes of your time to answer a series of questions relating to your PhD research topic and your MSc experiences. Some of the information gathered during this study will be used to compile student cases studies, which will be included as part of a Web Science eBook.

What happens if I change my mind?

You retain the right to withdraw from this study at any time without your legal rights being affected.

What happens if something goes wrong?

Should you have any concerns or complaints, please contact: Research Governance Manager (02380 595058, rgoinfo@soton.ac.uk)

Where can I get more information?

Should you wish to find out more about this study, please contact the author at: eac1g09@ecs.soton.ac.uk

Data Protection

I understand that information collected during my participation in this study is completely anonymous / will be stored on a password protected computer/secure University server and that this information will only be used in accordance with the Data Protection Act (1998). The DPA (1998) requires data to be processed fairly and lawfully in accordance with the rights of participants and protected by appropriate security.

17.1.1 I agree to take part in this study

17.1.2 I understand my participation is voluntary and I may withdraw at any time.

Please tick this box to indicate that you consent to taking part in this survey ☐

Appendix (II): Questionnaire

18 WAIS Fest Survey of Web Science MSc/PhD Students: Interview Questions

Crowdsourcing a Web Science Primer

18.11. eBook participation consent

This study involves the production of a case study for each participant which will contribute to a Web Science ebook. Your name, email and personal website link will be included in the case study in order to attribute your work to you, and is an opportunity for you to promote your work in the Web Science field.

Please indicate if you are happy to contribute to the ebook.

Question 1

Please select one of the following options:

- Yes I am happy to contribute to the ebook
- No I do not want to contribute to the ebook

Question 2

Please enter your name to indicate consent:

----

2. About your PhD

Please describe your PhD research project subject

____

Question 1.

What is your current year of PhD study?

Question 2.

What is your Research Project Title?

Question 3.

What keywords would you use to describe your research project?

If you were identifying your academic subject area or discipline/field of study to someone what words would you use? E.g. criminology, education, law

Question 4.

Please briefly describe the subject area of your research:

Question 5.

What are your research questions?

Question 6.

Please briefly describe your research methods?

Question 7.

How would you (briefly) describe the subject 'Web Science'?

WAIS Fest Survey of Web Science MSc/PhD Students: *Crowdsourcing a Web Science Primer*

3. The Web Science MSc

Question 1.

Did you complete the University of Southampton Web Science MSc?

Yes, No

Question 2.

What was the title of your MSc Dissertation?

Question 3.

What keywords would you use to describe your MSc Dissertation?

Question 4.

Please list the 3 aspects of the Web Science MSc that you found most valuable and relevant to your current PhD studies?

Question 5. (This question is optional)

Please list the 3 aspects of the Web Science MSc that you found most difficult/challenging?

Question 6. (This question is optional)

Are there any Subjects that you did not cover during the Web Science MSc that you feel would have benefited you in your current work?

Question 7. (This question is optional)

Why did you choose to study Web Science?

WAIS Fest Survey of Web Science MSc/PhD Students: *Crowdsourcing a Web Science Primer*

4. About You

Questions 1 is optional, but is required if you wish to contribute to the Web Science ebook, as the information provided will be attributed to you.

Question 1. (This question is optional)

Please give your Name:

Question 2. (This question is optional)

Your Email Address:

Question 3. (This question is optional)

What was your Undergraduate degree discipline(s)?

Question 4. (This question is optional)

Do you have a personal website - if yes, what is the URL?

APPENDIX (III): CONSENT FORM

Consent Form

Ethics reference number: **ERGO/Error! Unknown document property name./22953**	Version: 3	Date: Aug 2016
Study Title: Crowdsourcing a Web Science Primer		
Investigator: Elisabeth Coskun		

Please tick the box(es) if you agree with the statement(s):

This study involves the production of a case study for each participant which will contribute to a Web Science ebook. Your name, email and personal website link will be included in the case study in order to attribute your work to you, and is an opportunity for you to promote your work in the Web Science field.

Please indicate if you are happy to contribute to the ebook. ☐

I have read and understood the Participant Information (version 1 dated 2015-6-2) and have had the opportunity to ask questions about the study. ☐

I agree to take part in this study ☐

I consent to the interview being recorded (for transcription purposes only) ☐

I understand my participation is voluntary and I may withdraw at any time and for any reason. ☐

Data Protection

I understand that information collected during my participation in this study is completely anonymous / will be stored on a password protected computer/secure University server and that this information will only be used in accordance with the Data Protection Act (1998). The DPA (1998) requires data to be processed fairly and lawfully in accordance with the rights of participants and protected by appropriate security.

Name of participant (print name)………………………………………………….

Signature of participant…………………………………………………………

Date……………………………………………………………………………

APPENDIX (IV): DPA PLAN

18.2 DPA Plan

Ethics reference number: **ERGO/Error! Unknown document property name./22953**	Version: 3	Date: 23rd Aug 2016
Study Title: Crowdsourcing a Web Science Primer		
Investigator: Elisabeth Coskun Name of supervisor(s) (if student **investigator**(s)): Su White		

The questionnaire shown in appendix II of this document provides the complete list of all the data that will be collected.

Data Relevance: The data is relevant to the study purposes because it will gather data relating to the WAIS research groups' research subjects, and feedback relating to a Southampton MSc course. The data is adequate because it adequately identifies participants without gathering unnecessary personal data, and it also asks a range of carefully chosen questions relating to the participant's research topic, the Web Science MSc programme and the participants views of the MSc. The data is not excessive because it includes no unnecessary personal data.

The personal data requested by the study are the names and e-mail addresses of participants and their public website addresses and dissertation titles. Names and emails are gathered so that the information that the participant provides may be attributed to them, giving them the opportunity to promote their work and contribution to Web Science, by having a summary of their work published in a case study. Student's dissertation titles are required in order to ascertain the nature of what the student is/has studied.

If students do not consent to contributing to the ebook, their information will not be published.

Some questions in sections 2 and 3 are listed as 'optional;'. This allows participants to not answer if they are not comfortable with providing the information. For example, some people may not be happy to list 3 point that they didn't like about the Web Science MSc and have that information published. Questions 1 of section 3 (Participant's name) is optional, but it is stated that on the survey that it is required if participants wish to contribute to the Web Science ebook, as the information provided will be attributed to them.

Data Accuracy: The majority of the information required by the study is subjective and relates to the knowledge and opinions of the participant of their individual research area, and their personal experiences of completing an MSc course. Interviews conducted in person will be recorded with the consent of each participant. Should they not wish the interview to be recorded and this leads to incomplete information, it may be necessary to contact the participant a second time. If it is not possible to provide an accurate account for a selected participant, then their information will not be used to compile a case study.

Data Retention: By agreeing to participate in the study, students are agreeing to contribute to a case study which may be published as part of a Web Science ebook. The data will be archived on ePrints by the investigators at the end of the research period (end of December 2017) in accordance to university policy http://library.soton.ac.uk/researchdata Physical data will be shredded after it has been scanned and added to the collection of data on the researchers' computers. Sound recordings made from the interviews will be used for transcription purposes only, will be stored securely and will be deleted upon completion of the study.

Data Processing: The data will be processed fairly, and the participants will have given explicit consent to take part in the study. The data will be processed in accordance with the rights of the participants. Rights: Participants will have the right to access, correct, and/or withdraw their data at any time and for any reason. Participants will be able to exercise their rights by contacting the investigators (e-mail: eac1g09@ecs.soton.ac.uk) or the project supervisor (e-mail: saw@ecs.soton.ac.uk).

Consent forms will be attached to the survey, and participants will be required to fill these out before they complete the attached survey.

The following data will not be transferred outside the European Economic Area (EEA).

19 Analysis of Web Science Conference Papers

The following section includes the manually coded analysis of Web Science conference proceedings from 2009 to 2017. Highlighted sections are examples of thought process and brief notes/memos which were recorded during the coding process.

Web Science 2009:

http://journal.webscience.org/view/events/WebSci=2709=3A_Society_On-Line/paper.html

Papers 1: Teaching and Learning

Web and education, a successful open entanglement.
Authors: Minguillón, Julià
Key Topics: Education, e-learning, Web Society, Web Technologies, Hardware,
Paper Key Topics: Academic Publishing, User Behaviour, Online Engagement, Open Data,
Author Keywords:

Semantic Technologies for Learning and Teaching in the Web 2.0 era - A survey.
Authors: Thanassis Tiropanis, Hugh Davis, Dave Millard, Mark Weal,
Key Topics: Web Technologies, Semantic Web, Web 2.0, Education, e-learning, Linked Data,
Paper Key Topics: Academic Publishing, Online Engagement, Folksonomy,
Author Keywords:

The oreChem Project: Integrating Chemistry Scholarship with the Semantic Web.
Authors: Lagoze, Carl
Key Topics: Semantic Web, Web Architecture, Networks, Linked Data, Web Ontologies,
Paper Key Topics: Web Archiving, RDF,
Author Keywords:

Papers 2: Trust and Distrust

Why Bowl Alone When You Can Flashmob the Bowling Alley?: Implications of the Mobile Web for Online-Offline Reputation Systems.
Authors: Karpf, David
Key Topics: Mobile Web, Web Society, Quality Management, General Web Science, Privacy and Trust
Paper Key Topics: Recommender Systems, Algorithms,
Author Keywords:

Trust- and Distrust-Based Recommendations for Controversial Reviews
Authors: Victor, Patricia and Cornelis, Chris and De Cock, Martine and Teredesai, Ankur
Key Topics: Web Society, Web Analytics, Privacy and Trust, Quality Management, Data Mining,
Paper Key Topics: Recommender Systems, eCommerce, Algorithms,
Author Keywords:

The Devil's Long Tail: Religious Moderation and Extremism on the Web.
Authors: O'Hara, Kieron and Stevens, David
Key Topics: Web Society, Web Analytics, Psychology, Digital Sociology, Politics,
Paper Key Topics: Religion, User Behaviour, Virtual Community, Social Science,
Author Keywords:

Papers 3: Openness and Control

Lessons from the Net Neutrality Lobby: Balancing openness and control in a networked society.
Authors: Powell, Alison
Key Topics: Politics, Web Society, Privacy and Trust, Digital Sociology, Web History, Law, Networks, Cloud Computing
Paper Key Topics:
Author Keywords:
^ Not specifically network science?

The consumer and the Web: a critical revision of the contributions to Web science from the marketing and the consumer behaviour discipline
Authors: Rodríguez-Ardura, Inma and J. Martínez-López, Francisco and Luna, Paula
Key Topics: General Web Science, Marketing, Web Society, Web 2.0, Web Analytics,
Paper Key Topics: User Behaviour, eCommerce,
Author Keywords:

Can cognitive science help us make information risk more tangible online?
Authors: Creese, Sadie and Lamberts, Koen
Key Topics: Risk Management, Web Society, Web Analytics, Security, Psychology,
Paper Key Topics: Cognitive Science, Social Science,
Author Keywords:

Designing effective regulation for the "Dark side" of the Web.
Authors: Richter, Wolf and Brown, Ian
Key Topics: Law, Web Society, Risk Management, Web Architecture,
Paper Key Topics: Cybercrime, Social Science,
Author Keywords:

Papers 4: Tagging and Search

On Measuring Expertise in Collaborative Tagging Systems.
Authors: Man Au Yeung, Ching and Noll, Michael and Gibbins, Nicholas and Meinel, Christoph and Shadbolt, Nigel
Key Topics: Web Analytics, Big Data, Data Mining,
Paper Key Topics: Algorithms, Folksonomy, Knowledge Patterns,
Author Keywords:
^ not geo- tagging? Knowledge patterns? Big Data = data set?

Evaluating implicit judgements from Image search interactions
Authors: Smith, Gavin and Ashman, Helen
Key Topics: Web Search, Media, Quality Management, Big Data,
Paper Key Topics:
Author Keywords:

Ephemeral emergents and anticipation in online connected creativity.
Authors: Warren, Lorraine and Fuller, Ted and Lane, Giles and Bryan-Kinns, Nick and Roussos, George and Lesage, Frederik
Key Topics: Media, Web Society, Research Methodology, Information Communication Technology ICT, Business,
Paper Key Topics: Arts and Culture, Knowledge Patterns, Web Publishing, Online Engagement, Social Science,
Author Keywords:

Federating Distributed Social Data to Build an Interlinked Online Information Society.
Authors: Passant, Alexandre and Samwald, Matthias and Breslin, John and Decker, Stefan
Key Topics: Big Data, Linked Data, Web Society, Semantic Web, Social Networking, Web 2.0,
Paper Key Topics: Web Archiving, Social Machines,
Author Keywords: Social Semantic Web, Linked Data, SIOC, Collective Intelligence

Papers 5: Social Networking

KERNEL MODELS FOR COMPLEX NETWORKS.
Authors: Mihail, Milena and Amanatidis, Yorgos and Young, Stephen
Key Topics: Web Graph, Web Modelling, Networks, Network Theory, Web Ontologies, Web Architecture, Web Analytics,
Paper Key Topics: Network Science,
Author Keywords:

Six Degrees of Separation in Online Society.
Authors: Zhang, Lei and Tu, Wanqing
Key Topics: Networks, Network Theory, Web Analytics, Web Society, Social Networking, Web Modelling, Web Graph,
Paper Key Topics: Virtual Community, Content Analysis, Online-Offline communities,
Author Keywords: Six degrees of separation, social network, minimum diameter problem, small world

BFFE (Be Friends Forever): the way in which young adolescents are using social networking sites to maintain friendships and explore identity.
Authors: Clarke, Barbie
Key Topics: Psychology, Web Society, Social Networking, Sociology, Digital Sociology,
Paper Key Topics: Online Identity, User Behaviour, Online Offline Community, Communication Science,
Author Keywords: None

Papers 6: Web of Data

Class Association Structure Derived From Linked Objects.
Authors: Qu, Yuzhong and Ge, Weiyi and Cheng, Gong and Gao, Zhiqiang
Key Topics: Web Application, Web Graph, Semantic Web, Linked Data,
Paper Key Topics: RDF, Network Science,
Author Keywords: Object link, class association graph, complex network analysis

Social Meaning on the Web: From Wittgenstein To Search Engines.
Authors: Halpin, Harry
Key Topics: Semantic Web, Web Society, Social Networking, Web Search, Linked Data, Hypertext,
Paper Key Topics: Natural Language Processing,
Author Keywords:
^ URI – Web Protocols? = Semantic Web

Interactive Information Access on the Web of Data
Authors: Hardman, Lynda and van Ossenbruggen, Jacco and Troncy, Raphael and Amin, Alia and Hildebrand, Michiel
Key Topics: Big Data, Linked Data, Cloud Computing, Accessibility, Web Search, Web Analytics, Web Society,
Paper Key Topics: Web Archiving, Open Data,
Author Keywords:
^ Web Technologies?

Introducing new features to Wikipedia - Case studies for Web science.
Authors: Schindler, Mathias and Vrandecic, Denny
Key Topics: Web Society, General Web Science,
Paper Key Topics: Social Machines, Web Publishing,
Author Keywords:
^ Web Science – Ref in Lit Review? – described as an example of a sociotechnical process

Papers 7: Government, Citzens and Law on the Web

Is web-based interaction reshaping the organizational dynamics of public administration?: A comparative empirical study on eGovernment.
Authors: Batlle, Albert and Waksberg, Ana and Aibar, Eduard
Key Topics: Web Society, Web 2.0, General Web Science, Information Theory, Information Communication Technology, Information Systems,
Paper Key Topics: Government, User Behaviour,
Author Keywords:

Online Dispute Resolution: Designing new Legal Processes for Cyberspace.
Authors: Katsh, Ethan
Key Topics: Law, Risk Management, Web Society,
Paper Key Topics: Social Science,
Author Keywords:

Experiments for Web Science: Examining the Effect of the Internet on Collective Action
Authors: Margetts, Helen and John, Peter
Key Topics: General Web Science, Politics, Web Society, Web Analytics, Politics,
Paper Key Topics: Social Science, Online Engagement, User Behaviour,
Author Keywords:

Papers 8: Life On-line

LifeGuide: A platform for performing web-based behavioural interventions
Authors: Hare, Jonathon and Osmond, Adrian and Yang, Yang and Wills, Gary and De Roure, David and Joseph, Judith and Yardley, Lucy and Weal, Mark
Key Topics: Psychology, Web Development, Web Society,
Paper Key Topics: User Behaviour, Social Science, Online Engagement,
Author Keywords:

Securing Cyberspace: Realigning Economic Incentives in the ICT Value Net.
Authors: van Eeten, Michel and Bauer, Johannes
Key Topics: Economics, Security, Information Communications Technology, Web Society,
Paper Key Topics: Cybercrime,
Author Keywords: None

Cognition, Cognitive Technology and the Web
Authors: Carr, Leslie and Harnad, Stevan
Key Topics: Web Search, Web Society, HCI Human Computer Interaction,
Paper Key Topics: Cognitive Science, User Behaviour,
Author Keywords:

Web Science 2010:

http://journal.webscience.org/view/events/WebSci10=3A_Extending_the_Frontiers_of_Society_On-Line/paper.html

WS10 Paper Session 1: Web & Society

The Value (Driven) Web
Authors: Baken, Nico and van Oortmerssen, Gerard and Wiegel, Vincent
Key Topics: Ethics, Web Society, Digital Sociology, Privacy and Trust, Information Communication Technology,
Paper Key Topics: Social Science,
Author Keywords: Values, Value Sensitive Design, Holons, Trans-sector Innovation, Functional Decomposition, Internet, Ethics, Trust

Web 2.0 Vision for the Blind.
Authors: Baumgartner, Robert and Fayzrakhmanov, Ruslan and Gattringer, Rafael and Göbel, Max and Holzinger, Wolfgang and Klein, David and Kruepl, Bernhard
Key Topics: Web 2.0, Accessibility, Web Technologies, Web Society, Research Methodology, Human Computer Interaction,
Paper Key Topics: User Behaviour, Cognitive Science,
Author Keywords: accessibility, blind, navigation, screen reader

#iranElection: quantifying online activism.
Authors: Gaffney, Devin
Key Topics: Web 2.0, Data Mining, Social Networking, Web Society, Web Analytics, Politics, Media,
Paper Key Topics: Microblogging, User Behaviour, Government, Online Engagement, Social Science,
Author Keywords: Twitter, Online Activism, Iran Election 2009,
^ Look at data presentation!
From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search
Authors: Mustafaraj, Eni and Metaxas, Panagiotis
Key Topics: Web Search, Social Networking, Politics, Web Graph, Data Mining,
Paper Key Topics: Microblogging, User Behaviour,
Author Keywords: Social Web, Real-Time Web, US elections, Twitter, Twitter-bomb, Google

Intimacy 2.0: Privacy Rights and Privacy Responsibilities on the World Wide Web.
Authors: O'Hara, Kieron
Key Topics: Web 2.0, Web Technologies, Web Society, Web Analytics, Privacy and Trust, Politics, Social Networking,
Paper Key Topics: Social Science, User Behaviour, Online Identify,
Author Keywords:

Copyright on the Web: Looking for a Snap Answer to a Fundamental Conflict
Authors: Wilson, Caroline
Key Topics: Digital Copyright, Law, Web Society,
Paper Key Topics:
Author Keywords:

A Peep at Pornography Web in China.
Authors: Wu, Zhaohui and Jiang, Lu and Zheng, Qinghua and Tian, Zhenhua and Liu, Jun and Zhao, Junzhou
Key Topics: Web Society, Criminology, Psychology, Digital Sociology, Economics,
Paper Key Topics: User Behaviour, Geography, Cybercrime, Social Science,
Author Keywords: Pornography web, China, anti-pornography, power law,
^ power laws mentioned again (Does cybercrime cover porn?) – economics = ecommerce?

