Evolving a Web Science curriculum

Elisabeth Coskun
Web and Internet Science (WAIS)
University of Southampton
eac1g09@ecs.soton.ac.uk

Su White
Web and Internet Science (WAIS)
University of Southampton
saw@ecs.soton.ac.uk

ABSTRACT

Web Science is an emerging subject, which was formally defined in 2006. Since the inception of Web Science, a number of Web Science taught programs have emerged across the world. Despite this, there are still currently very few attempts to provide a formal curriculum definition for Web Science. This paper outlines a study which explores the scope of Web Science, by examining what is taught as part of Web Science MSc and Undergraduate programs worldwide. This paper presents some initial results of the study, including the most frequently occurring 'key topics' found from an analysis of Web Science taught modules. Plans for future work are then outlined; the ultimate aim of the study being to present a view of Web Science as it is taught across the world.

CCS Concepts

• **Social and professional topics~Computing education** • **Social and professional topics~Computing education programs** • **Social and professional topics~Information science education**

Keywords

Web Science; Education; Web Science Curriculum.

1. INTRODUCTION

Web Science was formally outlined as a subject by Berners-Lee et al in 2006. [1] [2] They make the case for a dedicated subject to study and understand the development of the Web, in order to influence and shape its development in ways which benefit society. Since 2006 a number of Web Science taught programmes have emerged across the world. Despite this, there are currently no formal curriculum guidelines available for Web Science, which is an issue if Web Science is to continue to gain popularity as a taught subject.

2. A 'Bottom up' approach

As explained by White et al. [3], traditional approaches to defining curricula such as the example of the ACM Joint Task Force for Computer Science [4] are typically undertaken utilizing a 'top down' or theoretical research approach, in which a panel of experts will agree a set of criteria, which are often based upon revisions of pre-existing curriculum definitions, and then seek evidence to validate the theory. The final agreed curriculum definition is then given to educators for their feedback, before it then becomes the formal model and set of standards upon which all other taught curricula are based. This approach is both onerous, time consuming and costly. In the case of a subject with evolution as rapid as that of Web Science, any definition created

utilizing this method will potentially become outdated by the time it is completed. There are currently no formal curriculum guidelines available for Web Science, yet there are a number of Web Science taught curricula in existence. This study conducts a bottom up examination of these current Web Science related curriculums and resources with the future aim of building a picture of Web Science according to what is taught as Web Science worldwide, in a 'bottom up', reverse engineering fashion. [5] Instead of the curriculum guidelines shaping the curricula, the curricula will be used to shape and develop a set of Web Science curriculum guidelines, which will ultimately be presented in the form of a taxonomy of key topics which are typically taught as part of Web Science programmes.

The proposed bottom up approach is suited to qualitative Grounded Theory, incorporating mixed methods, as outlined by Strauss and Corbin. According to Strauss and Corbin, [6] 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. [7] The Grounded theory approach suggested by Strauss and Corbin [6] 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.

3. Data Collection

Data collection began in with the compilation of a simple desk survey of Web Science teaching institutions, recorded in a spreadsheet. Research so far shows that Web Science is most commonly taught at postgraduate level, with 15 out of the 21 institutions identified offering MSc programmes, some of which offer both MSc and progression to PhD. There were 7 examples of institutions offering PhD study, and 3 examples of Web Science study or research at an unknown level. Out the 21 institutions, only the University of Southampton and Eindhoven University of Technology offer Web Science as undergraduate programmes. The initial list of Web Science teaching institutions identified was used to compile a more detailed survey of individual modules offered by the programmes. A total of 142 modules were identified across 15 teaching intuitions which offered publicly available module information on their websites.

4. Data Analysis

The 'topics taught' field of the modules spreadsheet was imported into the qualitative data analysis package, Nvivo. It was then analyzed utilizing a process of manual thematic analysis in order to identify the most frequently occurring 'key topics' across all the modules. The key topics will be refined through further coding iterations, and compared with the key topics identified through a process of automated keyword analysis. The refined key topics identified will ultimately form the basis of the proposed taxonomy of Web Science key subjects, providing a curriculum outline for Web Science. The list in table 1 shows the top 15 key topics taught across all the Web Science taught programmes identified. While this requires further study and

comparison with other sources this gives a good initial indication that these may be the most frequently occurring key topics taught as Web Science across the world. Further study is required in order to see how this compares with key topics published within other data sources such as the Web Science conference publications to date, however an existing study by Hooper et al's [8] study gives an initial indication of how the two compare. Hooper et al [8] conducted an analysis of the materials published during the first 3 Web Science conferences. Utilizing a process of 'bolometric mapping' or keyword analysis, they identified the 15 most frequently occurring topics within the first 3 Web Science conferences from 2009 to 2011. These are shown below in Table and compared with the top 15 most frequently occurring key topics identified during the manual coding process carried out during this current study. This provides an interesting comparison between what was published during the first 3 Web Science conferences, and what is currently being taught as part of Web Science programmes.

