UNIVERSITY OF SOUTHAMPTON
Faculty of Engineering, Science and Mathematics
School of Electronics and Computer Science
Intelligence, Agents, Multimedia Group

A mini-thesis submitted for transfer from MPhil to PhD

Supervisor: Professor Nigel R Shadbolt
2nd Supervisor: Dr David E Millard
Examiner: Dr Nicholas M Gibbins

An Investigation into Automatically Captured Autobiographical Metadata, and the Support for Autobiographical Narrative Generation

by Mischa M Tuffield

September 29, 2006

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE
INTELLIGENCE, AGENTS, MULTIMEDIA GROUP

A mini-thesis submitted for transfer from MPhil to PhD

by Mischa M Tuffield

Personal information and the act of publishing multimedia artifacts to the World Wide Web is becoming more and more observable. This report presents an infrastructure for the capturing and exploitation of personal metadata to drive research into context aware systems. I aim to expose ongoing research in the areas of capture of personal experiences, context aware systems, multimedia annotation systems, narrative generation, and that of Semantic Web enabling technologies. This report details the work underway towards the goal of creating a multidomain contextual log, and is followed by a discussion of how this work is being used to drive the development of a multimedia asset management system infrastructure.

Practical work already completed as part of this project is then described, focusing on the areas of data integration, infrastructure, and personal photo collection annotation, all in the Semantic Web context. The future plan draws together the work completed and presents how the autobiographical context captured will be evaluated to provide insight into the utility of the metadata harnessed to aid human memory management. This is followed by the proposed evaluation of how suitable semi-automatically captured metadata can be to facilitate the construction of a narrative of a person's observable life.

Contents

List of Figures

Acknowledgements

I would like the thank my supervisors Nigel Shadbolt and David Millard, for their irreplaceable guidance. I would also like to thank all the AKTors, and my friends in the IAM group for making me feel at home in Southampton. I would also like to give special thanks Antonis Loizou, Dave Dupplaw, and Hugh Glaser for their collaborative efforts during this last academic year.

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01.

Chapter 1

Introduction

1.1 Introduction

This report details an investigation into a novel approach for gathering and exploiting auto-biographical metadata. The metadata captured by our approach is presented as a contextual log of a user's digitally visible life. Ultimately, the plan is to use the acquired log in conjunction with the user's multimedia collection to facilitate auto-biographical narrative generation.

The term autobiographical narrative is used to describe a multimedia presentation of a person's life for a specified time interval. Narrative generation is presented as a means to evaluate the quality of the archived contextual log, the richness of the metadata gathered will be measured by its ability to correctly highlight the salient observable events of a user's life.

This report enumerates the data sources harnessed, the infrastructure developed, and elaborates an example use for the knowledge elicited - namely that of enriching personal photo-collections. The surreptitious nature of the logging approach presented has been core to our research agenda.

The end result is a methodology for the capture and storage of personal metadata and is proposed as a framework for *multimedia asset management*. This research also aims to provide insight into the utility of the metadata collected when considering problems of recalling and browsing personal multimedia artifacts.

This article presents the motivations in terms of our scope of research, background information, and related work, requirements and design issues considered. The details of initial experimental work, and the infrastructure developed are presented and evaluated. The research agenda is set forward, and the future work plan is defined.

1.2 Report Structure

Chapter 2 discusses the motivations behind this line of research, and divulges a critique on the existing literature. The foundations presented in the following section will include the relationship of the work to the Semantic Web (SW) Vision along with identifying the role our approach could play towards its uptake.

Chapter 3 starts out by enumerating of the sources of information captured and harnessed by the infrastructure developed, focusing on the issues discussed in chapter 2. The choice of metadata gathered is also grounded in the experimental work of annotating personal photo-collections. This section describes the technologies exploited by the approach taken, and presents the manner that the infrastructure facilitates the interrogation and integration of the captured metadata.

Given the infrastructure detailed, chapter 4 identifies the future direction of the work capitalising on the work completed, identifying the scope of our future research, while stressing the end product of this proposed research endeavour.

Chapter 2

Motivations and Related Work

2.1 Information Overload

A key problem faced by the information management community is that of information overload, or *infosmog* (Shadbolt and O'Hara, 2003). The ever growing and evolutionary nature of the World Wide Web (Web) is not making the task of information management any easier. The web is populated with an unmanageable amount of heterogeneous data, in forms as diverse as image, sound, video, and text. This coupled with the shift in the consumer electronics market from the analogue to the digital recording medium and the mass-market, low cost nature of the hardware industry has produced a rich *corpus* of personal multimedia artifacts.

The current trend of publishing personal information to the web, in the form of calendars, photo collections, GPS data, and the adoption of standards such as iCal files, the MBOX family of file formats[1], and the abundance of web accessible services such as flickr[2] or plazes[3], is presented as liberation of personal information. This liberation is seen as a social shift towards the self publishing and archival of personal information.

This report presents a study into the capture and aggregation of personal information, and is presented as the basis for *context based multimedia asset management*. The utility of the metadata enumerated in this work is to be evaluated as a method of annotating personal multimedia collections for browsing and retrieval purposes, and as a platform to generate autobiographical narratives.