WS10 Paper Session 2: Web & Communities

EXPLORING AREA-SPECIFIC MICROBLOGGER SOCIAL NETWORKS
Authors: Aksu Degirmencioglu, Ece and Uskudarli, Suzan
Key Topics: Social Networking, Web Society, Web Analytics, Web Applications,
Paper Key Topics: Microblogging, Social Science, Knowledge Patterns, Topic Modelling, Content Analysis, Virtual Community, Narrative, Social Machines,
Author Keywords: Interest Networks, User suggestion, Twitter, keyword co-occurrence, social network analysis, community extraction,
^ Recommender Systems?

Analysis of Strategies for Item Discovery in Social Sharing on the Web.
Authors: Au Yeung, Ching Man
Key Topics: Web Analytics, Web Society, Social Networking, Big Data, Data Mining, Media,
Paper Key Topics: Virtual Community, Web Publishing,
Author Keywords: Social sharing, item discovery, item adoption
^ Online sharing

Decomposing Discussion Forums using Common User Roles.
Authors: Chan, Jeffrey and Hayes, Conor and Daly, Elizabeth
Key Topics: Web 2.0, Web Society, Web Analytics, Web Applications, Web Technologies, Web Architecture,
Paper Key Topics: User Behaviour, Web 3.0, Communication Science,
Author Keywords:

Studying Scientific Discourse on the Web using Bibliometrics: A Chemistry Blogging Case Study.
Authors: Groth, Paul and Gurney, Thomas
Key Topics: Web Society, Web Analytics,
Paper Key Topics: Microblogging, Communication Science, Scientific Method,
Author Keywords: bibliometrics, science blogs, researchblogging.org, scientific discourse
^ bibliometrics (again?!)

Understanding how Twitter is used to spread scientific messages.
Authors: Letierce, Julie and Passant, Alexandre and Breslin, John and Decker, Stefan
Key Topics: Web Society, Semantic Web, Web 2.0, Web Technologies, Media,
Paper Key Topics: Scientific Method, Microblogging, Virtual Community, Communication Science,
Author Keywords:

Harnessing the Social Web: The Science of Identity Disambiguation
Authors: Rowe, Matthew and Ciravegna, Fabio
Key Topics: Web Society, Web Analytics, Social Networking, Digital Sociology,
Paper Key Topics: Online Identify, Machine Learning,
Author Keywords:

WS10 Paper Session 3: Web & Data

Exploiting a Web of Semantic Data for Interpreting Tables.
Authors: Finin, Tim and Syed, Zareen and Mulwad, Varish and Joshi, Anupam
Key Topics: Semantic Web, Linked Data, Web Technologies, Web Modelling, Web Analytics, Cloud Computing, Data Mining,
Paper Key Topics: RDF, Web Archiving, Web Publishing,
Author Keywords:

Standing on the shoulders of the trusted web: Trust, Scholarship and Linked Data.
Authors: Gamble, Matthew and Goble, Carole
Key Topics: Linked Data, Privacy and Trust, Web Society, Web Analytics, Research Methodology,
Paper Key Topics: Social Science, Web Publishing, Scientific Method, Virtual Community, Knowledge Patterns,
Author Keywords: Trust, Scholarship, Linked Data, Sharing, Science,

Global Integration of Public Sector Information.
Authors: Koumenides, Christos and Alani, Harith and Shadbolt, Nigel and Salvadores, Manue
Key Topics: Linked Data, Big Data, Web Architecture, Web Society, Digital Sociology,
Paper Key Topics: Web Archiving, Government, Open Data, Geography, Social Science,
Author Keywords:
Study Supreme Court Decision Making with Linked Data.
Authors: Li, Xian and Ding, Li and Hendler, James
Key Topics: Law, Linked Data, Semantic Web, General Web Science, Digital Sociology, Web Society, Web Analytics, Web Technologies,
Paper Key Topics: Social Science,
Author Keywords: Linked Data, Semantic Web, Web Science, Supreme Court, Decision Making,

Learning from Linked Open Data Usage: Patterns & Metrics.
Authors: Möller, Knud and Hausenblas, Michael and Cyganiak, Richard and Grimnes, Gunnar and Handschuh, Siegfried
Key Topics: Linked Data, Big Data, Web Analytics, Cloud Computing, Semantic Web, Data Mining,
Paper Key Topics: Open Data, RDF, Online Engagement,
Author Keywords: linked data, Web of Data, access, usage pattern

Lightweight Reasoning, and the Web of Data
Authors: Thomas, Edward and Pan, Jeff Z. and Taylor, Stuart and Ren, Yuan
Key Topics: Semantic Web, Web Ontologies, Web Modelling, Linked Data, Big Data, Web Technologies, Web Languages, Cloud Computing, Web Graph, Networks, Network Theory,
Paper Key Topics: RDF, Web Archiving, Network Science,
Author Keywords: Semantic Web, Reasoner, Ontology, RDF, Linked Data

WS10 Paper Session 4: Web & Intelligence

A Trust Model to Estimate the Quality of Annotations using the Web
Authors: Ceolin, Davide and van Hage, Willem R. and Fokkink, Wan
Key Topics: Semantic Web, Privacy and Trust, Quality Management, Web Technologies, Web Analytics, Media, Security, Digital Sociology, Web Modelling, Web Ontologies,
Paper Key Topics: RDF,
Author Keywords: Trust, Annotations, Subjective Logic, Semantic Web,

Peer-to-Peer Human Computation & "Help Me Decide": Enabling search users to help other users make purchase decisions.
Authors: Chandrasekar, Raman and Jain, Kamal
Key Topics: Web Search, Web Society, economics, Marketing,
Paper Key Topics: Recommender Systems, ecommerce,
Author Keywords: Purchase decisions, shopping, incentives, peer-to-peer human computation, search engines,

Crowdsourcing Scholarly Data.
Authors: Hoang, Diep Thi and Kaur, Jasleen and Menczer, Filippo
Key Topics: Web Society, Web Analytics,
Paper Key Topics: Academic Publishing, Crowdsourcing, Web Publishing, Content Analysis,
Author Keywords: Crowdsourcing, citation analysis, scholarly data, impact measures, discipline annotations, social tagging,

The Language Grid for Intercultural Collaboration.
Authors: Ishida, Toru
Key Topics: Web Society, Digital Sociology,
Paper Key Topics: Social Machines, Social Science, Geography, Virtual Community,
Author Keywords: Language Grid, intercultural collaboration, service grid,

VIVO: Enabling National Networking of Scientists.
Authors: Krafft, Dean B. and Cappadona, Nicholas A. and Caruso, Brian and Corson-Rikert, Jon and Devare, Medha and Lowe, Brian J. and VIVO Collaboration,
Key Topics: Semantic Web, Linked Data, Web Ontologies, Research Methodology, Web Society, General Web Science, Web Technologies,
Paper Key Topics: Open Data, Network Science, Virtual community,
Author Keywords:

Helping online communities to semantically enrich folksonomies
Authors: Limpens, Freddy and Gandon, Fabien and Buffa, Michel
Key Topics: Web Society, Web Ontologies, Social Networking, Semantic Web, Web Technologies,
Paper Key Topics: Folksonomies, Virtual Community, RDF, Social Machines,
Author Keywords:
^**Tagging** – not geotagging (again)?

Beyond Convergence: User Evaluations of Identical Syndicated News Content Across Media.
Authors: Treem, Jeffrey and Thomas, Kristin
Key Topics: Media, Web Society, Research Methodology, Quality Management,
Paper Key Topics: User Behaviour,
Author Keywords: Blogs, Blog Readers, Convergence, Credibility, Online News, Syndication,

WS10 Session 5: Web & Methodology

Could the Web be a Temporary Glitch?
Authors: Carr, Leslie and Pope, Cathy and Halford, Susan
Key Topics: General Web Science, Web Society, Web Technologies, Networks,
Paper Key Topics: Virtual Community, Social Science,
Author Keywords: Web science, open access, information sharing

Anchors in Shifting Sand: the Primacy of Method in the Web of Data.
Authors: De Roure, David and Goble, Carole
Key Topics: Linked Data, Big Data, Web 2.0, Research Methodology, Social Networking, Web Society, Privacy and Trust,
Paper Key Topics: Virtual Community,
Author Keywords: Reproducible research, linked data, myExperiment, scientific workflows

A Manifesto for Web Science
Authors: Halford, Susan and Pope, Catherine and Carr, Leslie
Key Topics: General Web Science, Research Methodology, Web Society,
Paper Key Topics: Social Science, User Behaviour,
Author Keywords: Web Science,

Towards a Philosophy of the Web: Representation, Enaction, Collective Intelligence.
Authors: Halpin, Harry and Clark, Andy and Wheeler, Michael
Key Topics: General Web Science, Web Society, Semantic Web, Web 2.0,
Paper Key Topics: Philosophy, Cognitive Science, Knowledge Patterns,
Author Keywords: philosophy, representations, enaction, collective intelligence

Teasing Apart and Piecing Together: Towards Understanding Web-based Interactions.
Authors: Hooper, Clare J. and Millard, David E.
Key Topics: General Web Science, Psychology, Web Society, Web Analytics, Web Design, Human Computer Interaction,
Paper Key Topics: User Behaviour,
Author Keywords: web-based interactions, physical to digital, understanding, analysis, design

Behavior Change Support Systems: The Next Frontier for Web Science.
Authors: Oinas-Kukkonen, Harri
Key Topics: Psychology, Web Society, General Web Science, Information Systems, Web 2.0,
Paper Key Topics: User Behaviour,
Author Keywords: Socio-technical system, behavioral outcomes, psychological outcomes, behavioral change, persuasive technology.

Web Science 2011: https://www.websci11.org/www.websci11.org/program/index.html

Session 1: Analysis of the Web and Web Users

ATT: Analyzing Temporal Dynamics of Topics and Authors in Social Media.
Authors: Nasir Naveed, Sergej Sizov and Steffen Staab
Key Topics: Social Networking, Web Society, Web Analytics, Web Modelling, Statistical Analysis,
Paper Key Topics: Microblogging, Topic Modelling, User Behaviour,
Author Keywords: Probabilistic Models, Topic Modeling, Text Categorization,

Analyzing Temporal Dynamics in Twitter Profiles for Personalized Recommendations in the Social Web.
Authors: Fabian Abel, Qi Gao, Geert-Jan Houben and Ke Tao
Key Topics: Social Networking, Web Analytics, Web Modelling, Web Society, Psychology,
Paper Key Topics: Microblogging, User Behaviour, Content Analysis, Recommender Systems,
Author Keywords:

Dengue surveillance based on a computational model of spatio-temporal locality of Twitter.
Authors: Janaína Gomide, Adriano Veloso, Wagner Meira Jr., Fabrício Benevenuto, Virgílio Almeida, Fernanda Ferraz and Mauro Teixeira
Key Topics: Social Networking, Web Analytics, Web Society, Data Mining, Web Modelling, Statistical Analysis,
Paper Key Topics: Microblogging, Content Analysis, Health,
Author Keywords: Twitter, Dengue, Surveillance, Spatio-temporal data mining,

Classifying Queries Submitted to a Vertical Search Engine.
Authors: Richard Berendsen, Bogomil Kovachev, Edgar Meij, Maarten de Rijke and Wouter Weerkamp
Key Topics: Web Search, Web Analytics,
Paper Key Topics: Content Analysis, Topic Modelling,
Author Keywords: People search, query log analysis, classification

Using TAPT as an Analytical Method for Understanding Online Experiences.
Authors: Clare J. Hooper
Key Topics: Mobile Web, Web Analytics, Web Society,
Paper Key Topics: Virtual Community, User behaviour, Geography, Geo-tagging, Sentiment Analysis,
Author Keywords: Web-based interactions, online lives, analysis, TAPT, TAGMA,
^ More textual analysis (Geo-caching!)

Towards a diversity-minded Wikipedia.
Authors: Fabian Flöck, Denny Vrandecic and Elena Simperl:
Key Topics: Web Society, Web Analytics, Data Mining, Semantic Web, Research Methodologies,
Paper Key Topics: User behaviour, Social Machines, Sentiment Analysis, Academic Publishing, Content Analysis,
Author Keywords: Wikipedia, diversity, community-driven content creation, social dynamics, opinion mining, sentiment analysis
^ More textual analysis

Session 2a: Microlevel Processes on the Web

Social Comparisons at Your Fingertips: The Importance of Majority/Minority Status.
Authors: Silvia Knobloch-Westerwick and Axel Westerwick
Key Topics: Psychology, Web Society, Social Networking, Statistical Analysis,
Paper Key Topics: Demographics, User Behaviour,
Author Keywords: Selective exposure, social comparison, optimal distinctiveness, ingroup, outgroup, self-categorization, self-esteem

Accountability and Deterrence in Online Life (Extended Abstract).
Authors: Joan Feigenbaum, James Hendler, Aaron D. Jaggard, Daniel Weitzner and Rebecca Wright
Key Topics: Security, Privacy and Trust, Web Society,
Paper Key Topics: Cybercrime, User Behaviour,
Author Keywords: Accountability, Deterrence, Privacy, Security

Bad News Travel Fast: A Content-based Analysis of Interestingness on Twitter.
Authors: Nasir Naveed, Thomas Gottron, Jérôme Kunegis and Arifah Che Alhadi
Key Topics: Data Mining, Social Networking, Web Analytics, Web Society, Media,
Paper Key Topics: Content Analysis, Microblogging, Sentiment Analysis,
Author Keywords:

A Shifting Boundary: the dynamics of internal cognition and the web as external representation.
Authors: Alan Dix
Key Topics: Human Computer Interaction, Web Society, Psychology,
Paper Key Topics: Cognitive Science, Philosophy,
Author Keywords: embodiment, external representation, distributed cognition

Social Practices around Personal Videos using the Web.
Authors: Rodrigo Laiola Guimaraes, Pablo Cesar, Dick C.A. Bulterman, Ian Kegel and Peter Ljungstrand:
Key Topics: Web Society, Media, Privacy and Trust, Social Networking, Web Analytics,
Paper Key Topics: User Behaviour, Content Analysis, Web Publishing,
Author Keywords: Social Multimedia, Video Sharing, Personal Videos, Web- Mediated Communication, Connectedness.

Session 2b: Web Science Tools and Technologies

Digital Hermeneutics: Agora and the Online Understanding of Cultural Heritage.
Authors: Chiel van den Akker, Susan Legêne, Marieke van Erp, Lora Aroyo, Roxane Segers, Lourens van der Meij, Jacco van Ossenbruggen, Guus Schreiber, Bob Wielinga, Johan Oomen and Geertje Jacobs
Key Topics: Web Society, Web History, Web Technologies, Web Modelling,
Paper Key Topics: Arts and Culture, Web Archiving, Narrative,
Author Keywords: Digital hermeneutics, online cultural heritage, collection enrichment

WSCD: Negotiating the Web Science Curriculum Development through Shared Educational Artefacts.
Authors: Madalina Croitoru, Stephane Bazan, Stefano Cerri, Hugh Davis, Clement Jonquet, Gianfranco Prini, Francois Scharffe, Steffen Staab, Michalis Vafopoulos and Su White
Key Topics: General Web Science, Education, e-learning, Web Society,
Paper Key Topics: Academic Publishing, Web Publishing, Web Archiving
Author Keywords: Web Science Education, Web Science Curriculum, Educational Repository, Negotiated Curriculum, Co-Creation
^ Repository = Web Archiving

The Past Issue of the Web.
Authors: Helen Hockx-Yu:
Key Topics: Law, Big Data, Web Technologies, Web Analytics, General Web Science, Visualisation,
Paper Key Topics: Web Archiving, Content Analysis,
Author Keywords: Heritage, Academic Research and the Web, Web Archiving. Library Information Management, Digital Libraries, Web Archive, Web Harvesting, Electronic Legal Deposit, Digital Preservation.

Session 2c: Governance & Trust

Survey on Governance of User-generated Content in Web Communities.
Authors: Felix Schwagereit, Ansgar Scherp and Steffen Staab
Key Topics: Web Governance, Media, Social Networking, Web Society, Web Analytics,
Paper Key Topics: Virtual Community, User Behaviour, Content Analysis,
Author Keywords: Governance, Web Community, Rating System

A Farm in Every Window: A Study into the Incentives for Participation in theWindow Farm Virtual Community.
Authors: Dominic Difranzo and Alvaro Graves
Key Topics: Web Society, Web Analytics, Web Modelling, Research Methodology, Psychology,
Paper Key Topics: Virtual Community, Online Engagement, User Behaviour, Content Analysis,
Author Keywords: Motivation, virtual communities, contribution, knowledge sharing

Quality, Trust and Utility of Scientific Data on the Web: Towards a Joint Model.
Authors: Matthew Gamble and Carole Goble
Key Topics: Quality Management, Privacy and Trust, Big data, Web Modelling, Web Analytics,
Paper Key Topics: Scientific Method, Content Analysis,
Author Keywords: Science, Data Quality, Trust, Data Sharing, Decision Networks

Session 2d: Web Technologies

Visualising the Past: Annotating a Life with Linked Open Data.
Authors: Ashley Smith, Kieron O'Hara and Paul Lewis:
Key Topics: Linked Data, Big Data, Web Technologies, Web Society, Web Analytics, Visualisation, Privacy and Trust,
Paper Key Topics: Open Data, Content Analysis, User Behaviour, Geo-tagging, RDF, Provenance,
Author Keywords:

The communication infrastructure during the learning process in web based collaborative learning systems.
Authors: Cathleen M. Stuetzer, Kathleen M. Carley, Thomas Koehler and Gerhard Thiem
Key Topics: Web Society, Networks, Web Architecture, Education, e-learning, Statistical Analysis,
Paper Key Topics: Social Science, Communication Science, Virtual Communities, Online Engagement, Network Science,
Author Keywords: Learning networks, Role analysis, Emerging roles,
Communication roles, Network position, Social learning, Social Network Analysis (SNA), Web-based collaborative learning,

Sharing innovative teaching experience in higher education on the Web: An interdisciplinary study on a contextualized Web 2.0 application for community building and teacher training.
Authors: Stéphane Bernard Bazan, Christophe Varin and Sabrine Saad:
Key Topics: Web 2.0, Education, e-learning, Web Society, Media, General Web Science, Web Technologies, Statistical Analysis,
Paper Key Topics: User Behaviour, Virtual Community,
Author Keywords: Innovation, Education, Web 2.0, Collaborative projects, Teachers learning, Teaching experience, CMS, Web design, Digital Power
^ Collective Intelligence = knowledge patterns? Long tail

Social Media on the Job: An exploration of the potential legal consequences of employer's social media usage whilst during the course of employment.
Authors: Sarosh Khan, Roksana Moore and Dr Mark Weal
Key Topics: Social Networking, Law, Web Society,
Paper Key Topics: Microblogging,
Author Keywords: Social media, Twitter, Facebook, Weblogs, Employment Law, Harassment, Defamation

Web Technologies for Open Innovation.
Authors: Darko Jesic, Jovana Kovacevic and Milan Stankovic
Key Topics: Web Technologies, Web Search, Social Networking,
Paper Key Topics: User Behaviour, Virtual Community,
Author Keywords: Open Innovation, Web, Key Words Matching, Social Propagation and Expert Search.

Session 3a: Advances in Network Studies of the Web

Small Worlds with a Difference: New Gatekeepers and the Filtering of Political Information on Twitter.
Authors: Pascal Juergens, Andreas Jungherr and Harald Schoen
Key Topics: Social Networking, Politics, Web Society, Web Analytics, Data Mining,
Paper Key Topics: Network Science, Microblogging, Topic Modelling,
Author Keywords: Bundestagswahl 2009, Entropy, Gatekeepers, Network Analysis, Political Communication, Twitter.