It seems that contrary to Hooper et al's [8] [9] observations regarding a lack of subject coverage within conference publications, the opposite appears to be true for the key topics identified in from the Web Science taught modules. For example, subjects such as law and economics, while deemed to be under represented within the Web Science conferences, appear to be fairly well represented within the Web Science taught modules. There are some direct overlaps evident between the two lists, notably, social networking and semantic web. However, many of the terms do not exhibit direct overlaps. This may be indicative of differences in what is published and what is taught in practice; however further investigation is needed to confirm this.

Table 1. Comparison of results – Hooper et al [8] topics and key topics from this study

Top 15 Topics Hooper et al [8]	Freq.	Top 15 Key Topics (This Study)	Freq.
Computer Science	31	Networks	24
Semantic Web	26	Web Society	23
Real world	18	Semantic Web	21
Social networking	15	Sociology	17
Network Analysis	12	Network Theory	16
Social web	12	Research Methodology	16
Open source	11	Web Languages	15
Information retrieval	11	Social Networking	14
Open data	10	Web Architecture	12
Web users	9	Business	11
Web data	9	Economics	11
Search results	8	Law	10
Information science	8	Big Data	9
Web search	8	Information and Communications Technology	9
Knowledge management	7		

5. Conclusions and Future Work

This paper has detailed the bottom up, Grounded Theory, mixed method research approach used during this study, explaining why this method is suited to the research of Web Science, due to the subject's rapid evolution. A series of Web Science teaching institutions and modules have been identified; these were then analyzed and coded utilizing the qualitative data analysis package NVivo in order to identify key topics. The key topics were identified from manual coding of the topics taught within Web Science modules. The Key topics were compared with a study [8] of the most commonly published topics within Web Science, in which a keyword analysis of the first 3 conferences were presented. This initially suggests that there is greater diversity in the subjects taught than the range of subjects published in the conference, although further research is required in order to confirm this. Key topics will be compared and contrasted with the

only existing example of a formal attempt to create a Web Science curriculum (the Web Science Subject Categorization) and other Web Science literature, for example, the Web Science conference series. This will be supplemented by further study into the backgrounds of Web Science students. By interviewing students who have completed the Web Science MSc programme at the University of Southampton, it will be possible to gain insight into the spread of disciplines from which Web Scientists originate, as well as an insight into first hand experiences of completing a Web Science programme. These findings will then be compared with the key topics identified from Web Science taught programmes.

The final intended outcome of this work will be to create a curriculum taxonomy and associated set of resources for Web Science, evolved from Web Science as it is taught. Once refined, the key topics identified will ultimately form the basic structure of this proposed taxonomy. This will be presented in a wiki format so that it may be accessed and contributions may be made beyond the initial lifespan of the project. It is hoped that this will encourage further interest and debate within the wider Web Science community, in an attempt to encourage further work into the development of a formal curriculum for Web Science.

6. References

[1] T. Berners-Lee, D. J. Weitzner, W. Hall, K. O'Hara, N. Shadbolt, and J. a. Hendler, "A Framework for Web Science," *Found. Trends® Web Sci.*, vol. 1, no. 1, pp. 1–130, 2006.

[2] T. Berners-Lee, W. Hall, J. Hendler, N. Shadbolt, and D. J. Weitzner, "Creating a Science of the Web," *Science (80-.).*, vol. 313, no. 5788, pp. 769–771, 2006.

[3] S. White and M. N. Vafopoulos, "Web Science: expanding the notion of Computer Science," in *43rd ACM Technical Symposium on Computer Science Education, 28th February-3rd March 2012*, 2012.

[4] The Joint Task Force on Computing Curricula (Association for Computing Machinery IEEE-Computer Society), *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*. 2013.

[5] S. White, M. Croitoru, S. S. Bazan, S. Cerri, H. C. Davis, C. Jonquet, G. Prini, F. F. Scharffe, S. S. Staab, T. Tiropanis, M. Vafopoulos, M. Croitoru, S. S. Bazan, R. Folgieri, F. F. Scharffe, S. S. Staab, T. Tiropanis, M. Vafopoulos, S. Cerri, H. C. Davis, C. Jonquet, G. Prini, F. F. Scharffe, S. S. Staab, T. Tiropanis, and M. Vafopoulos, *Negotiating the Web Science Curriculum Development through Shared Educational Artefacts*. Koblenz, Germany, NY, USA: ACM, 2011, pp. 11:1–11:8.

[6] A. Strauss and J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, vol. 3. 1998.

[7] K. Bryant, A, & Charmaz, *The Sage handbook of grounded theory*. Sage Publications, 2007.

[8] C. J. Hooper, N. Marie, and E. Kalampokis, "Dissecting the butterfly: representation of disciplines publishing at the web science conference series.," *WebSci*, pp. 137–140, 2012.

[9] C. J. Hooper, G. Bordea, and P. Buitelaar, "Web Science and the two (hundred) cultures: representation of disciplines publishing in Web Science," May 2013.