The Web as an Adaptive Network: Coevolution of Web Behavior and Web Structure.
Authors: Connor Mccabe, Richard A. Watson, Jane Prichard and Wendy Hall:
Key Topics: General Web Science, Networks, Social Networking, Web Architecture,
Paper Key Topics: Network Science, User Behaviour,
Author Keywords: Networks, User Behaviour, Adaptive, Structure, Dynamics, Simulation,

The Effect of User Features on Churn in Social Networks.
Authors: Marcel Karnstedt, Matthew Rowe, Jeffrey Chan, Harith Alani and Conor Hayes:
Key Topics: Social Networking, Web Society, Psychology, Web Analytics, Statistical Analysis,
Paper Key Topics: User behaviour, Virtual Community, Online Engagement,
Author Keywords:
Sic Transit Gloria Mundi Virtuali? Promise and Peril at the Intersection of Computational Social Science and Online Clandestine Organizations
Authors: Brian Keegan, Muhammad Aurangzeb Ahmad, Dmitri Williams, Jaideep Srivastava and Noshir Contractor
Key Topics: Social Networking, Web Society, Web Technologies, Criminology, Virtual Environment,
Paper Key Topics: Social Science, Computer Science, User Behaviour, Virtual Community,
Author Keywords: Gold farming, clandestine organization, social network analysis, massively multiplayer online game, drug trafficking, risk,
^ Gaming ?

Flocks, Herds, and Stories: temporal coherence and the long tail.
Authors: Mark Bernstein
Key Topics: Hypertext, Web Analytics, Web Technologies, Web Society, Statistical Analysis, Psychology, Media,
Paper Key Topics: Narrative, Web Publishing, Content Analysis, User Behaviour, Online Engagement, Philosophy,
Author Keywords: Web; hypertext; browsing; narrative; flocks; herds

Session 3b: Broad Views of Web Science

Is there anybody out there? – Social Media as a new social fetish.
Authors: Karolin Eva Kappler and Ricard Riuz de Querol
Key Topics: Social Networking, Web Society, Research Methodology, Web 2.0,
Paper Key Topics: User Behaviour,
Author Keywords: Interpassivity, interactivity, user profiles, Social Media, social fetish,

Is (Web) Science Ready for Empowerment?
Authors: Hans Akkermans, Nana Baah Gyan, Anna Bon, Wendelien Tuyp, Aman Grewal Stéphane Boyera and Mary Allen
Key Topics: General Web Science, Web Society, Web Technologies, Accessibility, Mobile Web
Paper Key Topics: Geography, Scientific Method,
Author Keywords: The pro-human Web, evolving technologies, knowledge production, Web futures

Hacktivism: a theoretical and empirical exploration of China's cyber warriors.
Authors: Michael Yip and Craig Webber
Key Topics: Criminology, Web Society, Web Analytics, Digital Sociology, Politics, Statistical Analysis,
Paper Key Topics: Cybercrime, User Behaviour, Content Analysis, Social Science,
Author Keywords: Hacktivism, National Humiliation, Ressentiment, Relative Deprivation, China,

Web Science 2012

This year is alphabetical by author !!!!

Web Dynamics as a Random Walk: How and Why Power Laws Occur
Authors: Hans Akkermans
Key Topics: Networks, Network Theory, Social Networks, Web Graph, Web Modelling, Web Ontologies, Statistical Analysis
Paper Key Topics: Network Science
Author Keywords: network formation models, system dynamics, degree distributions, scale-free networks, power laws

This work investigates the conditions under which power laws occur in networks. The work is based on a new network model termed the 'exciton' model, which includes "the processes of link creation, link removal, node creation and node loss". Results are compared with existing models in the network science literature. The authors conclude that power laws in networks can be framed as the result of random graph theory.

Topical Anomaly Detection From Twitter Stream
Authors: Pramod Anantharam, Krishnaprasad Thirunarayan, Amit Sheth
Key Topics: Social Networking, Privacy and Trust, Web Analytics
Paper Key Topics: Microblogging, Content Analysis
Author Keywords: Anomaly detection; spam and off-topic content detection; binary classification; twitter stream analysis

Some users of microblogging platforms will post corrupt and off-topic links for selfish reasons, in an attempt to influence opinion, advertise, bias political opinion, etc.

This work argues for the importance of trustworthiness on microblogging platforms as people are increasingly using posts on such platforms to make informed decisions.

The paper details an analysis which spots anomalies in Tweet streams by analysing the content of the URLs contained in the Tweets. They explain that this is an approach that has so far been neglected by other analysis. The research detailed specifically measures the divergence between subject hashtags suggesting the content of Tweets with the actual content of the documents provided by the URLs.

Synthesis Ranking with Critic Resonance
Authors: Fred S. Annexstein, Kenneth A. Berman
Key Topics: Social Networking, Web Search, Web Society
Paper Key Topics: Synthesis ranking, Algorithms,
Author Keywords: Social Media, Social Choice

Look Who I Found: Understanding the Effects of Sharing Curated Friend Groups Social Networks, Online communities,
Authors: Lujun Fang, Alex Fabrikant, Kristen LeFevre
Key Topics: Social Networking**,** Web Society
Paper Key Topics: Network Science, Online Identify, Virtual community
Author Keywords: Social Network, Google+, Circle Sharing

This work suggests that for users joining a social networking site, building a circle of contacts can be time consuming. This work examines 'Circle sharing', a feature introduced by Google+, which enables users to share their group of contacts directly with another user, who may choose to adopt either all or just some of the contacts. This paper claims to be the first large scale study investigating and reporting on the impact that circle sharing had on the social network Google+. Cluster analysis is used to examine two naturally occurring types of 'circles', those consisting primarily of celebrities and those consisting of members of a community. The authors observed that the circles sharing feature had a measurable impact upon growth. The work also investigates the feasibility of suggesting to users which of their circles they should share with friends.

Recipe recommendation using ingredient networks
Authors: Chun-Yuen Teng, Yu-Ru Lin, Lada A. Adamic
Key Topics: Social Networking, sociology
Paper Key Topics:
Author Keywords: ingredient networks, recipe recommendation
Recipe sharing on the web
Trends in food popularity
Measure the popularity of a specific ingredient

Nutrition – too niche to code?

Recipe sharing online is one of the earliest examples of online collaborative community on the Web. Online recipes are submitted by users, and feature crowdsources rankings and reviews which enable users to assess whether they will like a given recipe, and also include suggestions as to how to improve given recipes. A wealth of information is collected about each recipe, which can reveal a substantial amount of detail about the preferences of individual users, and the overall popularities of dishes, as well as individual ingredients.

This work gathers collective knowledge and preference about cooking through mining a popular recipe-sharing website. The authors construct two types of networks: The complement network captures which ingredients tend to co-occur frequently, and the substitute network, derived from user-generated suggestions for modifications. Experiments suggest that it is possible to predict recipe ratings based upon a combination of ingredient networks and nutrition information.

Personality and Patterns of Facebook Usage
Authors: Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli, David Stillwell
Key Topics: Social Networks, Psychology, Sociology
Paper Key Topics: Social Science, User behaviour, Online Identity
Author Keywords: Social Networks, Personality, Big Five Personality Model

Five Factor model
Users activity on social networking sites reflects their personality

This work examines correlations between users' personality and the properties of their Facebook. For example, number of friends, number of photos uploaded, amount of tags in photos, group membership, etc.

The results of this work show significant relationships between personality traits and various features of Facebook profiles. It is also demonstrated that it is possible to predict personality traits of a user based on Facebook activity.

Four Degrees of Separation
Authors: Lars Backstrom, Paolo Boldi, Marco Rosa, Johan Ugander, Sebastiano Vigna
Key Topics: Networks, Social networks, Web Graph, Network theory, Web Society, Web Analytics, Web Ontologies
Paper Key Topics: Online Identity**,** Algorithms,
Author Keywords: None

Friendship links on social media!
Digital sociology?

The term six degrees of separation suggests that any two people are separated by a maximum of six friendship links. This study claims to be the first world scale study of a social network graph, using the entire network of active users on Facebook to observe friendship links. The findings suggest that the average distance of Facebook is 4:74, that is, 3:74 "degrees of separation", which suggests that on Facebook people are separated by a total of 4 friends. The authors also provide details of some more general study relating to the distance distribution of Facebook, providing detailed statistical metadata which proves the accuracy of their measurements.

Belief Surveillance with Twitter
Authors: Sanmitra Bhattacharya, Hung Tran, Padmini Srinivasan, Jerry Suls
Key Topics: Networks, Social networks, Web Graph, Network theory, Web Society, Web Analytics, Web Ontologies
Paper Key Topics: Online Identity
Author Keywords: microblogs, large-scale data collection, public belief analysis

This work outlines a collaboration between authors from both Computer Science and Psychology. The result is a paper which explores the prospect of utilising people's responses on social media to measure their levels of belief, disbelief and doubt in relation to a specific topic or proposition. They coin the term 'belief surveillance', and propose a mythological framework for conducting this via Twitter. It is suggested that it is possible to measure the level of belief on any proposition, as long as it is specifiable in a form the authors term 'probes'. They define a probe as a "sentence reflecting a directed binary relationship linking two concepts." Given two concepts in a Tweet, e.g. 'autism' and 'vaccine', they then assess the opinion of response to the concepts on a scale of *'support, oppose, doubt and other'*.

We Love Rock 'n' Roll: Analyzing and Predicting Friendship Links in Last FM
Authors: Kerstin Bischoff
Key Topics: Social Networking, Web Graph, Networks, Network Theory, Psychology, Web Ontologies
Paper Key Topics: Online/offline community
Author Keywords: social networks; tie strength; social links; online vs. offline; statistical analysis; machine learning

Digital sociology?
Friendship links on social media!

People are said to gravitate to 'like-minded' people, or people with similar tastes, which is termed 'homophily'. This paper outlines a study of friendship links on the social media site, Last FM. The work searches for similarities in music taste between users, as well as information on demographic attributes and local network structure.

This work looks at the predicting of ties, characterizing online and offline ties and learning to predict both, as well as their strength. The results indicate the predicative power of analysing mutual friends.

A Network Pruning Based Approach for Subset-Specific Influential Detection
Authors: Praphul Chandra, Arun Kalyanasundaram
Key Topics: Network Theory, Social Networks, Web Analytics, Web Modelling,
Paper Key Topics: Network Science, Algorithms,
Author Keywords: influentials, social networks, information diffusion

This work relates to the identification of the most influential nodes in a network. As networks grow larger, it becomes necessary to study the influence of a given node, not on the whole network, but on a subset of nodes within the network. The authors term this ‘subset specific top-k influential problem’.

The outcome of the study makes three key contributions. The first is an iterative network pruning algorithm, the purpose of which is to to find the subset specific top-k influential and compare the effectiveness of these with existing algorithms. The second is to extend the existing analytical framework for top-k, and the third involves the analysis of the analytical frame- work in order to show that the influence spread function continues to be sub-modular.

A Study of Human Flesh Search with Epidemic Models
Authors: Long Cheng, Lei Zhang, Jinchuan Wang
Key Topics: Web Governance, Privacy and Trust, Web Analytics, Social Networks
Paper Key Topics: Health
Author Keywords: Human Flesh Search; HFS; SIR Model; Epidemic Model.

Contrary to what the title may initially suggest, this paper is not related to medical research. Instead, Human Flesh Search (HFS) indicates a collaborative effort to highlight an online user who may have committed some form of misbehaviour online, exposing their personal information.

The authors use a similar mathematical model to those used to track the spread of epidemics in order to track the phenomenon of HFS, and their results suggest that the authors model matches HFS cases effectively.

(NB. Is this still violating their Data protections rights even though they have committed a misbehaviour)

Web supported emplotment: Using object and event descriptions to facilitate storytelling online and in galleries
Authors: Trevor Collins, Paul Mulholland, Annika Wolff
Key Topics: Web Applications, Web Ontologies, Data Mining
Paper Key Topics: Arts and Culture, narrative
Author Keywords: Cultural heritage, emplotment, event-based representation, interpretation, narrative.

This paper looks at a process termed ‘emplotment’ which involves identifying the key events within a story and the relationship between them with the aim of of producing a plot interpreting these evening. The authors then apply this principle to the desktop based Web tools, specifically those used by galleries to design and present exhibitions in order to demonstrate how it is possible to construct a narrative around exhibitions.

‘Digital Natives? Investigating young people’s critical skills in evaluating web based information’
Authors: Huw C. Davies, Susan J. Halford, Nick Gibbins
Key Topics: Digital Sociology, Web Society, Digital Ethnography
Paper Key Topics: Digital Native, Digital literacy
Author Keywords: Digital Natives, digital literacy, critical skill, ethnography, habitus

This work explores the suggestion that young people who have been failed by the education system are incapable of critical thinking and online, and need support to become more ‘savy’ Web users.

The work suggests that popular stereotypes are unhelpful in understanding young people’s relationship with the Web. The authors present the findings of a qualitative study of young people on the web, comparing youths from two different social backgrounds in order to observe differences in their behaviour online. They suggest that it is possible to measure the fact that differing levels of Web Skills can provide an indication of young peoples’ environment, and contextual influences such as family, peers and education.

Tracking Twitter for Epidemic Intelligence
Authors: Ernesto Diaz-Aviles, Avare Stewart
Key Topics: Social Networks, Web Society, Sociology, Digital Sociology, Web Analytics,
Paper Key Topics: Health, Disaster Response, Network Science, User Behaviour, Social Science, Microblogging,
Author Keywords: Epidemic Intelligence; Medicine 2.0; Twitter.

Social Networks, epidemics, Web Analytics, Health, Epidemiology, Sociology

Network Science?
This work suggest that social media may be used in order to strengthen surveillance of viral outbreaks and epidemics. The authors propose using Twitter as an ‘early warning’ system for detecting disease outbreaks.

Stigmergy in Web 2.0: a Model for Site Dynamics
Authors: Aiden Dipple, Prof. Kerry Raymond, Assoc. Prof. Michael Docherty
Key Topics: Web 2.0, HCI Human Computer Interaction, Psychology, Web Applications, Web Design,
Paper Key Topics: Academic Publishing?
Author Keywords: Guides; conference publication

This work examines what makes a successful Website by gaining insight from biologically I inspired design patterns, with the aim of discovering what site traits contribute to a successful Website.

EventShop: From Heterogeneous Web Streams to Personalized Situation Detection and Control
Authors: Mingyan Gao, Vivek K. Singh, Ramesh Jain
Key Topics: Big Data, Web Analytics, Linked Data, Data Mining,
Paper Key Topics: Real-time data
Author Keywords: None

This paper look at the possibilities for combing data gathered from multiple heterogeneous data sources including physical sensors or data observed by human-sensors. It suggests that when combined data from various sources can provide meaningful information and potentially provide early detection for situations to aid millions of users.

Measuring Tie Strength in Implicit Social Networks
Authors: Mangesh Gupte, Tina Eliassi-Rad
Key Topics: Social Networks, Digital Sociology, Network Theory
Paper Key Topics: Network Science, Online Identity, Online Offline Community
Author Keywords: Social networks, tie strength, axiomatic approach

This work examines the feasibility of measuring strength between friendships on social media sites. The authors utilise event information where users register as participating in the same event, with the aim of measuring the strength of ties based upon mutual event attendance by users.

^ Thoughts - Problem with this is that people can often say that they're going to an event when they're not! Could maybe take into account people being tagged at the event location after the event?

From e-government to Social Network Government: Towards a Transition Model
Authors: Daniel Halpern, James E. Katz
Key Topics: Social Networking, Politics, Web 2.0, Psychology, Web Society, Web Analytics,
Paper Key Topics: User Behaviour, Online Engagement, Content Analysis, Democracy,
Author Keywords: Civic Participation, e-Government, Social Media, Online Deliberation,

Transparency, Open Data and Trust in Government: Shaping the Infosphere
Authors: Kieron O'Hara
Key Topics: Privacy and Trust, Big Data, Linked Data, Digital Sociology,
Paper Key Topics: Open Data, Government, Social Science,
Author Keywords: Transparency; trust; open data; crime data; criminal justice data, rational choice, social capital, deliberative democracy**.**

Dissecting the Butterfly: Representation of Disciplines Publishing at the Web Science Conference Series
Authors: Clare J. Hooper, Nicolas Marie, Evangelos Kalampokis
Key Topics: General Web Science, Education, Visualisation,
Paper Key Topics: Virtual Community, Academic Publishing,
Author Keywords: Web Science discipline, community analysis, bibliometrics, disciplines

Web Science and Human-Computer Interaction: When Disciplines Collide
Authors: Clare J. Hooper, Alan Dix
Key Topics: General Web Science, HCI Human Computer Interaction,
Paper Key Topics: Virtual Community,
Author Keywords: HCI discipline; Web Science discipline; Methodology

Colonel Blotto On Facebook: The Effect of Social Relations On Strategic Interaction
Authors: Pushmeet Kohli, Yoram Bachrach, David Stillwell, Michael Kearns, Ralf Herbrich, Thore Graepel
Key Topics: Social Networking, Web Society, Web Analysis, Psychology, Game Theory, Data Mining,
Paper Key Topics: Gamification, User Behaviour,
Author Keywords: None

Crowd IQ: Measuring the Intelligence of Crowdsourcing Platforms
Authors: Michal Kosinski, Yoram Bachrach, Gjergji Kasneci, Jurgen Van-Gael, Thore Graepel
Key Topics: Quality Management, Performance Management,
Paper Key Topics: Crowdsourcing, Quality Control,
Author Keywords: Crowdsourcing, Psychometrics, Incentive Schemes

Controlling Opinion Bias in Online Social Networks
Authors: Chris J. Kuhlman, V. S. Anil Kumar, S. S. Ravi
Key Topics: Politics, Social Networking, Web Society, Web Analytics, Web Modelling, Digital Sociology,
Paper Key Topics: Government, User Behaviour, Democracy, Algorithms,
Author Keywords: Voter model; Countering Bias; Online Interactions.

When a Friend in Twitter is a Friend in Life
Authors: Ee-Peng Lim, Wei Xie, Cheng Li, Feida Zhu, Xueqing Gong
Key Topics: Social Networking, Web Society, Web Analysis, Digital Sociology,
Paper Key Topics: Online Offline Community, Microblogging, Social Science, Algorithms,
Author Keywords: None

The Configuration of Networked Publics on theWeb: Evidence from the Greek Indignados Movement
Authors: Xuesong Lu, Giorgos Cheliotis, Xiyue Cao, Yi Song, Stephane Bressan
Key Topics: Social Networking, Web Society, Web Analytics, Web History, Web Modelling
Paper Key Topics: Social Science, Communication Science, User Behaviour, Content Analysis, Network Science, Online Engagement, Topic Modelling
Author Keywords: None

Guard Your Connections: Infiltration of a Trust/Reputation Based Network
Authors: Malik Magdon-Ismail, Brian Orecchio
Key Topics: Privacy and Trust, Web Modelling, Web Ontology, Web Analytics, Web Architecture, Networks, Network Theory,
Paper Key Topics: Network Science, Algorithms, User Behaviour,
Author Keywords: Cascade, phase transition, social networks, LinkedIn

Political Speech in Social Media Streams: YouTube Comments and Twitter Posts
Authors: Yelena Mejova, Padmini Srinivasan
Key Topics: Politics, Media, Social Networking, Web Society, Web Analytics,
Paper Key Topics: Sentiment Analysis, Microblogging,
Author Keywords: Social media; Political discourse; Sentiment analysis
How to Translate a Book Within an Hour
Authors: Patrick Minder, Abraham Bernstein
Key Topics: Networks, Web Languages,
Paper Key Topics: Crowdsourcing, Computer Science, Algorithms, Network Science,
Author Keywords: Human Computation, Crowdsourcing, CrowdLang

Containment of Misinformation Spread in Online Social Networks
Authors: Nam P. Nguyen, Guanhua Yan, My T. Thai , Stephan Eidenbenz
Key Topics: Social Networking, Web Modelling, Web Analytics, Networks, Network Theory,
Paper Key Topics: Content Analysis, Quality Control, Microblogging, Virtual Community, Algorithms,
Author Keywords: Misinformation containment, Online social networks

When Daily Deal Services Meet Twitter: Understanding Twitter as a Daily Deal Marketing Platform
Authors: Jaimie Y. Park, Chin-Wan Chung
Key Topics: Data Mining, Social Networking, Marketing, Hypertext, Web Analysis,
Paper Key Topics: Microblogging, Content Analysis,
Author Keywords: Twitter, Daily Deal Service, Social Media Marketing, Electronic Commerce, Electronic Word-of-Mouth (eWOM), Consumer Behavior, Microblogging, Online Social Networking

Designing for Vigilance during Intermittent Use
Authors: M. Giles Phillips,
Key Topics: HCI Human Computer Interaction, Social Networking, Performance Management, Web Design, Psychology, Web Society,
Paper Key Topics: User Behaviour, Online Engagement,
Author Keywords: Vigilance, Intermittent Use, Social Media, Mobile

Fairness on theWeb: Alternatives to the Power Law
Authors: Jerˆ ome Kunegis, Julia Preusse
Key Topics: Web Analytics, Network Theory, Networks, Social Networking, Web Modelling, Web Ontologies, Web Architecture,
Paper Key Topics: Communication Science, Network Science,
Author Keywords: Network analysis, Power-lawexponent, Gini coefficient, Fair- ness, Entropy

TweetLDA: Supervised Topic Classification and Link Prediction in Twitter
Authors: Daniele Quercia, Harry Askham, Jon Crowcroft
Key Topics: Hypertext, Linked Data, Social Networking, Web Analytics, Web Modelling, Web Applications,
Paper Key Topics: Geography, Microblogging, Topic Modelling, Content Analysis,
Author Keywords: None

Loosing "Friends" on Facebook
Authors: Daniele Quercia, Mansoureh Bodaghi, Jon Crowcroft
Key Topics: Social Networking, Psychology, Digital Sociology, Sociology, Web Society, Web Analytics,
Paper Key Topics: Microblogging, Content Analysis, Online-Offline Community, Social Science, Virtual Community, Online Engagement,
Author Keywords: social networks, tie decay, facebook, personality

Behaviour analysis across different types of Enterprise Online Communities
Authors: Matthew Rowe, Miriam Fernandez, Harith Alani, Inbal Ronen, Conor Hayes and Marcel Karnstedt
Key Topics: Social Networking, Web Analysis, Web Society,
Paper Key Topics: Virtual Community, Microblogging, Content Analysis, Network Science,
Author Keywords: Community Analysis, User Behaviour, Enterprise Communities, Web Science
^ Virtual Community the same as online community

Tracing the Birth of an OSN: Social Graph and Profile Analysis in Google+
Authors: Doris Schioberg, Fabian Schneider, Harald Schioberg, Stefan Schmid, Steve Uhlig, Anja Feldmann
Key Topics: Web Graph, Social Networking, Web Society, Networks, Network Theory,
Paper Key Topics: Geography, Demographics, Virtual Community, Network Science,
Author Keywords: Online Social Networks, Measurements,Google+, Dynamic Graphs, Asymmetric and symmetric Links, User Profile

Rumoring During Extreme Events: A Case Study of Deepwater Horizon 2010
Authors: Emma S. Spiro, Jeannette Sutton, Matt Greczek, Sean Fitzhugh, Nicole Pierski, Carter T. Butts
Key Topics: Web Analytics, Web Modelling, Web Society, Social Networking, Psychology,
Paper Key Topics: Disaster Response, Social Science, Microblogging, User Behaviour, Communication Science,
Author Keywords: informal online communication, rumouring, disaster response, microblogging

Tracking Twitter for Epidemic Intelligence
Authors: Ernesto Diaz-Aviles, Avar´ e Stewart
Key Topics: Social Networking, Web Society, Web Analytics, Data Mining, Digital Sociology, Web Modelling,
Paper Key Topics: Health, Crowdsourcing, Algorithms, Microblogging, Network Science, Disaster Response,
Author Keywords: Epidemic Intelligence; Medicine 2.0; Twitter

Buying unlicensed slimming drugs from the Web: a virtual ethnography
Authors: Lisa Sugiura, Catherine Pope, Craig Webber
Key Topics: Digital Ethnography, Research Methodology, Digital Sociology, General Web Science, Web Society,
Paper Key Topics: Health, Social Science, Cybercrime
Author Keywords: Web Science; Criminology; Virtual Ethnography,
^ tagged with cybercrime, because details crime facilitated by the internet – although the crime is not directly damaging the technology, it is still crime facilitated by technology. Technology provides the accessibility to the drugs, as described by the paper.

Coevolution of Network Structure and Content
Authors: Chun-Yuen Teng, Liuling Gong, Avishay Livne EECS, Celso Brunetti, Lada Adamic
Key Topics: Networks, Network Theory, Web Architecture, Web Modelling, Social Networking, Data Mining,
Paper Key Topics: Network Science, Microblogging,
Author Keywords: social media, information networks, network evolution

Mixing Methods and Theory to Explore Web Activity
Authors: Ramine Tinati, Susan Halford, Les Carr, Cathy Pope,
Key Topics: Linked Data, Semantic Web, Research Methodology, General Web Science
Paper Key Topics: Social Science, Open Data,
Author Keywords: Web Science, Methodology, Open Government Data

Pay-less Entity Consolidation – Exploiting Entity Search User Feedbacks for Pay-as-you-go Entity Data Integration
Authors: Thanh Tran, Yongtao Ma, Gong Cheng
Key Topics: Linked Data, Semantic Web, Hypertext, Web Modelling, Web Development,
Paper Key Topics: Algorithms,
Author Keywords: Clustering, entity consolidation, entity search, implicit user feedback, pay-as-you-go data integration

Network, Personality and Social Capital
Authors: Jayant Venkatanathan, Evangelos Karapanos, Vassilis Kostakos, Jorge Gonçalves
Key Topics: Social Networking, Web Society, Web Analytics, Psychology, Digital Sociology, Networks, Network Theory,
Paper Key Topics: User Behaviour, Network Science, Social Science,
Author Keywords: Social networks analysis, ego networks, social capital, personality traits.

MiningWeb Query Logs to Analyze Political Issues
Authors: Ingmar Weber, Venkata Rama Kiran Garimella, Erik Borra
Key Topics: Data Mining, Politics, Digital Sociology, Web Analytics, Web Society, Web Search,
Paper Key Topics: Sentiment Analysis, User Behaviour**,** Demographics,
Author Keywords: web search logs; political leaning; partisanship; opinion mining and sentiment analysis

The Evolution of a Crawling Strategy for an Academic Document Search Engine: Whitelists and Blacklists
Authors: JianWu, Pradeep Teregowda, Juan Pablo Fernandez Ramırez, Prasenjit Mitra, Shuyi Zheng and C. Lee Giles
Key Topics: Data Mining, Web Search, Web Analytics, Semantic Web,
Paper Key Topics: Content Analysis, Academic Publishing,
Author Keywords: Information retrieval; web crawling; search engine

You are where you E-mail: Using E-mail Data to Estimate International Migration Rates
Authors: Emilio Zagheni, IngmarWeber
Key Topics: Digital Ethnography, Digital Sociology, Sociology, Web Analytics, Web Society, Statistical Analysis,
Paper Key Topics: Demographics, Geography, Social Science,
Author Keywords: Demographics, Migration, Mobility, E-mail data

Web Science 2013

This year is alphabetical by Title !!!!

There's no such thing as raw data'. Exploring the sociotechnical life of a government dataset
Tim Davies, Mark Frank
Key Topics: Big Data, Linked Data,
Paper Topics: Government**,** Open Data**,** Web Archiving,
Author Keywords: Open Data, Public Sector Information

A Categorization Scheme for Socialbot Attacks In Online Social Networks
Silvia Mitter, Claudia Wagner, Markus Strohmaier
Key Topics: Social Networking
Paper Key Topics: Socialbot
Author Keywords: Socialbots; attack; Taxonomy; categorization scheme; Twitter; online social networks;
^ WS 2014????? – short paper?

A Comparison between Online and Offline Prayer
Fabian Eikelboom, Paul Groth, Victor de Boer, Laura Hollink
Key Topics: Sociology, Social Networking, Web Society
Paper Key Topics: Religion, Theology
Author Keywords: Practical Theology; Religion on the Web; Prayer.

Aemoo∗ : exploring knowledge on the Web
Key Topics: Semantic Web, Linked Data
Paper Topics: knowledge patterns
Author Keywords: None

^ Useful for me and my thesis?

AltOA: A Framework for Dissemination Through Disintermediation
Richard William Fyson, Simon Coles, Les Carr
Key Topics: Information systems, Information and communications technology. Media, Web Society, Education
Paper Key Topics Open Access, Academic Publishing**,** Web Publishing
Author Keywords: Academic publishing; disintermediation; open access
Distributing academic material via the Web – the Web has provided a new platform upon which to distribute academic materials. Yet the authors claim that the Web has failed to revolutionise academic publishing, instead it simply better facilitates what were existing processes.

An Empirical Analysis of Characteristics of Useful Comments in Social Media
Authors: Elaheh Momeni, Gerhard Sageder
Key Topics: Social Networking, Media, Quality Management
Paper Topics: User behaviour**,** Content analysis
Author Keywords: social media, user-generated comments, usefulness prediction
ACM
This paper examines the feasibility of identifying and potentially utilising useful comments on social media. For example, a descriptive comment on a picture might be used to assign useful information to a picture, such as where the photo was taken. However, it is necessary to distinguish between useful information and irrelevant comments.

An Investigation into Correlations between Financial Sentiment and Prices in Financial Markets
Authors: Paul Gaskell, Frank McGroarty, Thanassis Tiropanis
Key Topics: Business, Economics, Social Media, Data Mining
Paper Topics: eCommerce**,** Content Analysis
Author Keywords: Finance; Text Analysis; Sentiment Analysis.
Can the content of Tweets and other social media posts be used to gauge the health of financial investments and be used to influence potential investments – e.g. is it worth buying shares in a particular company?
^Finance changed to ecommerce for consistence and optimisation

Are User-contributed Reviews Community Property? Exploring the Beliefs and Practices of Reviewers
Authors: Frank M. Shipman, Catherine C. Marshall
Key Topics: Digital Copyright, Social Media, Digital Sociology
Paper Topics: Narrative, User Behaviour, Web Publishing
Author Keywords: Product reviews, information rights, reuse, removal.
ACM
Looks at online reviewers – and their intellectual property?

Art As A Source For Innovation In Knowledge Processing
Authors: Robert Tolksdorf, Markus Luczak-Rosch
Key Topics: Information Systems, Media**,** Sematic Web, Linked Data, Web ontologies**,** Artificial Intelligence,
Paper Topics: Arts and Culture, Knowledge Patterns,
Author Keywords: Knowledge processing, Artists work, Knowledge representation
ACM
Art as an inspiration for innovation in knowledge modelling, as art visualises world views…

Assessing the Educational Linked Data Landscape
Authors: Mathieu d'Aquin, Alessandro Adamou, Stefan Dietze
Key Topics: Education, Linked Data, Networks, Information and Communications Technology, Data Mining, Semantic Web, Web Search,
Paper Topics: Network Science?, MOOCs,
Author Keywords: Computers and Education: Computer Uses in Education
^ Relevant to my research? Education, Linked Data

Automatically Extracting Frames from Media Content using Syntacting Analysis
Sentence structure - mentions the term 'communication science'
Authors: Wouter van Atteveldt, Tamir Sheafer, Shaul Shenhav
Key Topics: Media, Data Mining, Networks, Semantic Web, Research Methods
Paper Topics: Content Analysis, Communication Science, Computer Science
Author Keywords: None

Beyond Positivism in Computer Science
Authors: Markus Luczak-Rösch
Key Topics: Research Methodology, Web Ontologies, Web Society
Paper Topics: Philosophy, Computer Science
Author Keywords: Computer Science; Web Science; Philosophy; Positivism; Modeling; Engineering

Beyond The Trowel's Edge: Provenance-Based Collective Archaeological Interpretations
Provenance
Authors: Michael O. Jewell, Tom Frankland, Enrico Costanza, Graeme Earl, Luc Moreau
Key Topics: Research methodologies, Sematic Web, HCI Human Computer Interaction, Web Applications
Paper Topics: Provenance, archaeology,
Author Keywords: None

BlueFinder: Recommending Wikipedia Links Using DBpedia Properties
Knowledge,
Authors: Diego Torres, Hala Skaf-Molli, Pascal Molli, Alicia Diaz
Key Topics: Semantic Web, Web Search, Web Society, Web Analytics
Paper Topics: Algorithms,
Author Keywords: Information Search and Retrieval: Information filtering, DBpedia, Wikipedia, Recommendation

Can simple social copying heuristics explain tag popularity in a collaborative tagging system?
Authors: Jared Lorince, Peter M. Todd
Key Topics: Social Networking, Computer Science, Web Society,
Paper Topics: Cognitive Science
Author Keywords: Collaborative tagging, folksonomy, decision-making, ecological rationality, heuristics, cognitive science

Challenges and Opportunities of Local Journalism: A Case Study of the 2012 Korean General Election
Authors: Souneil Park, Minsam Ko, Jaeung Lee, Aram Choi, Junehwa Song
Key Topics: Media, Politics, Social Networking, Data Mining
Paper Topics: Journalism**,** Microblogging
Author Keywords:

Collabmap: Crowdsourcing Maps for Emergency Planning
Authors: Sarvapali D. Ramchurn, Bing Shi
Key Topics: Privacy and Trust, Information Systems, Systems Design
Paper Topics: Crowdsourcing**,** Geographic Information Systems (GIS)
Author Keywords: None

Considering People with Disabilities as ¨ Uberusers for Eliciting Generalisable Coping Strategies on the Web
Authors: Markel Vigo, Simon Harper
Key Topics: Accessibility, Human Computer Interaction, Web Society, Web Technologies, Hypertext, Web Search, Web Analytics
Paper Topics: User Behaviour, Health, Communication Science, Algorithms,
Author Keywords: Coping strategies, Web, uberusers, behavioural sciences

Content-Based Similarity Measures of Weblog Authors
Authors: Christopher Wienberg, Melissa Roemmele, and Andrew S. Gordon
Key Topics: Social Networking**,** Psychology, Sociology, Digital Sociology, Web Analytics
Paper Topics: Social Science**,** Online Identity, Communication Science, Content Analysis, User Behaviour, Web Publishing,
Author Keywords: Similarity measures; Weblogs; Personal pronouns
^ Blogging = Web Publishing

Crowd Truth: Harnessing disagreement in crowdsourcing a relation extraction gold standard
Authors: Lora Aroyo, Chris Welty
Key Topics: Quality Management,
Paper Topics: Crowdsourcing, Natural Language Processing, Knowledge Patterns
Author Keywords: Relation Extraction, Gold Standard Annotation

Cyberbullying 2.0: Implications of Web development for the prevention of cyberbullying
Authors: Lambros Lazuras, Michalis Vafopoulos
Key Topics: Social Networking, Psychology, Sociology, Digital Sociology, Web 2.0,
Paper Topics: Cyberbullying, Online/Offline Community, User Behaviour, Virtual Community, Web 3.0,
Author Keywords: Cyberbullying, prevention guidelines, youth, social media
^ First instance of Web 3.0?

Debanalizing Twitter: The Transformation of an Object of Study
Authors: Richard Rogers
Key Topics: Social Networking**,** Research Methodology, Web Analytics, Data Mining, Web 2.0,
Paper Topics: Microblogging, Virtual Community,
Author Keywords: Social Media, Twitter, Digital Methods

Democracy and Trust in the Age of the Social Web
Authors: Gloria Origgi
Key Topics: Privacy and Trust, Social Networking, Psychology**,** Web Society,
Paper Topics: Democracy**,** Social Science, User Behaviour,
Author Keywords: Democracy, Social Networks, Trust, Epistemic Trust

Designing the W3C Open Annotation Data Model
Authors: Robert Sanderson, Paolo Ciccarese, Herbert Van de Sompel
Key Topics: Web Architecture, Linked Data, Hypertext, Semantic Web, Web Languages, Web Technologies, Web 2.0,
Paper Topics:
Author Keywords: Annotation; Web Architecture; Interoperability

Detecting Cyberbullying: Query Terms and Techniques
Authors: April Kontostathis, Kelly Reynolds, Andy Garron, Lynne Edwards
Key Topics: Social Networking, Web Analytics,
Paper Topics: Cyberbullying, Machine Learning, Content Analysis,
Author Keywords: Machine Learning; Cyberbullying Detection; Term Analysis; Latent Semantic Indexing
^ Should machine learning come under AI?
Does the Web Extend the Mind?
Authors: Harry Halpin
Key Topics: Psychology**,** Web Society, Media**,**
Paper Topics: Philosophy, Cognitive Science, Crowdsourcing, User Behaviour, Digital Natives,
Author Keywords: extended mind, philosophy, cognitive integration, Web, distributed cognition, collective intelligence

Don't Worry, Be Happy: The Geography of Happiness on Facebook
Authors: Daniele Quercia
Key Topics: Social Networking, Psychology, Research Methodology, Statistical Analysis, Digital Sociology, Sociology, Web Society,
Paper Topics: Geography**,** Social Science,
Author Keywords: Psychology, geography, quantitative methods, Facebook, statistics

Experiences Surveying the Crowd: Reflections on Methods, Participation, and Reliability
Authors: Catherine C. Marshall, Frank M. Shipman
Key Topics: Research Methodology, Media, Social Networking, Data Mining, Statistical Analysis, Web Analytics, Quality Management,
Paper Topics: Crowdsourcing**,** demographics, Content Analysis
Author Keywords: Crowdsourcing; surveys; demographics; reliability
MTurk – employment?

Filling the Gaps Among DBpedia Multilingual Chapters for Question Answering
Authors: Julien Cojan, Elena Cabrio and Fabien Gandon
Key Topics: Linked Data, Semantic Web, Artificial Intelligence
Paper Topics: Natural Language Processing, Web Publishing,
Author Keywords: Linked Data, DBpedia, Property Alignment, Question Answering

From Information Delivery to Interpretation Support: Evaluating Cultural Heritage Access on the Web
Authors: Chiel van den Akker, Marieke van Erp, Lora Aroyo, Ardjan van Nuland, Lourens van der Meij, Susan Legene, Guus Schreiber
Key Topics: Web History, Web Society, Media, Web Applications, Human Computer Interaction,
Paper Topics: Arts and Culture, Philosophy,
Author Keywords: Digital Hermeneutics, online cultural heritage, evaluation, Agora, events

From networked publics to issue publics: Reconsidering the public/private distinction in web science
Authors: Andreas Birkbak
Key Topics: General Web Science, Privacy and Trust, Social Networking, Networks, Security, Risk management, Digital Sociology, Politics,
Paper Topics: Philosophy, Open Access,
Author Keywords: networked publics; issue publics; public/private; social media; democracy; Facebook, Dewey

Head Start: Improving Academic Literature Search with Overview Visualizations based on Readership Statistics
Authors: Peter Kraker, Christoph Trattner, Kris Jack, Stefanie Lindstaedt, Christian Schlogl
Key Topics: Web Society, Visualisation, Digital Sociology, Statistical Analysis, Education,
Paper Topics: Academic Publishing**,** User Behaviour, Content Analysis, bibliometrics
Author Keywords: literature search; co-readership analysis; library statistics; interactive map; educational technology; alternative metrics
^ Looking at subject/discipline/curriculum similarity
^ bibliometrics = new, is it needed? Does it occur anywhere else?

How confident are you? Worker classification in Crowdsourcing
Authors: Yuko Sakurai, Masaaki Oka, Satoshi Oyama, Makoto Yokoo
Key Topics: Human Computer Interaction,
Paper Topics: Quality Control, Crowdsourcing, User Behaviour,
Author Keywords: Crowdsourcing; Human computation; Mechanism design

Identifying Research Talent Using Web-Centric Databases
Authors: Anca Dumitrache, Paul Groth, Peter van den Besselaar
Key Topics: Web Graph, Data Mining, Web Search,
Paper Topics: bibliometrics, Academic Publishing, Social Science, Computer Science, Knowledge Patterns,
Author Keywords: Scholarly networks, bibliometrics, altmetrics, online vs. offline databases, independence indicators
Bibliometrics = an author topic here!

Location Tracking via Social Networking Sites
Authors: Lisa Thomas, Pam Briggs, Linda Little
Key Topics: Web Society, Web Analytics, Social Networking, Psychology, Mobile Web,
Paper Topics: Geography, Geo-tagging, User Behaviour,
Author Keywords: Location- Based Services; Location-Tracking; Social Networking; Disclosure; Trust; Intention

Market-based SPARQL Brokerage: Economic Incentives for Linked Data
Authors: Mengia Zollinger, Cosmin Basca, Abraham Bernstein
Key Topics: Linked Data, Big Data, Semantic Web, Marketing, Web Languages, Web Analysis, Systems Design,
Paper Topics: ecommerce, Open Data, RDF,
Author Keywords: Web of Data, Free and Commercial Data, Data Markets
^ Semantic Web – covers lower level topics such as RDF – RDF is now also a topic!

Measuring Social Media Quotations in Journalism
Authors: J. Nathan Matias
Key Topics: Social Networking, Web Analytics, Web Society, Media,
Paper Topics: Content Analysis, Journalism, Microblogging,
Author Keywords: Twitter; Journalism; Content Analysis; Social Media; Curation; Quotation; Ecosystems; Text Analysis; Arab Spring
^ Content Analysis also covers Text Analysis,

Mechanical Turk as an Ontology Engineer?
Authors: Natalya F. Noy, Jonathan Mortensen, Paul R. Alexander, Mark A. Musen
Key Topics: Web Ontologies, Semantic Web, Quality Management,
Paper Topics: Crowdsourcing,
Author Keywords: Semantic Web, ontology, human computation, crowdsourcing, Amazon Mechanical Turk

Mining User Behaviours: A Study of Check-in Patterns in Location Based Social Networks
Authors: Daniel Preo¸ tiuc-Pietro, Trevor Cohn
Key Topics: Social Networking, Data Mining, Web Analysis, Web Society, Psychology, Digital Sociology,
Paper Topics: Geography, Geo-tagging, User Behaviour,
Author Keywords: Social networks, Location Based Social Networks, Foursquare, Mobility patterns, Clustering, User behaviour, User movement prediction, Data mining

Modeling Movements in Oil, Gold, Forex and Market Indices using Search Volume Index and Twitter Sentiments
Authors: Tushar Rao, Saket Srivastava
Key Topics: Business, Social Networking, Web Modelling, Statistical Analysis, Web Analytics,
Paper Topics: Sentiment Analysis, ecommerce, Microblogging,
Author Keywords: Stock market; sentiment analysis; Twitter; microblogging; social network analysis; oil; gold; forex

On Measuring the Impact of Hyperlinks on Reading
Authors: Gemma Fitzsimmons, Mark Weal, Denis Drieghe
Key Topics: Hypertext, General Web Science, Psychology, HCI Human Computer Interaction,
Paper Topics: Cognitive Science,
Author Keywords: Hyperlinks, Reading, Web Science, Psychology, Human Computer Interaction, Eye movements, Visual cognition
^ Hyperlinks = hypertext

Personality Traits and Microblogging Behavior of Weibo Users: Onlies versus Others
Authors: Dong Nie, Ang Li, Bibo Hao, Tingshao Zhu
Key Topics: Psychology, Social Networking, Web Society, Digital Sociology,
Paper Topics: Microblogging, User Behaviour, Social Science,
Author Keywords: only child, personality, microblogging behavior, significant difference

Petition Growth and Success Rates on the UK No. 10 Downing StreetWebsite
Authors: Scott A. Hale, Helen Margetts, Taha Yasseri
Key Topics: Big Data, Politics, Digital Sociology, Web Society,
Paper Topics: Online Engagement, Democracy, Government,
Author Keywords: petition; mobilization; trace data; big data; leptokurtic; bursty growth

Preferential Attachment in Online Networks: Measurement and Explanations
Authors: Jerome Kunegis, Marcel Blattner, Christine Moser
Key Topics: Big Data, Networking, Network Theory, Web Analytics, General Web Science, Web Architecture, Web Modelling, Data Mining,
Paper Topics: Network Science,
Author Keywords: Network analysis; preferential attachment

Producing a Unified Graph Representation from Multiple Social
Authors: Derek Greene, Padraig Cunningham,
Key Topics: Web Analytics, Data Mining, Web Graph, Social Networking, Web 2.0, Web Architecture, Visualisation, Networks, Network Theory,
Paper Topics: Microblogging, Demographics, Network Science,
Author Keywords: Social network analysis, Data integration, Social media
^ Network Theory and Web Graph the same thing? Or at least too similar?

R-energy for Evaluating Robustness of Dynamic Networks
Authors: Ming Gao, Ee-Peng Lim, David Lo
Key Topics: Networks, Network Theory, Web Graph,
Paper Topics: Network Science, Microblogging,
Author Keywords: R-energy; network robustness; normalized Laplacian matrix

Rethinking Measurements Of Social Media Use By Charities: A Mixed Methods Approach
Authors: Christopher Phethean, Thanassis Tiropanis, Lisa Harris
Key Topics: Social Networking, Web Society, Web Analytics, Research Methodology, General Web Science, Marketing,
Paper Topics: Content Analysis, Online Engagement,
Author Keywords: Social media; charities; marketing; web science

Reverse Privacy Engineering
Authors: Julien Pierre
Key Topics: Privacy and Trust, Social Networking, Web Society, Sociology, Digital Sociology, Web Analytics, Semantic Web, Web 2.0,
Paper Topics: Communication Science, User Behaviour, Content Analysis,
Author Keywords: Privacy, social process, web-service, regulation, personal data, computer-mediated communication, privatory framework, metadata, browser

Semantic Tagging on Historical Maps
Authors: Bernhard Haslhofer, Werner Robitza, Carl Lagoze, Francois Guimbretiere
Key Topics: Linked Data, Semantic Web, Social Networking, Web Society, Web Analytics, Digital Sociology, Digital Ethnography,
Paper Topics: Geo-tagging, Social Science, User Behaviour, Content Anlsysis,
Author Keywords: Tagging, Linked Data, Digital Humanities
^ Tagging, not just geo-tagging – Digital Humanities – covers digital soc and others!

Sentiment and Topic Analysis on Social Media: A Multi-Task Multi-Label Classification Approach
Authors: Shu Huang, Wei Peng , Jingxuan Li, Dongwon Lee
Key Topics: Social Networking, Web Society, Web Analytics, Marketing, Data Mining
Paper Topics: Sentiment Analysis, Topic Modelling, User Behaviour, Content Analysis, Microblogging, Algorithms,
Author Keywords: multi-task; multi-label; classification; sentiment analysis; topic analysis

Simultaneously Detecting Fake Reviews and Review Spammers using Factor Graph Model
Authors: Yuqing Lu, Lei Zhang, Yudong Xiao, and Yangguang Li
Key Topics: Web Graph, Quality Management, Data Mining, Web Analytics, Linked Data, Web Modelling, Network Theory, Networks,
Paper Topics: Network Science, Content Analysis, Algorithms, ecommerce, User Behaviour,
Author Keywords: opinion spam, fake review, factor graph

Social Media as a Measurement Tool of Depression in Populations
Authors: Munmun De Choudhury, Scott Counts, Eric Horvitz
Key Topics: Social Networking, Psychology, Digital Sociology, Web Society, Web Analytics, Data Mining, Big Data, Statistical Analysis
Paper Topics: User Behaviour, Social Science, Health, Crowdsourcing, Sentiment Analysis, Micorblogging,
Author Keywords: behavior, depression, emotion, health, language, social media, mental health, public health, Twitter, wellness

Socio-Technical Transitions Pathways for UK Open Government Data
Authors: Chris J. Martin, Tim G. Davies, Jo Bates
Key Topics: Web Society, Web Analytics, General Web Science, Linked Data, Big Data,
Paper Topics: Government, Web Archiving, Online Engagement, Open Data,
Author Keywords: Open government data; socio-technical transitions; pathways

Sprint Methods for Web Archive Research
Authors: Hugo C. Huurdeman, Anat Ben-David, Thaer Sammar
Key Topics: Research Methodology, Web History, Web Society, Media, Big Data, Data Mining, Web Analytics,
Paper Topics: Web Archiving, Social Science, Computer Science,
Author Keywords: Web Archives; Digital Methods; Web History; Temporal Analysis; News Analysis; Information Retrieval; Search Interface; Web Collections
^ Refers to data retrieval, is this necessarily the same as data mining?

Studying Facebook via Data Extraction: The Netvizz Application
Authors: Bernhard Rieder
Key Topics: Research Methodology, Social Networking, Media, Data Mining, Big Data, Web Analytics, Privacy and Trust,
Paper Topics: Content Analysis, Algorithms,
Author Keywords: research tool, social networking services, Facebook, data extraction, social network analysis, media studies
^ Mentions Web Crawler, not tagged with 'web society' cos studies the data extraction process not the people

The Performativity of Data: Re-conceptualizing the Web of Data
Authors: Marie Joan Kristine Gloria, Dominic DiFranzo, Marco Fernando Navarro, Jim Hendler
Key Topics: Big Data, Linked Data, Semantic Web, General Web Science, Web Analytics, Web Society,
Paper Topics: Content Analysis, RDF, Open Data,
Author Keywords: Experimental methods, Semantic Web, Web Science Theory

The Rise and the Fall of a Citizen Reporter
Authors: Panagiotis Metaxas, Eni Mustafaraj
Key Topics: Web Society, Web Analytics, Social Networking, Media, Digital Sociology, Sociology, General Web Science,
Paper Topics: Virtual Community, Journalism, Online Engagement, Crowdsourcing, Microblogging, Disaster Response, User Behaviour, Algorithms,
Author Keywords: webscience; social computing; social media; citizen reporters; civic media; crisis informatics; crowdsourcing; news; drug war; microblogging; narcotweets; Twitter; Mexico.
^ Disaster Management = crisis informatics?

The Utility of Social and Topical Factors in Anticipating Repliers in Twitter Conversations
Authors: Johannes Schantl, ClaudiaWagner, Rene Kaiser, Markus Strohmaier
Key Topics: Social Networking, Web Society, Web Analytics, Psychology, Media, Data Mining,
Paper Topics: Microblogging, Communication Science, Online Engagement, User Behaviour, Content Analysis,
Author Keywords: Twitter, social media communication, reply behavior, reply prediction

The Web Science Curriculum at work: The Digital Economy Master Program at USJ-Beirut
Authors: Stéphane B. Bazan, Michalis Vafopoulos
Key Topics: General Web Science, Economics, Education,
Paper Topics: ecommerce,
Author Keywords: Economics; Higher Education; Pedagogy; Web; Digital economy

There's no such thing as raw data'. Exploring the socio- technical life of a government dataset
Authors: Tim Davies, Mark Frank
Key Topics: Big Data, Linked Data, Data Mining, Web Society, Web Analytics,
Paper Topics: Web Archiving, Government, Content Analysis, Open Data,
Author Keywords: Open Data, Public Sector Information

Toward a Next Generation of Network Models for the Web
Authors: Hans Akkermans, Rena Bakhshi
Key Topics: Data Mining, Network, Network Theory, Web Modelling,
Paper Topics: Network Science,
Author Keywords: Power law; dynamic network models; nonlinear preferential attachment; degree distributions; Wikipedia hyperlink network
^ Data Gathering not data mining? – Power laws?
Includes Web Science under ACM classification! – is it now?!

Toward Google Borders
Authors: Antoine Mazieres, Samuel Huron
Key Topics: Web Search, Web Analysis, Web Society, Digital Sociology, Data Mining,
Paper Topics: Social Science, Geography, Computer Science, Content Analysis,
Author Keywords: Digital humanities. Digital Studies. Human Factors. Cultural Trends. Auto completion. Suggestion. Web.
^Digital humanities – data collection again-

Towards A Redefinition of Time in Information Networks?
Authors: Sebastien Heymann, Bedicte Le Grand
Key Topics: Network Theory, Networks, Web Applications, Social Networking, Data Mining, Web Graph, Web Modelling, Web Ontologies,
Paper Topics: Network Science, Web Archiving,
Author Keywords: Time; dynamics; measurement; complex networks; social network; sliding window.
^ Web Repositories = Git Hub, is that 'web archiving'?

Traditional media seen from social media
Authors: Jisun An, Daniele Quercia, Meeyoung Cha, Krishna Gummadi, Jon Crowcroft
Key Topics: Social Networking, Data Mining, Visualisation, Web Society, Web Analytics, Media,
Paper Topics: Microblogging, Content Analysis,
Author Keywords: Social media, Media study, Visualization, Structural hole

Uncovering the Wider Structure of Extreme Right Communities Spanning Popular Online Networks
Authors: Derek O'Callaghan, Derek Greene, Maura Conway, Joe Carthy, Padraig Cunningham
Key Topics: Social Networking, Web Analytics, Web Society, Data Mining, Sociology, Digital Sociology,
Paper Topics: Network Science, Content Analysis, User Behaviour, Microblogging, Social Science,
Author Keywords: Social network analysis; extreme right; heterogeneous online networks.

Understanding The Impact Of Socialbot Attacks In Online Social Networks
Authors: Silvia Mitter, Claudia Wagner, Markus Strohmaier
Key Topics: Social Networking, Web Analytics, Web Society, Web Graph,
Paper Topics: User Behaviour, Cybercrime, Microblogging, Socialbot,
Author Keywords: socialbots; attack; Twitter; online social networks;

Unpicking the Privacy Paradox: Can Structuration Theory Help to Explain Location-Based Privacy Decisions?
Authors: Aristea M. Zafeiropoulou, David E. Millard, Craig Webber, Kieron O' Hara
Key Topics: Social Networking, Web 2.0, Privacy and Trust, Web Analytics, Web Society, Psychology,
Paper Topics: Geography, Geo-tagging, User Behaviour, Content Analysis, Online Engagement,
Author Keywords: Location data; privacy paradox; privacy trade-off; structuration
^ User motivation,

Unveiling the link between logical fallacies and Web Persuasion
Authors: Antonio Lieto, Fabiana Vernero
Key Topics: HCI Human Computer Interaction, Web Technologies
Paper Topics: ecommerce, Content Analysis,
Author Keywords: Web persuasion; logical fallacies; captology; human computer interaction,

Voice-based Web access in rural Africa
Authors: Nana Baah Gyan, Victor de Boer, Anna Bon, Chris van Aart, Hans Akkermans, Stephane Boyera, Max Froumentin, Aman Grewal, Mary Allen
Key Topics: Semantic Web, HCI Human Computer Interaction, Accessibility, Mobile Web, Web Development, Web Society,
Paper Topics: Geography, Demographics,
Author Keywords: Web access; Multi-modality; Developing countries; Voice-based interfaces

Web Science and Human-Computer Interaction: When Disciplines Collide
Authors: Clare J. Hooper, Alan Dix,
Key Topics: General Web Science, HCI Human Computer Interaction, Research Methodology,
Paper Topics: Virtual Communitiy, User Behaviour,
Author Keywords: HCI discipline; Web Science discipline; Methodology

Web Science for Ancient History: Deciphering Proto-Elamite Online
Authors: Terhi Nurmikko, Dr Jacob Dahl, Dr Kirk Martinez, Dr Graeme Earl
Key Topics: General Web Science, Web History, Web Society, Digital Sociology,
Paper Topics: Social Science, Virtual Community, Online Engagement,
Author Keywords: Digital heritage, citizen science, paleography, proto- Elamite, archaeology, decipherment, Elam, Iran, Reflectance Transformation Imaging (RTI)

When Politicians Tweet: A Study on the Members of the German Federal Diet
Authors: Mark Thamm, Arnim Bleier
Key Topics: Social Networking, Web Analytics, Web Society, Politics,
Paper Topics: Microblogging, Government, Sentiment Analysis,
Author Keywords: German Parliament; Political Communication; Microblogging; Twitter; Time Series; Sentiment Analysis

Who Wants To Get Fired?
Authors: Ricardo Kawase, Bernardo Pereira Nunes, Eelco Herder, Wolfgang Nejdl, Marco Antonio Casanova
Key Topics: Web Analytics, Web Society, Privacy and Trust, Social Networking,
Paper Topics: Microblogging, User Behaviour,
Author Keywords: Twitter, privacy awareness, user issues

Why don't we trust health websites that help us help each other? An analysis of online peer-to-peer healthcare
Authors: Elizabeth Sillence, Claire Hardy, Pam Briggs
Key Topics: Web Analytics, Web Society, Privacy and Trust, Quality Control,
Paper Topics: Health, Online Advertising, User Behaviour,
Author Keywords: eHealth, health informatics, trust, credibility, advertising, health, patient experience, patient communities, smoking cessation

Why Forums? An Empirical Analysis into the Facilitating Factors of Carding Forums
Authors: Michael Yip, Nigel Shadbolt, Craig Webber
Key Topics: Social Networking, Web Analytics, Web Society, Web 2.0,
Paper Topics: Cybercrime,
Author Keywords: cybercrime; carding; underground economy; social computing; web 2.0

Why individuals seek diverse opinions (or why they don't)
Authors: Jisun An, Daniele Quercia, Jon Crowcroft
Key Topics: Privacy and Trust, Quality Management, Psychology, Web Applications, Web Society, Social Networking, Digital Sociology, Web Analytics,
Paper Topics: User Behaviour, Microblogging, Social Science,
Author Keywords: None

Web Science 2014 Conference

SESSION: Keynote addresses

Observing the web
Wendy Hall
Key Topics: General Web Science
Paper Key Topics:
Author Keywords: Keynote Talk; Computers and Society

Web science: how is it different?
Daniel Tunkelang
Key Topics: General Web Science
Paper Key Topics: Scientific Method
Author Keywords: Scientific Method, Web Science.

The global war for internet governance
Laura DeNardis
Key Topics: Web Governance, Internet Protocols, Semantic Web, Privacy and Trust, Social Networking
Paper Key Topics: None
Author Keywords: Internet governance; Internet protocols; domain name system; interconnection

SESSION: Session 1: methods (full papers)

Translating surveys to surveillance on social media: methodological challenges & solutions
Chao Yang, Padmini Srinivasan
Key Topics: Research Methodology, Social Networking, Web Society, Digital Sociology, Web Analytics, Data Mining
Paper Key Topics: Social Science, Content Analysis,
Author Keywords: Life Satisfaction, Information Retrieval, Crowdsourcing

Rolling through tumblr: characterizing behavioral patterns of the microblogging platform
Jiejun Xu, Ryan Compton, Tsai-Ching Lu, David Allen
Key Topics: Social Networking, Research Methodology, Psychology, Web Analytics, Data Mining
Paper Key Topics: Microblogging**,** Narrative, Geo-tagging, Content Analysis, User behaviour,
Author Keywords: Tumblr, Online Social Network, Quantitative Methods, Location-based Patterns
^ Add geo-tagging as a keyword?

Identifying and analyzing researchers on twitter
Asmelash Teka Hadgu, Robert Jäschke
Key Topics: Social Networking, Web Analytics, Web Society, Digital Sociology, Data Mining
Paper Key Topics: Microblogging, Computer Science, Demographics, Content Analysis, Algorithms,
Author Keywords: Twitter; Computer Science; Classification; Social Network

Twitter: who gets caught? observed trends in social micro-blogging spam
Abdullah Almaatouq, Ahmad Alabdulkareem, Mariam Nouh, Erez Shmueli, Mansour Alsaleh, Vivek K. Singh, Abdulrahman Alarifi, Anas Alfaris, Alex (Sandy) Pentland
Key Topics: Social Networking, Web Analytics, Digital Sociology, Psychology, Data Mining
Paper Key Topics: Microblogging, Content Analysis, User behaviour,
Author Keywords: Spam; Online Social Networks; Microblogging; Account Abuse

Have coded all the papers which include studies which do data mining – even though the topic is not data mining – should this be the case?

Virtual community and online offline community – are they the same thing?

SESSION: Session 2: geographies (full papers)

The impact of visual attributes on online image diffusion

Luam Catao Totti, Felipe Almeida Costa, Sandra Avila, Eduardo Valle, Wagner Meira, Jr., Virgilio Almeida

Key Topics: Social Networking, Computer Graphics, Data Mining, Psychology,

Paper Key Topics: Content Analysis, Algorithms,

Author Keywords: Content diffusion, popularity prediction, image popularity.

The new blocs on the block: using community forums to foster new neighbourhoods

Elizabeth M. Daly, Dominik Dahlem, Daniele Quercia

Key Topics: Web Society, Social Networking, Web Analytics, Data Mining, Psychology, Big Data, Digital Sociology,

Paper Key Topics: Online Offline Community, Virtual Community, User behaviour,

Author Keywords: Online Communities; Social Capital; Citizen Engagement

Good analysis of topics and keywords – refer to diagram, note presentation of results (p57)

Mapping the UK webspace: fifteen years of british universities on the web

Scott A. Hale, Taha Yasseri, Josh Cowls, Eric T. Meyer, Ralph Schroeder, Helen Margetts

Key Topics: Big data, Web Analytics, Linked Data, Web Architecture, Networks, Network Theory,

Paper Key Topics: Virtual Community, Content Analysis,

Author Keywords: Web Archives; WorldWideWeb; Network Analysis; Hyper- link Analysis; Big Data; Academic Web

Country-level spatial dynamics of user activity: a case study in location-based social networks

Anh Le, Konstantinos Pelechrinis, Prashant Krishnamurthy

Key Topics: Social Networking, Network Theory, Data Mining, Web Analytics, Big Data, Psychology, Digital Sociology,

Paper Key Topics: Geography, Geo-tagging, User Behaviour, Content Analysis,

Author Keywords: Spatial dynamics; Location-based social networks

SESSION: Session 3: engagements (full papers)

Evolution of online user behavior during a social upheaval

Onur Varol, Emilio Ferrara, Christine L. Ogan, Filippo Menczer, Alessandro Flammini

Key Topics: Social Networking, Politics, Psychology, Digital Sociology, Web Analytics, Data Mining, Web Society,

Paper Key Topics: Microblogging, Democracy, Geo-tagging, Online Offline Community, User Behaviour, Content Analysis,

Author Keywords: Social media analysis, social protest, political mobilization, online user behaviour

"I always feel it must be great to be a hacker!": the role of interdisciplinary work in social media research
Katharina Kinder-Kurlanda, Katrin Weller
Key Topics: Research Methodology, Social Networking, Sociology, Digital Sociology, Psychology, Web Analytics,
Paper Key Topics: Social Science, Computer Science,
Author Keywords: Interdisciplinarity, research methods, collaboration, social science, computational social science, social media, qualitative research, interviews.

Multilinguals and Wikipedia editing
Scott A. Hale
Key Topics: Web Society, Social Networking, Sociology, Digital Sociology,
Paper Key Topics: Geography**,** Web Publishing,
Author Keywords: Social Media; Information Discovery; Social Network Analysis; Information Diffusion; Cross-language; Wikipedia; Multilingual

Motivating online engagement and debates on energy consumption
Lara Schibelsky Godoy Piccolo, Harith Alani, Anna De Liddo, Cecília Baranauskas
Key Topics: Web Society, Psychology, Sociology, Digital Sociology, Social Networking,
Paper Key Topics: Environmental Science**,** Government**,** Online engagement
Author Keywords: Engagement, motivation, energy awareness, online debate.

Environmental Science – a significant subject area not yet represented in the keywords list. (check:) Number of papers about green IT and how the Web can be used to promote environmental awareness, but the Web itself can also be used to have a positive impact on the environment. (E.g. reducing the need for physical commuting)
Online engagement – not a subject or science in its own right, but a notable feature worth recording, because obtaining user participation about a real life topic via the Web is something which is becoming more common and does have an effect on society.

SESSION: Session 4: networks (full papers)

Graph structure in the web: aggregated by pay-level domain
Oliver Lehmberg, Robert Meusel, Christian Bizer
Key Topics: Networks, Network Theory, Web Analytics, Data Mining, Web Graph, Linked Data,
Paper Key Topics: Content Analysis, Network Science,
Author Keywords: World Wide Web, Web Graph, Network Analysis, Graph Analysis, Web Mining, Web Science

"Supertagger" behavior in building folksonomies
Jared Lorince, Sam Zorowitz, Jaimie Murdock, Peter M. Todd
Key Topics: Social Networking, Psychology, Digital Sociology,
Paper Key Topics: Folksonomy**,** Crowdsourcing, Online Engagement,
Author Keywords: Collaborative tagging, Folksonomy, Supertaggers

Reading the source code of social ties
Luca Maria Aiello, Rossano Schifanella, Bogdan State
Key Topics: Digital Sociology, Social Networking, Linked Data, Web Analytics,
Paper Key Topics: Content Analysis, Communication Science,
Author Keywords: Computational sociology; social exchange; domains of interaction; aNobii; Flickr

Centrality rankings in multiplex networks
Albert Solé-Ribalta, Manlio De Domenico, Sergio Gómez, Alex Arenas
Key Topics: Networks, Network Theory, Social Networking, Web Analytics,
Paper Key Topics: Network Science, Microblogging, Content Analysis,
Author Keywords: Betweenness centrality; Multiplex networks; Multilayer net- works

Folksonomy – justification? – difference between that and social machine?

- Need to go through tagging networks and network theory for all that have network science?
- Web Archiving and folksonomy are under tagged?

SESSION: Session 5: interactions (short papers)

Noticing the other gender on Google+
Diego Couto de Las Casas, Gabriel Magno, Evandro Cunha, Marcos André Gonçalves, César Cambraia, Virgilio Almeida
Key Topics: Social Networking, Digital Sociology, Psychology, Web Analytics, Web Society,
Paper Key Topics: Content Analysis, User Behaviour, Online Identity, Social Science
Author Keywords: Online social networks; Gender issues; Google+

Latent dirichlet allocation: stability and applications to studies of user-generated content
Sergei Koltcov, Olessia Koltsova, Sergey Nikolenko
Key Topics: Web Analytics, Research Methodology, Web Society, Data Mining, Linked Data,
Paper Key Topics: User Behaviour, Content Analysis, Web Publishing, Social Science,
Author Keywords: Latent Dirichlet Allocation, topic modelling, social analysis

Do ordinary bloggers really differ from blog celebrities?
Olessia Koltsova, Sergei Koltcov, Svetlana Alexeeva
Key Topics: Web Society, Web Analytics, Psychology, Digital Sociology,
Paper Key Topics: User behaviour, Web Publishing, Content Analysis, Social Science
Author Keywords: Live Journal, blogger, public opinion, topic modelling.

Multimodal communication on tumblr: "i have so many feels!"
Elli Bourlai, Susan C. Herring
Key Topics: Social Networking, Media, Computer Graphics, Psychology, Web Analytics, Digital Sociology,
Paper Key Topics: Content Analysis, User Behaviour, Web Publishing,
Author Keywords: Communication, GIF, image analysis, meme, multimodality, sarcasm, sentiment, social media.

Friends you haven't met yet: a documentary short film
Jesse Vigil, Asa Shumskas Tait, Christopher Wienberg, Andrew S. Gordon
Key Topics: Social Networking, Psychology, Digital Sociology, Web Society, Web Analytics,
Paper Key Topics: Content Analysis, User Behaviour, Online Engagement,
Author Keywords: Professional Ethics – Ethical dilemmas.

SESSION: Session 6: activities (short papers)

Challenging social media analytics: web science perspectives
Ramine Tinati, Olivier Phillipe, Catherine Pope, Les Carr, Susan Halford
Key Topics: Social Networking, Web Analytics, Web Society, General Web Science, Research Methodology, Data Mining, Big Data,
Paper Key Topics: Content Analysis, Social Science, Computer Science, Microblogging,
Author Keywords: Social Theory, Social Media, Twitter, Methodology, Interdisciplinarity

How "big vs" dominate chinese microblog: a comparison of verified and unverified users on sina weibo
Ning Wang, James She, Junting Chen
Key Topics: Social Networking, Web Analytics, Psychology,
Paper Key Topics: Microblogging, User Behaviour, Content Analysis, Topic Modelling, Network Science,
Author Keywords: Sina Weibo; Microblog; Verification; Online Social Network.

An activity-based information-theoretic annotation of social graphs
Arun V. Sathanur, Vikram Jandhyala
Key Topics: Web Graph**,** Social Networking, Web Society, Sociology, Digital Sociology, Linked Data, Web Analytics, Psychology, Networks, Network Theory,
Paper Key Topics: Content Analysis, User Behaviour, Network Science,
Author Keywords: Time Series, Transfer Entropy, Information Theory, Social Networks, Causality, Directed Influence, Delay Distribution

Detecting and forecasting domestic political crises: a graph-based approach
Yaser Keneshloo, Jose Cadena, Gizem Korkmaz, Naren Ramakrishnan
Key Topics: Web Graph, Politics, Web Analytics, Data Mining, Digital Sociology, Research Methodology, Networks, Network Theory,
Paper Key Topics: Social Science, Government, Network Science,
Author Keywords: GDELT, event forecasting, graph mining, domestic political crises.

Pelagios and the emerging graph of ancient world data
Leif Isaksen, Rainer Simon, Elton T.E. Barker, Pau de Soto Cañamares
Key Topics: Linked Data, Web History, Big Data, Web Analytics,
Paper Key Topics: Geographic Information Systems, Geo-tagging, Content Analysis, Arts and Culture, Crowdsourcing, Open Access, Open Data,
Author Keywords: Linked Open Data; Humanities; Geospatial, Social Science,

SESSION: Session 7: content (full papers)

It's all in the content: state of the art best answer prediction based on discretisation of shallow linguistic features
George Gkotsis, Karen Stepanyan, Carlos Pedrinaci, John Domingue, Maria Liakata
Key Topics: Social Networking, Quality Management, Web Analytics,
Paper Key Topics: Online Engagement, Content Analysis, Real-time Data, Quality Control, Crowdsourcing,
Author Keywords: Community Question Answering, Social Media

Skim reading: an adaptive strategy for reading on the web
Gemma Fitzsimmons, Mark J. Weal, Denis Drieghe
Key Topics: Hypertext, Psychology, General Web Science, Web Society, Web Search, Human Computer Interaction,
Paper Key Topics: User Behaviour, Digital Literacy,
Author Keywords: Hyperlinks; Reading; Skim reading; Web Science; Psychology; Human Computer Interaction; Eye movements

Towards tracking and analysing regional alcohol consumption patterns in the UK through the use of social media
Daniel Kershaw, Matthew Rowe, Patrick Stacey
Key Topics: Social Networking, Web Society, Psychology, Research Methodology, Web Analytics, Digital Sociology,
Paper Key Topics: Microblogging, User Behaviour, Content Analysis, Health, Social Science
Author Keywords: Twitter, SNS, Keyword Analysis, Alcohol, Trend Detection

Mining and comparing engagement dynamics across multiple social media platforms
Matthew Rowe, Harith Alani
Key Topics: Social Networking, Web Society, Web Analytics, Data Mining, Psychology, Digital Sociology, Linked Data,
Paper Key Topics: Online Engagement, User Behaviour, Microblogging, Open Data
Author Keywords: Social Media, Engagement, Data Mining

Web Science 2015

http://dl.acm.org/citation.cfm?id=2786451&picked=prox

SESSION: Politics and Culture

Sustainability Implications of Open Government Data: A Cross-Regional Study
Alison Koczanski, Marta Sabou
Key Topics: Big Data, Linked Data, Web Analytics, Web Society, Web Technologies,
Paper Key Topics: Open Data, Web Archiving, Government, Geography, Online Engagement, User Behaviour,
Author Keywords: Open Government Data, Sustainability, User Studies, Survey
^ Web Archiving because data must be stored in repository
Government and politics aren't always the same thing like here… even though the session says Politics and culture, should all the papers be tagged with it?

Unveiling the Political Agenda of the European Parliament Plenary: A Topical Analysis
Derek Greene, James P. Cross
Key Topics: Web Analytics, Web Society, Politics, Data Mining,
Paper Key Topics: Topic Modelling, Government, Content Analysis,
Author Keywords: Topic modelling, Text mining, Political speech, EU politics
^ Specifically mentions data mining

Mining cross-cultural relations from Wikipedia: A study of 31 European food cultures
Paul Laufer, Claudia Wagner, Fabian Flöck, Markus Strohmaier
Key Topics: Data Mining, Web Analytics, Web Society, Digital Sociology,
Paper Key Topics: Computer Science, Algorithms, Social Machine, Social Science, Content Analysis, User Behaviour, Geography,
Author Keywords: None
^ Wikipedia – social machine?

Information and Communication Technologies (ICTs) and Peacebuilding: a Conceptual Framework
Jennifer R. Welch, Susan Halford, Mark Weal
Key Topics: Information and Communications Technology, Web Society, Web Analysis, Media, Digital Sociology, Politics,
Paper Key Topics: Geographic Information Systems GIS, Geo-tagging, Geography, Social Science,
Author Keywords: Information and Communication Technologies (ICTs); peacebuilding; web; liberal peace; affordances.
^ Does Geotagging come under geography or are the two not necessarily related?

SESSION: Data challenges

Avoiding Chinese Whispers: Controlling End-to-End Join Quality in Linked Open Data Stores
Jan-Christoph Kalo, Silviu Homoceanu, Jewgeni Rose, Wolf-Tilo Balke
Key Topics: Semantic Web, Linked Data, Quality Management, Big Data,
Paper Key Topics: Web Archiving, Open Data, Natural Language Processing,
Author Keywords: Semantic Web, Linked Open Data, Entity Resolution, Instance Matching, Distributed Query Processing, Joins
^Web Graph and networking???

Big Data?: Big Issues Degradation in Longitudinal Data and Implications for Social Sciences
Matthew S. Weber, Hai Nguyen
Key Topics: Big Data, Quality Management, Social Networking, Web Society, Web Analytics,
Paper Key Topics: Social Science, Health, Web Archiving,
Author Keywords: None,

A Linked Data Scalability Challenge: Concept Reuse Leads to Semantic Decay
Paolo Pareti, Ewan Klein, Adam Barker
Key Topics: Linked Data, Big Data, Quality Management, Web Analytics,
Paper Key Topics: Web Archiving,
Author Keywords: None
^ Semantic Information – not the same as Semantic Web

Ranking Buildings and Mining the Web for Popular Architectural Patterns
Ujwal Gadiraju, Stefan Dietze, Ernesto Diaz-Aviles
Key Topics: Data Mining, Digital Sociology, Web Analytics, Web Society, Social Networking,
Paper Key Topics: Crowdsourcing, Sentiment Analysis, Algorithms,
Author Keywords: Web Mining, Crowdsourcing, Architectural Structures,
Influential Factors, Perception.
^ Real life architecture, not Web Architecture!

SESSION: Online Social Behaviour

An Ethnomethodologically-Informed Approach to Interface Design to Support Collective Web Practice Around Video
Anna Zawilska, Steven Albury
Key Topics: Media, Social Networking, Web Society, Web Analytics, Web 2.0, Data Mining, HCI Human Computer Interaction,
Paper Key Topics: User Behaviour, Content Analysis, Online Engagement,
Author Keywords: Ethnomethodology, Social Interaction, Video Annotation

Self Curation, Social Partitioning, Escaping from Prejudice and Harassment: the Many Dimensions of Lying Online
Max Van Kleek, Dave Murray-Rust, Amy Guy, Daniel A. Smith, Kieron O'Hara, Nigel R. Shadbolt
Key Topics: Privacy and Trust, Quality Management, Web Analytics, Web Society, Social Networking, Psychology,

Paper Key Topics: Virtual Community, User Behaviour, Microblogging,
Author Keywords: Lying online; privacy; digital identity; online communities

Anonymity and Online Commenting: The Broken Windows Effect and the End of Drive-by Commenting
Rolf Fredheim, Alfred Moore, John Naughton
Key Topics: Quality Management, Web Analytics, Web Society, Data Mining, Social Networking, Privacy and Trust,
Paper Key Topics: Online Engagement, User Behaviour, Algorithms,
Author Keywords: None

Time to Introduce Myself!: Impact of Self-disclosure Timing of Newcomers in Online Discussion Forums
Di Lu, Rosta Farzan
Key Topics: Web Society, Web Analytics, Social Networking, Psychology, Data Mining,
Paper Key Topics: Virtual Community, Online Engagement, User Behaviour, Content Analysis,
Author Keywords: Self-disclosure, Newcomers, Online communities
^ This and the last one say online community not virtual community – they mean same thing?

Observing Social Machines Part 2: How to Observe?
David De Roure, Clare Hooper, Kevin Page, Ségolène Tarte, Pip Willcox
Key Topics: Internet of Things, Web Society, Web Analytics, Social Networking, Psychology,
Paper Key Topics: Social Machines, Web Archiving, User Behaviour, Content Analysis,
Author Keywords: Social Machines; Web Observatories; Internet of Things.
^ Web Archiving covers Web Observatories

What can be Found on the Web and How: A Characterization of Web Browsing Patterns
Alexey Tikhonov, Liudmila Ostroumova Prokhorenkova, Arseniy Chelnokov, Ivan Bogatyy, Gleb Gusev
Key Topics: Hypertext, Web Modelling, Web Ontology, Web Analysis,
Paper Key Topics: User Behaviour,
Author Keywords: user browsing behaviour, browsing patterns, clustering of web pages

Online Footsteps to Purchase: Exploring Consumer Behaviors on Online Shopping Sites
Munyoung Lee, Taehoon Ha, Jinyoung Han, Jong-Youn Rha, Ted Taekyoung Kwon
Key Topics: Psychology, Web Society, Web Analytics, HCI Human Computer Interaction, Web Search,
Paper Key Topics: E-commerce, User Behaviour, Content Analysis,
Author Keywords: Online Markets, E-commerce, Internet Economy, Consumer Behaviors, Behavior Trajectory, Purchase Prediction

SESSION: Innovating methods

Building a Social Machine: Co-designing a TimeBank for Inclusive Research
Clare J. Hooper, Melanie Nind, Sarah Parsons, Andrew Power, Anne Collis
Key Topics: Social Networking, Research Methodology, Web Analytics, Web Society, Web Technologies, Psychology, HCI Human Computer Interaction, General Web Science, Education,
Paper Key Topics: Social Machine, User Behaviour, Communication Science, Computer Science, Geography, Gamification,
Author Keywords: Social Machine, Inclusive Research; TimeBanking

From Chirps to Whistles: Discovering Event-specific Informative Content from Twitter
Debanjan Mahata, John R. Talburt, Vivek Kumar Singh
Key Topics: Social Networking, Data Mining, Web Analytics,
Paper Key Topics: Microblogging, Algorithms, Content Analysis, Topic Modelling,
Author Keywords: social media mining, text mining, twitter, mutual reinforcement, event, information retrieval, ranking, event-specific information

Assembling thefacebook: Using Heterogeneity to Understand Online Social Network Assembly
Abigail Z. Jacobs, Samuel F. Way, Johan Ugander, Aaron Clauset
Key Topics: Social Networking, Data Mining, Web Analytics, Web Society, Web Graph, Web Architecture, Web Development,
Paper Key Topics: Content Analysis, Network Science, Virtual Community,
Author Keywords: None

Taming a Menagerie of Heavy Tails with Skew Path Analysis
Josh Introne, Sean Goggins
Key Topics: Social Networking, Data Mining, Web Analytics, Web Society, Statistical Analysis,
Paper Key Topics: Content Analysis, Health,
Author Keywords: Power Law, Social Media, Diversity, Dynamics

SESSION: Ethics

Developing the 'Pro-human' Web
Michael J. Day, Leslie Carr, Susan Halford
Key Topics: General Web Science, Web Society, Ethics, Privacy and Trust,
Paper Key Topics: Social Science, User Behaviour,
Author Keywords: Web Science, Pro-human Web, Democracy, Digital Rights
^Relevant for bibliography
Suggests that educating users in the Web and its use and dynamics is vital to creating the Web we want….Points out that programming is being introduced in schools from an early age in order to promote digital literacy, and suggests that the same should be true for Web Education. They state that as with any subject, education is key, and starting at a young age promotes greater awareness.

RoboCode-Ethicists: Privacy-friendly robots, an ethical responsibility of engineers?
Christoph Lutz, Aurelia Tamò
Key Topics: Ethics, Privacy and Trust, Artificial Intelligence, Internet of Things, HCI Human Computer Interaction,
Paper Key Topics: Algorithms,
Author Keywords: Privacy, Robots, Ethical Coding, Applied Ethics

Insights on Privacy and Ethics from the Web's Most Prolific Storytellers
Christopher Wienberg, Andrew S. Gordon
Key Topics: Ethics, Privacy and Trust, Digital Ethnography, Web Analytics, Web Society, Social Networking, Research Methodology,
Paper Key Topics: User Behaviour, Narrative,
Author Keywords: Privacy, Weblogs, Research Ethics, Human Subjects, Research, Ethnography

SESSION: Digital Narratives

Storyscope: Supporting the authoring and reading of museum stories using online data sources
Paul Mulholland, Annika Wolff, Eoin Kilfeather
Key Topics: Web Society, Linked Data, Web Development,
Paper Key Topics: Narrative, Knowledge Patterns, Arts and Culture,
Author Keywords: Museums, digital storytelling, story themes, story settings, events

Archetypal Narratives in Social Machines: Approaching Sociality through Prosopography
Ségolène Tarte, Pip Willcox, Hugh Glaser, David De Roure
Key Topics: Social Networking, Web Analytics, Web Society, Psychology,
Paper Key Topics: Narrative, Social Machines, Virtual Community, User Behaviour,
Author Keywords: Theory, Human Factors, Design
^ Introduction section - Good explanation of social machine

How much is said in a microblog?: A multilingual inquiry based on Weibo and Twitter
Han-Teng Liao, King-Wa Fu, Scott A. Hale
Key Topics: Social Networking, Web Society, Web Analytics, Data Mining, Media, Digital Sociology,
Paper Key Topics: Microblogging, User Behaviour, Content Analysis, Sentiment Analysis,
Author Keywords: Microblogs, Language, Design, Social Networking

'/Command' and Conquer: Analysing Discussion in a Citizen Science Game
Ramine Tinati, Markus Luczak-Roesch, Elena Simperl, Nigel Shadbolt, Wendy Hall
Key Topics: Data Mining, Big Data, Web Analytics, Web Society, Accessibility, Psychology,
Paper Key Topics: Virtual Community, Communication Science, Content Analysis, User Behaviour, Sentiment Analysis, Online Engagement,
Author Keywords: Citizen Science; Gamification; Player Behaviour; Online Communities
^ Should Citizen Science be a topic? – Only in this paper? Also discusses games – but not game theory or gamification?

Analyzing Discourse Communities with Distributional Semantic Models
Igor Brigadir, Derek Greene, Pádraig Cunningham
Key Topics: Web Analytics, Politics, Web Society, Social Networking,
Paper Key Topics: Content Analysis, Microblogging, Sentiment Analysis, Social Science, User Behaviour,
Author Keywords: None
^ Another one looking at textual analysis – should this be a theme or too general?

How much is Wikipedia Lagging Behind News?
Besnik Fetahu, Abhijit Anand, Avishek Anand
Key Topics: Media, Web Society, Web Analytics, Statistical Analysis,
Paper Key Topics: Social Machines, Content Analysis,
Author Keywords: entity lag, event lag, news reference density, emergent entity, density, wikipedia, news corpora

Considering a Wider Web?: Employing Multimodal Critical Discourse Analysis in Exploration of Multiple Online Spaces
Rebecca Nash
Key Topics: Big Data, Linked Data, Research Methodology, Data Mining, Web Analytics, Ethics, Web Society,
Paper Key Topics: Health,
Author Keywords: Aesthetic Surgery, Big Data, Discourse, Multimodality, Qualitative Research, Web
^ Interesting to my work?! (Is this talking about textual analysis too? Is discourse analysis the same thing?)

SESSION: Social Safety & Wellbeing

Habits vs Environment: What Really Causes Asthma?
Mengfan Tang, Pranav Agrawal, Ramesh Jain
Key Topics: Risk Management, Web Analytics, Web Society, Data Mining,
Paper Key Topics: Health, Social Science,
Author Keywords: Asthma, Feature extraction, Asthma risk analysis, Gradient Boosting Tree
^ Collecting sensor data, is this the same as data mining?

Emotional States vs. Emotional Words in Social Media
Asaf Beasley, Winter Mason
Key Topics: Social Networking, Web Analytics, Web Society, Statistical Analysis, Psychology,
Paper Key Topics: Content Analysis, Sentiment Analysis, Microblogging, User Behaviour,
Author Keywords: emotions, sentiment analysis, social media

Assessing the Value of Social Media for Organisations: The Case for Charitable Use
Christopher Phethean, Thanassis Tiropanis, Lisa Harris
Key Topics: General Web Science, Social Networking, Web Society, Marketing, Media,
Paper Key Topics: Communication Science, Online Advertising,
Author Keywords: Social media, charities, marketing, communication, web science

Web Science 2016 Conference

SESSION: Behaviour

Beyond the MOOC platform: gaining insights about learners from the social web
Guanliang Chen, Dan Davis, Jun Lin, Claudia Hauff, Geert-Jan Houben
Key Topics: e-learning, Web Analytics, Web Society, Social Networking, Data Mining, Big Data, Research Methodology,
Paper Key Topics: MOOC, Content Analysis, User Behaviour, Online Engagement, Gamification,
Author Keywords: Human-centred computing, User models, Social networks, Applied computing, Learning management systems,

Understanding video-ad consumption on YouTube: a measurement study on user behavior, popularity, and content properties
Mariana Arantes, Flavio Figueiredo, Jussara M. Almeida
Key Topics: Media, Social Networking, Psychology, Web Society, Digital Sociology, Web Analytics, Marketing, Data Mining, Big Data,
Paper Key Topics: User behaviour, Content Analysis, Algorithms,
Author Keywords: YouTube; Video Ads; Popularity; User-Behavior

Teens are from mars, adults are from venus: analyzing and predicting age groups with behavioral characteristics in instagram
Kyungsik Han, Sanghack Lee, Jin Yea Jang, Yong Jung, Dongwon Lee
Key Topics: Social Networking, Psychology, Big Data, Web Analytics, Sociology, Digital Sociology, Data Mining, Web Society,
Paper Key Topics: User Behaviour, Social Science, Content Analysis, Online Identity, Demographics
Author Keywords: Teens in social media; behavioral patterns and detection of teens in social media; comparative analysis

Because science is awesome: studying participation in a citizen science game
Ramine Tinati, Markus Luczak-Roesch, Elena Simperl, Wendy Hall
Key Topics: Crowdsourcing, Web Society, Sociology, Digital Sociology, Psychology, Web Analytics,
Paper Key Topics: User Behaviour, Online Engagement, Social Science, Content Analysis, Gamification, Communication Science,
Author Keywords: Citizen Science, Crowdsourcing, Online Communities,

SESSION: Relationships and identity

Identity assurance in the UK: technical implementation and legal implications under the eIDAS regulation
Niko Tsakalakis, Kieron O'Hara, Sophie Stalla-Bourdillon
Key Topics: Privacy and Trust, Web Applications, Ethics, Web Society,
Paper Key Topics: Government, Online Identity,
Author Keywords: eID, eIDM, electronic identity, trust services, Gov.UK Verify, German nPA, eIDAS

Privacy and twitter in qatar: traditional values in the digital world
Norah Abokhodair, Sofiane Abbar, Sarah Vieweg, Yelena Mejova
Key Topics: Privacy and Trust, Research Methodology, Web Society, Social Networking, Sociology, Digital Sociology, Psychology,
Paper Key Topics: Online Identity, Social Science, User Behaviour, Religion,
Author Keywords: Privacy; Islam; Qatar; Twitter; Mixed Methods; Arab world; Gender; Honor.

The social ties of immigrant communities in the United States
Amaç Herdağdelen, Bogdan State, Lada Adamic, Winter Mason
Key Topics: Web Society, Sociology, Digital Sociology, Research Methodology, Social Networking, Psychology, Web Analytics, Data Mining,
Paper Key Topics: Demographics, Social Science, User Behaviour, Online Identify, Content Analysis, Online Offline Community, Network Science,
Author Keywords: migration, social networks, integration

SESSION: Information spreading and engagement

Talking climate change via social media: communication, engagement and behaviour
Miriam Fernandez, Lara S. G. Piccolo, Diana Maynard, Meia Wippoo, Christoph Meili, Harith Alani
#ISISisNotIslam or #DeportAllMuslims?: predicting unspoken views
Walid Magdy, Kareem Darwish, Norah Abokhodair, Afshin Rahimi, Timothy Baldwin
Key Topics: Social Networking, Psychology, Web Analysis, Data Mining, Web Society,
Paper Key Topics: Microblogging, Network Science, Content Analysis, User Behaviour, Crowdsourcing,
Author Keywords: Network analysis, Twitter data analysis, Stance prediction, Paris attacks, Homophily, Social influence,

Spreading the news: how can journalists gain more engagement for their tweets?
Claudia Orellana-Rodriguez, Derek Greene, Mark T. Keane
Key Topics: Media, Social Networking, Digital Sociology, Web Analysis, Web Society, Psychology,
Paper Key Topics: Online engagement, Microblogging, Journalism, Content Analysis, User Behaviour,
Author Keywords: Computational journalism; social media; audience engagement; news events

Economic value of web keyword advertising campaigns
Sergej Sizov, Sarah Piller
Key Topics: Economics, Marketing,
Paper Key Topics: User Behaviour, Content Analysis, Online Advertising,
Author Keywords: keyword advertising, keyword auctions, revenue analysis, AdWords

SESSION: Conceptualization

ACROSS: A framework for multi-cultural interlinking of web taxonomies

Natalia Boldyrev, Marc Spaniol, Gerhard Weikum

Key Topics: Big Data, Linked Data, Web Analytics, Web Architecture, Semantic Web, Data Mining,

Paper Key Topics: Network Science,

Author Keywords: Multicultural Knowledge Bases, Knowledge Taxonomies, Alignment Methods, Integer Linear Programming, Integer Quadratic Programming, Label Propagation

Towards a theoretical approach for analysing music recommender systems as sociotechnical cultural intermediaries

Jack Webster, Nicholas Gibbins, Susan Halford, Brian J. Hracs

Key Topics: Big Data, Linked Data, Web Analysis, Media, Web Society,

Paper Key Topics: Content Analysis, Computer Science, Social Science,

Author Keywords: Big Data; Recommender Systems; Sociotechnical Systems; Cultural Intermediaries; Bourdieu; Actor-Network Theory

^ NB. Could include 'recommender system' as a paper keyword, but is it too niche? It's a specific topic as opposed to a subject area?

Reconceptualizing imitation in social tagging: a reflective search model of human web interaction

Paul Seitlinger, Tobias Ley

Key Topics: HCI Human Computer Interaction, Web Society, Big Data, Social Networking, Psychology,

Paper Key Topics: User Behaviour,

Author Keywords: Social Tagging, Semantic Stabilization, Organism-Environment Dynamics, Search of human memory

^ Tagging, not geotagging?

Social machines in practice: solutions, stakeholders and scopes

Clare J. Hooper, Brian Bailey, Hugh Glaser, James Hendler

Key Topics: Linked Data, Social Networking, Semantic Web, Internet of Things, Privacy and Trust, Web Society,

Paper Key Topics: Social Machines, Crowdsourcing, User Behaviour,

Author Keywords: Social machines; stakeholders; linked data.

SESSION: Digital rights and public access

A content-linking-context model for "notice-and-take-down" procedures

Pei Zhang, Sophie Stalla-Bourdillon, Lester Gilbert

Key Topics: Social Networking, Media, Web Analysis, Web Applications, Web Architecture,

Paper Key Topics: User behaviour, Topic Modelling,

Author Keywords: Social media comments; Reddit; News

^ Topic Modelling as a paper keyword? Yes

A manifesto for data sharing in social media research
Katrin Weller, Katharina E. Kinder-Kurlanda
Key Topics: Linked Data, Big Data, Privacy and Trust, Web Analytics, Data Mining, Accessibility, Social Networking,
Paper Key Topics: Open Data, Open Access, Online Engagement, Web Publishing,
Author Keywords: Reproducibility; methodology; social media; archiving; data sharing; data archives; privacy; data protection; legal issues
Can we find documents in web archives without knowing their contents?
Khoi Duy Vo, Tuan Tran, Tu Ngoc Nguyen, Xiaofei Zhu, Wolfgang Nejdl
Key Topics: Web Search, Research Methodology, Web Analytics, Web Graph,
Paper Key Topics: Web Archiving,
Author Keywords: Web Archive Search, Temporal Ranking, Feature Analysis

Twitter as a first draft of the present: and the challenges of preserving it for the future
Axel Bruns, Katrin Weller
Key Topics: Web History, Social Networking,
Paper Key Topics: Microblogging, Social Science, Arts and Culture
Author Keywords: Twitter; social media; user-generated content; cultural heritage; archiving; history; historical sources.

SESSION: Politics and the web

Using the web to support political analysis: identifying legislative bill ideology in the chilean parliament
Pablo Loyola, Francisco Szederkenyi, Yutaka Matsuo
Key Topics: Politics, Web Society, Web Analytics, Data Mining, Digital Sociology,
Paper Key Topics: Topic Modelling, User Behaviour, Social Science,
Author Keywords: Political Analysis, Word Embeddings

Quote RTs on Twitter: usage of the new feature for political discourse
Kiran Garimella, Ingmar Weber, Munmun De Choudhury
Key Topics: Politics, Web Analytics, Web Society, Social Networking, Digital Sociology,
Paper Key Topics: Microblogging, Social Science,
Author Keywords: Networks, Online Social Networks

Internet use, in- and exclusion in decision-making processes within political parties
Gefion Thuermer, Silke Roth, Markus Luczak-Rösch, Kieron O'Hara
Key Topics: Politics, Web Technology, Internet Protocols, Social Networking, Web Society, Digital Sociology,
Paper Key Topics: Democracy, Online Engagement, Social Science,
Author Keywords: Decision-making; e-democracy; political parties; process design; inclusion,

Virtual caucasus on VK social networking site
Daniel Alexandrov, Alexey Gorgadze, Ilya Musabirov

Key Topics: Social Networking, Digital Sociology, Sociology, Politics, Web Society, Data Mining, Web Analytics,
Paper Key Topics: Topic Modelling, Social Science, Vital Community, Network Science, Content Analysis, User Behaviour, Geography,
Author Keywords: Ethnic groups, LDA, Topic modeling, SNA, On-line social networks, VKontakte

SESSION: Categorization and predictions

LlamaFur: learning latent category matrix to find unexpected relations in Wikipedia
Paolo Boldi, Corrado Monti
Key Topics: Data Mining, Linked Data, Hypertext, Semantic Web, Web Architecture, Web Analytics,
Paper Key Topics: Topic Modelling, Information Theory, Content Analysis,
Author Keywords: None

The psychology of job loss: using social media data to characterize and predict unemployment
Davide Proserpio, Scott Counts, Apurv Jain
Key Topics: Social Networking, Economics, Psychology, Web Analytics, Web Society,
Paper Key Topics: Social Science, User Behaviour, Content Analysis
Author Keywords: Social science, social media, economics, unemployment

Content attribution ignoring content
Mattia Samory, Enoch Peserico
Key Topics: Privacy and Trust, Web Analytics, Data Mining, Social Networking, Ethics
Paper Key Topics: Content Analysis, User Behaviour, Natural Language Processing,
Author Keywords: authorship, attribution, structural features, forum, identification, timing, social, privacy
^Authorship?

Towards detection of influential sentences affecting reputation in wikipedia
Yiwei Zhou, Alexandra I. Cristea
Key Topics: Data Mining, Web Analytics, Quality Management,
Paper Key Topics: Content Analysis,
Author Keywords: Wikipedia; Cross-domain classification; Reputation-influentia
^ Sentiment Analysis?

SESSION: Information spreading

Seeing the forest for the trees: new approaches to forecasting cascades
Siddharth Krishnan, Patrick Butler, Ravi Tandon, Jure Leskovec, Naren Ramakrishnan
Key Topics: Social Networking, Web Graph, Semantic Web, Web Analytics,
Paper Key Topics: Microblogging, Network Science, User Behaviour,
Author Keywords: None
^ Information Propagation

Information dissemination in heterogeneous-intent networks
Abhimanyu Das, Sreenivas Gollapudi, Emre Kıcıman, Onur Varol
Key Topics: Social Networking, Psychology, Web Society, Web Analytics,
Paper Key Topics: Topic Modelling, User Behaviour, Microblogging, Online Engagement, Communication Science, Social Science,
Author Keywords: information dissemination, user modelling, topic modelling, social media

Using social network analysis to predict online contributions: the impact of network diversity in cross-cultural collaboration
Jenna Mittelmeier, YingFei Héliot, Bart Rienties, Denise Whitelock
Key Topics: Social Networking, Web Analytics, Web Society, Psychology, Education, Research Methodology,
Paper Key Topics: Online Offline Community, User Behaviour, Online Engagement, Crowdsourcing, Social Science, Content Analysis, Communication Science,
Author Keywords: Social Network Analysis, online collaboration, social networks, group work, cross-cultural collaboration, online contributions

Anticipated shocks in online activity: response functions of attention and word-of-mouth processes
Sebastian Stommel, David Garcia, Adiya Abisheva, Frank Schweitzer
Key Topics: Web Society, Web Analytics, Media, Psychology,
Paper Key Topics: Social Science, User Behaviour, Microblogging, Content Analysis,
Author Keywords: Social systems; collective response; attention; word-of-mouth;

SESSION: Information gathering

Do it yourself diagnosis: a study on acquiring health-related information online
Duwaraka Murugadas, Sergej Sizov
Key Topics: Web Society, Web Search, Psychology, Digital Sociology,
Paper Key Topics: Health, Digital Literacy, User Behaviour,
Author Keywords: online disease diagnosis, health-related information, search strategy, information literacy

Finding diverse needles in a haystack of comments: social media exploration for news
Hang Zhang, Vinay Setty
Key Topics: Web Analytics, Social Networking, Data Mining,
Paper Key Topics: Content Analysis, Topic Modelling, Narrative,
Author Keywords: Social media comments; Reddit; News

Analyzing web archives through topic and event focused sub-collections
Gerhard Gossen, Elena Demidova, Thomas Risse
Key Topics: Web History, Web Analytics, Data Mining, Hypertext, Linked Data,
Paper Key Topics: Content Analysis, Topic Modelling,
Author Keywords: Web archive; sub-collection; topics; events

Using social media to reveal social and collective perspectives on music
Deniz Iren, Cynthia C. S. Liem, Jie Yang, Alessandro Bozzon
Key Topics: Social Networking, Web Society, Web Analytics, Media, Psychology, Data Mining,
Paper Key Topics: Microblogging, Content Analysis, User Behaviour,
Author Keywords: social media, data analysis, music preferences

Private and public online groups in apartment buildings of St. Petersburg
Vadim Voskresenskiy, Ilya Musabirov, Daniel Alexandrov
Key Topics: Social Networking, Web Society, Psychology, Privacy and Trust, Web Analytics,
Paper Key Topics: User Behaviour, Online Engagement, Content Analysis,
Author Keywords: computer-mediated communication; local social media; citi- zen engagement; urban communities; place-based communities; community-oriented social media,
^ Coded as far as here in NVIVO!

Web Science 2017 Conference

http://dl.acm.org/citation.cfm?id=3091478&picked=prox

SESSION: Long Session I: Aggression, Controversy, Crime

Beyond Cyberbullying: Self-Disclosure, Harm and Social Support on ASKfm
Zahra Ashktorab, Eben Haber, Jennifer Golbeck, Jessica Vitak
Key Topics: Social Networking, Privacy and Trust, Web Society, Web Analytics, Psychology, Digital Sociology,
Paper Key Topics: User Behaviour, Online Engagement, Cyber Bullying, Topic Modelling,
Author Keywords: ASKfm; cyberbullying; self-disclosure; topic modeling
Permission

The Effect of Collective Attention on Controversial Debates on Social Media
Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, Michael Mathioudakis
Key Topics: Social Networking, Web Society, Web Analytics, Digital Sociology,
Paper Key Topics: Microblogging, User Behaviour, Content Analysis, Network Science,
Author Keywords: None

Mean Birds: Detecting Aggression and Bullying on Twitter
Despoina Chatzakou, Nicolas Kourtellis, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, Athena Vakali
Key Topics: Social Networking, Web Society, Web Analytics, Psychology, Digital Sociology,
Paper Key Topics: Cyber Bullying, User Behaviour, Content Analysis, Microblogging, Network Science,
Author Keywords: None

Stateless Puzzles for Real Time Online Fraud Preemption
Mizanur Rahman, Ruben Recabarren, Bogdan Carbunar, Dongwon Lee
Key Topics: Social Networking, Web Applications, Web Graph, Web Technologies,
Paper Key Topics: Cybercrime**,**
Author Keywords: Stateless Puzzle, Online Fraud Preemption

The Signals and Noise: Actionable Information in Improvised Social Media Channels During a Disaster
Xingsheng He, Di Lu, Drew Margolin, Mengdi Wang, Salma El Idrissi, Yu-Ru Lin
Key Topics: Social Networking, Web Society, Web Analytics, Data Mining,
Paper Key Topics: Online Engagement, Crowdsourcing, Communication Science. Quality Control,
Author Keywords: collective intelligence, disaster response, social media, improvised logistical channel, crowd behaviors, self-organized systems
^ Disaster response? – not specifically Web Science?

SESSION: Long Session II: Groups & Individuals, Attention, Opinion

Predicting Trust Relations Within a Social Network: A Case Study on Emergency Response
Nikhita Vedula, Srinivasan Parthasarathy, Valerie L. Shalin
Key Topics: Web Society, Privacy and Trust, Psychology, Sociology, Digital Sociology, Social Networking,
Paper Key Topics: Communication Science, User Behaviour, Network Science, Sentiment Analysis, Microblogging, Disaster response,
Author Keywords: None

Privacy for Children and Teenagers on Social Networks from a Usability Perspective: A Case Study on Facebook
Cristiana S. Silva, Glívia A.R. Barbosa, Ismael S. Silva, Tatiane S. Silva, Fernando Mourão, Flávio Coutinho
Key Topics: Privacy and Trust, Social Networking, Security, Web Society, Web Analytics, Psychology, Human Computer Interaction, Digital Sociology,
Paper Key Topics: User Behaviour, Social Science,
Author Keywords: Privacy; Usability; Online Social Networks; Children; Teenagers.

Closed-Loop Opinion Formation
Larissa Spinelli, Mark Crovella
Key Topics: Psychology, Web Society, Web Analytics, Marketing, Web Modelling,
Paper Key Topics: Content Analysis, Online Engagement, Recommender Systems, User Behaviour
Author Keywords: Filter Bubble, User Experience, Metrics, User Behaviour, Recommender Systems
User Experience? – too similar?

"(Weitergeleitet von Journalistin)": The Gendered Presentation of Professions on Wikipedia
Olga Zagovora, Fabian Flöck, Claudia Wagner
Key Topics: Web Society, Web Search, Sociology, Digital Sociology, Psychology, Web Analytics, Data Mining,
Paper Key Topics: Demographics**,** Social Science, Content Analysis,
Author Keywords: Wikipedia; gender inequality; professions; gender bias

Where Could We Go?: Recommendations for Groups in Location-Based Social Networks
Frederick Ayala-Gómez, Bálint Daróczy, Michael Mathioudakis, András Benczúr, Aristides Gionis
Key Topics: Social Networking, Web Society, Data Mining, Web Analytics,
Paper Key Topics: Recommender Systems, User Behaviour, Online Engagement, Geography, Content Analysis, Geo-tagging,
Author Keywords: Group Recommendation; Location-Based Social Networks; Recommender Systems

SESSION: Long Session III: Talking, Thinking, and Living Online

Using Twitter Data to Estimate the Relationship between Short-term Mobility and Long-term Migration
Lee Fiorio, Guy Abel, Jixuan Cai, Emilio Zagheni, Ingmar Weber, Guillermo Vinué
Key Topics: Web Society, Web Analytics, Web Modelling, Statistical Analysis, Sociology, Digital Sociology, Social Networking
Paper Key Topics: Demographics, Geography, Microblogging, Social Science,
Author Keywords: Twitter, Migration, Mobility, Demographic research

ELEVATE: A Framework for Entity-level Event Diffusion Prediction into Foreign Language Communities
Govind, Marc Spaniol
Key Topics: Web Society, Accessibility, Social Networking, Media, Web Analytics, Digital Sociology, Digital Ethnography,
Paper Key Topics: Virtual Communities, Content Analysis, Geography, Social Science,
Author Keywords: Societal Events Analysis, Multilingual Web Data

Predicting Rising Follower Counts on Twitter Using Profile Information
Juergen Mueller, Gerd Stumme
Key Topics: Psychology, Web Society, Web Analytics, Social Networking, Digital Sociology,
Paper Key Topics: Microblogging, User Behaviour, Social Science, Online Identity, Content Analysis,
Author Keywords: Prediction; Correlational Analysis; Classification; Social Network Analysis; Experimental Study; Onomastics

Sharing Means Renting?: An Entire-marketplace Analysis of Airbnb
Qing Ke
Key Topics: Marketing, Economics, Web Analytics, Data Mining,
Paper Key Topics: Content Analysis, Recommender Systems, Geography,
Author Keywords: Airbnb; measurement; sharing economy; online marketplace

Understanding Citizens' and Local Governments' Digital Communications During Natural Disasters: The Case of Snowstorms
Lingzi Hong, Cheng Fu, Paul Torrens, Vanessa Frias-Martinez
Key Topics: Web Society, Web Analytics, Data Mining, Linked Data, ,
Paper Key Topics: Disaster Response, Government, Microblogging, Communication Science, Topic Modelling, Narrative
Author Keywords: crisis communication, disaster analytics, topic models, spatio-temporal analysis

SESSION: Long Session IV: Networks -- Structure, Identifiers, Search

Breaking Cycles In Noisy Hierarchies
Jiankai Sun, Deepak Ajwani, Patrick K. Nicholson, Alessandra Sala, Srinivasan Parthasarathy
Key Topics: Web Graph, Web Applications, Web Analytics, Web Ontologies,
Paper Key Topics: Network Science,
Author Keywords: Directed Acyclic Graph, Graph Hierarchy, TrueSkill, Social Agony, Cycle Edges

ε - WGX: Adaptive Edge Probing for Enhancing Incomplete Networks
Sucheta Soundarajan, Tina Eliassi-Rad, Brian Gallagher, Ali Pinar
Key Topics: Web Graph, Web Analytics, Data Mining, Web Ontologies, Web Architecture,
Paper Key Topics: Network Science, (Narrative?)
Author Keywords: incomplete networks, adaptive probing, graph exploration

Hierarchical Change Point Detection on Dynamic Networks
Yu Wang, Aniket Chakrabarti, David Sivakoff, Srinivasan Parthasarathy
Key Topics: Social Networking, Web Analytics, Web Graph, Web Architecture, Web Modelling,
Paper Key Topics: Network Science, Virtual Communities, Algorithms,
Author Keywords: Anomaly Detection, Dynamic Social Networks, Community Detection

Analyzing the Keystroke Dynamics of Web Identifiers
Andrew G. West
Key Topics: Web Analytics, HCI Human Computer Interaction, Social Networking, Web 2.0, Web Society,
Paper Key Topics: Content Analysis, Crowdsourcing, Microblogging,
Author Keywords: typeability; keystroke dynamics; keyboards; typos; web identifier; domain names; hashtags; usernames
^ More on this one?

What's Happening and What Happened: Searching the Social Web
Omar Alonso, Vasileios Kandylas, Serge-Eric Tremblay, Jake M. Hofman, Siddhartha Sen
Key Topics: Web Society, Social Networking, Linked Data, Data Mining, Web Analytics, Web Search,
Paper Key Topics: User Behaviour, Content Analysis, Microblogging,
Author Keywords: None

SESSION: Short Session I

ASSIST: Automatic Summarization of Significant Structural Changes in Large Temporal Graphs
Charalampos Chelmis, Reshul Dani
Key Topics: Web Graph, Big Data, Linked Data, Web Analytics, Web Modelling, Web Architecture,
Paper Key Topics: Network Science, Content Analysis,
Author Keywords: Anomaly summarization; structural change interpretation; dynamic graph; change attribution

Automatic Generation of Event Timelines from Social Data
Omar Alonso, Serge-Eric Tremblay, Fernando Diaz
Key Topics: Social Networking, Web Society, Web Analytics, Artificial Intelligence, Web Modelling,
Paper Key Topics: Content Analysis, Narrative, Microblogging,
Author Keywords: Timelines, social pseudo relevance feedback, social media, Twitter

Cultural Fault Lines and Political Polarization
Yongren Shi, Kai Mast, Ingmar Weber, Agrippa Kellum, Michael Macy
Key Topics: Social Networking, Web Society, Politics, Digital Ethnography, Web Analytics, Digital Sociology,
Paper Key Topics: Demographics, Microblogging, Content Analysis, User Behaviour, Geography,
Author Keywords: Polarization, Culture, Social Media, Networks

EDSV: Emerging Defect Surveillance for Vehicles
Jiejun Xu, Daniel Xie, Tsai-Ching Lu, John Cafeo
Key Topics: Social Networking, Quality Management, Web 2.0, Data Mining, Web Analytics, Statistical Analysis,
Paper Key Topics: Content Analysis, Real-time Data, Topic Modelling,
Author Keywords: Measurement; Online Social Media; User Generated Content; Business Intelligence; Quality Management

Studying Personality through the Content of Posted and Liked Images on Twitter
Sharath Chandra Guntuku, Weisi Lin, Jordan Carpenter, Wee Keong Ng, Lyle H. Ungar, Daniel Preoţiuc-Pietro
Key Topics: Social Networking, Media, Web Society, Psychology, Web Modelling, Marketing, Data Mining, Web Analytics,
Paper Key Topics: Content Analysis, Microblogging, User Behaviour, Computer Science,
Author Keywords: None

A Large Labeled Corpus for Online Harassment Research
Jennifer Golbeck, Zahra Ashktorab, Rashad O. Banjo, Alexandra Berlinger, Siddharth Bhagwan, Cody Buntain, Paul Cheakalos, Alicia A. Geller, Quint Gergory, Rajesh Kumar Gnanasekaran, Raja Rajan Gunasekaran, Kelly M. Hoffman, Jenny Hottle, Vichita Jienjitlert, Shivika Khare, Ryan Lau, Marianna J. Martindale, Shalmali Naik, Heather L. Nixon, Piyush Ramachandran, Kristine M. Rogers, Lisa Rogers, Meghna Sardana Sarin, Gaurav Shahane, Jayanee Thanki, Priyanka Vengataraman, Zijian Wan, Derek Michael Wu
Key Topics: Big Data, Linked Data, Data Mining, Artificial Intelligence, Web Analytics, Digital Sociology,
Paper Key Topics: Microblogging, Cyberbullying, Cybercrime, User Behaviour, Social Science,
Author Keywords: online harassment; datasets

The Fake News Spreading Plague: Was it Preventable?
Eni Mustafaraj, Panagiotis Takis Metaxas
Key Topics: Politics, Social Networking, Web Society, Media,
Paper Key Topics: Microblogging, Cybercrime, Virtual Community, User Behaviour, Algorithms,
Author Keywords: fake news; misinformation spreading; Facebook; Twitter; Google

SESSION: Short Session II

Web Science: Mapping the Curriculum
Authors: Elisabeth Coskun, Su White
Key Topics: General Web Science, Education, Web Society,

Paper Key Topics: Content Analysis,
Author Keywords: Web Science, Education, Web Science Curriculum

Young People's Policy Recommendations on Algorithm Fairness
Authors: Elvira Perez Vallejos, Ansgar Koene, Virginia Portillo, Liz Dowthwaite, Monica Cano
Key Topics: Web Society, Privacy and Trust, Digital Sociology, Education,
Paper Key Topics: Digital Literacy, Social Science, Algorithms,
Author Keywords: Youth jury; algorithm fairness; youth opinion; deliberation; digital literacy; digital citizenship, privacy; policy

Using Facebook Ads Audiences for Global Lifestyle Disease Surveillance: Promises and Limitations
Authors: Matheus Araujo, Yelena Mejova, Ingmar Weber, Fabricio Benevenuto
Key Topics: Social Networking, Marketing, Web Society, Web Analytics, Digital Sociology, Statistical Analysis,
Paper Key Topics: User Behaviour, Health, Geography, Demographics, Online Advertising,
Author Keywords: Facebook, Advertising, Epidemiology, Social Media, Health

Ad-blocking: A Study on Performance, Privacy and Counter-measures
Authors: Kiran Garimella, Orestis Kostakis, Michael Mathioudakis
Key Topics: Privacy and Trust, Web Analytics,
Paper Key Topics: Online Advertising,
Author Keywords: None

Factors in Recommending Contrarian Content on Social Media
Authors: Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, Michael Mathioudakis
Key Topics: Social Networking, Web Analytics, Media, Digital Sociology,
Paper Key Topics: Democracy, Recommender Systems, Social Science, Microblogging, Online Advertising, Algorithms,
Author Keywords: None

InstaCan: Examining Deleted Content on Instagram
Authors: Ramine Tinati, Aastha Madaan, Wendy Hall
Key Topics: Social Networking, Big Data, Media, Privacy and Trust, Data Mining, Web Analytics,
Paper Key Topics: Microblogging, Web Publishing, Web Archiving,
Author Keywords: Micro-blogging, Social media, Twitter, Instagram, Photo-sharing, Deletion

Web Science Challenges in Researching Bug Bounties
Authors: Huw Fryer, Elena Simperl
Key Topics: Security, General Web Science, Web Development,
Paper Key Topics: Crowdsourcing, Cybercrime,
Author Keywords: Crowdsourcing; Vulnerability research; Bug bounties

SESSION: Long Session V: Time, Space, Archives

Characterizing Regional and Behavioural Device Variations Across the Twitter Timeline: A Longitudinal Study
Authors: Laura Cruz-Albrecht, Jiejun Xu, Kang-Yu Ni, Tsai-Ching Lu
Key Topics: Hardware, Social Networking, Web Society, Web Analytics, Data Mining,
Paper Key Topics: Microblogging, Geography, Demographics, User Behaviour,
Author Keywords: Twitter; Mobile Device; Longitudinal Study; Behavioral Variation; Device Homophily

Exploring Web Archives Through Temporal Anchor Texts
Authors: Helge Holzmann, Wolfgang Nejdl, Avishek Anand
Key Topics: Big Data, Web Search, Web Analytics, Data Mining, Web Graph, Web Modelling,
Paper Key Topics: Web Archiving,
Author Keywords: Web Archives; Temporal Information Retrieval; Big Data Analysis
^ Main topic of paper

Observing Web Archives: The Case for an Ethnographic Study of Web Archiving
Authors: Jessica Ogden, Susan Halford, Leslie Carr
Key Topics: Digital Ethnography, Big Data,
Paper Key Topics: Web Archiving, Algorithms,
Author Keywords: web archiving, knowledge production, STS, materiality, information labour
Web Archiving ^ Main topic of paper

Generating Stories From Archived Collections
Authors: Yasmin AlNoamany, Michele C. Weigle, Michael L. Nelson
Key Topics: Web Analytics, Data Mining, Big Data, Social Networking,
Paper Key Topics: Web Archiving, Narrative,
Author Keywords: Web Archiving, Storytelling, Information Retrieval, Document Similarity, Archived Collections, Web Content mining, Internet Archive

Deconstructing Diffusion on Tumblr: Structural and Temporal Aspects
Authors: Nora Alrajebah, Leslie Carr, Markus Luczak-Roesch, Thanassis Tiropanis
Key Topics: Social Networking, Web Society, Web Analytics, Data Mining,
Paper Key Topics: Web Publishing, Microblogging, Network Science, User Behaviour,
Author Keywords: Social Network Analysis, Information Diffusion, Cascades, Tumblr

SESSION: Long Session VI: Reflecting, Thinking, Understanding

A Deep Study into the History of Web Design
Authors: Bardia Doosti, David J. Crandall, Norman Makoto Su
Key Topics: Web Design, Web Analytics, Web History, Computer Graphics, Web Development,
Paper Key Topics: Content Analysis, Arts and Culture,
Author Keywords: Web Design, Deep Learning, Convolutional Neural Networks, Cultural Analytics

The Ethical Challenges of Publishing Twitter Data for Research Dissemination
Authors: Helena Webb, Marina Jirotka, Bernd Carsten Stahl, William Housley, Adam Edwards, Matthew Williams, Rob Procter, Omer Rana, Pete Burnap
Key Topics: Ethics, Web Analytics, Big Data, Data Mining,
Paper Key Topics: Microblogging, Content Analysis, Web Publishing, Research Methodology,
Author Keywords: Research ethics, Twitter, social media, informed consent

Broad, Interdisciplinary Science In Tela: An Exposure and Child Health Ontology

Authors: James P. McCusker, Sabbir M. Rashid, Zhicheng Liang, Yue Liu, Katherine Chastain, Paulo Pinheiro, Jeanette A. Stingone, Deborah L. McGuinness
Key Topics: Linked Data, Big Data, Web Analytics, Web Modelling, Web Ontology,
Paper Key Topics: Health,
Author Keywords: None

A Turn for the Scruffy: An Ethnographic Study of Semantic Web Architecture
Authors: Lindsay Poirier
Key Topics: Semantic Web, Digital Ethnography, Web Architecture, Web Ontology, Internet Protocols, Web History,
Paper Key Topics:
Author Keywords: Semantic Web, Logic, Science and Technology Studies, Thought Styles, Design Logics

Understanding Temporal Backing Patterns in Online Crowdfunding Communities
Authors: Yiming Liao, Thanh Tran, Dongwon Lee, Kyumin Lee
Key Topics: Web Analytics, Data Mining, Big Data, Economics,
Paper Key Topics: User Behaviour,
Author Keywords: Crowdfunding; User Behavior Analysis; Temporal Pattern