There is no method of querying the Web that would result in the creation of an autobiographical breakdown of your web accessible digital life. The metadata identified in this work will be evaluated based on its ability to deliver accurate and engaging overviews

[1] Mbox Specification: http://www.qmail.org/man/man5/mbox.html/
[2] Flickr: http://www.flickr.com/
[3] Plazes: http://www.plazes.com/

of a user perceived digital trail. Ultimately, our plan is to relate personal multimedia artifacts together developing narratives from the annotations captured. The aim is to provide insight into the robustness of automatically generated autobiographical metadata to support narrative generation.

2.2 The Abundance of Personal Information

As eluded to in the previous section we are currently faced with a dramatic change in the way people publish personal multimedia artifacts. The popularity of social software sites like flickr, plazes, del.icio.us[4], and Last.fm[5] can be perceived as a new found desire to publish personal information on publicly accessible websites. The benefits of this social nature, i.e. the interaction amongst peers, of the aforementioned sites have been presented as one of the key drivers behind their success (Marlow et al., 2006). Other factors such as: the ease of publishing, and vanity have also been cited in the literature (House and Davis, 2005). The number of people adopting shared vocabularies to document certain phenomena, flickr for images, iCal for calendars, is increasing by the day. This apparent willingness to post personal information on the web, was a key motivator of this line of research.

The focus of the work has been on the identification and integration of personal information into a web-accessible Knowledge Base (KB). The observed trend in self publishing has also influenced our decision to make sure the described *'metadata capture and storage process'* is as surreptitious as possible. Assuming that people have found a new motivation to publish personal information does not justify a user's desire to recreate any information they have authored before, hence stressing the serreptitious nature of this approach.

The surreptitious chronological capture and archival of personal information is presented as the foundations for our study into personal *'context'*. The autobiographical metadata curated by this work has had its utility grounded in the application of a context aware multimedia asset management system , an image annotation system called Photocopain (see section 3.2).

The end product of this work is seen as a personal 'black-box' for the storage of autobiographical metadata. The web-accessible nature of the system is motivated in the following section.

[4]Social Bookmarking site, del.icio.us: http://del.icio.us/
[5]Last.fm: http://www.last.fm/

2.3 The Semantic Web and Personal Metadata

The work carried out has built upon a number of Semantic Web enabling technologies (see section 3.1.2). The information identified and captured as a result of our metadata acquisition system is represented by models adhering to as many W3C[6] recommendations as possible. This decision was taken to ensure that the metadata published would be exploited by as many web services as possible, building on the Semantic Web Vision (Berners-Lee et al., 2001) of interoperable web accessible resources.

Our system builds on the ideas brought forward in the original *Scientific American* Semantic Web article, with a particular focus on the notion of assembling, and integrating personal information into web accessible resources (Shadbolt et al., 2006). At his keynote speech during the International Semantic Web Conference 2003 (Berners-Lee, 2003) Tim Berners-Lee identified the *'Killer App for the Semantic Web'*, not as a single application but the successful integration of information, or to use his blunt words, *'Its the integration, stupid!'*.

In an attempt to avoid sounding too evangelical I present work that integrates a number of sources of information (identified in Section 3.1.3), to build up personal metadata chronology. It is important to allow users to select how much information they wish logged, or wish exposed to the Semantic Web. Since this is a matter of preference it can not be dictated, albeit this directly affects the richness of the metadata assembled.

This work integrates a variety of tools (see section 3.1.2) from the Advanced Knowledge Technologies project[7] (AKT), a six-year interdisciplinary collaboration working in the general area of technology and infrastructure development for the Semantic Web. AKT's aim is to extend and develop integrated technologies and services, using the SW as a unified basic framework, for acquiring, retrieving and publishing content over the whole of the knowledge life cycle.

This research is presented as a means to populate the Semantic Web with personal metadata, by exposing information in a structured and standard form, i.e. by using the Resource Description Framework (RDF) (Manola and Miller, 2004) accessible through SPARQL endpoints (World Wide Web Consortium, 2005). The system uses a Universal Resource Identifier (URI) to point to a user's Friend of a Friend (FOAF)[8] file. In the case that the user does not have a FOAF file the system can generate a basic one upon registration, allowing them to edit it as they see fit. Each users FOAF file serves as a unique identifier for their RDF data. The user's FOAF URI is employed to log the provenance of all the information asserted in a personal KB.

[6]The World Wide Web Consortium: http://www.w3c.org
[7]The AKT project: http://www.aktors.org/
[8]Friend of a Friend: http://www.foaf-project.org/

In an attempt to realise some of promises road-mapped by the Semantic Web community: the seamless integration of heterogeneous data, and that of services exploiting existing machine-accessible knowledge (Shadbolt et al., 2006), a decision to create an easy to use infrastructure allowing users to store, update, visualise, and query their own personal knowledge base(s) through the web, seemed a pragmatic course of action and is described in full in section 3.1.

It comes without saying that such a system can never be omniscient, our aim is to identify how much knowledge can be generated through the integration of the sources of information identified. It is important to stress that a user is not required to expose any of the personal information presented in this report.

The sensitive nature of this metadata-chronology, along with the current trend in using web-based social software, implied that the system developed had to allow users to decide whether any information captured was to be posted for public consumption or not. This guided the design insofar as each user is provided with two different KBs: a public, and a private *persona*. A user's private KB is presented as a means to enrich their own multimedia assets, and to run personal context aware services from. Whereas the knowledge in the public KB is presented as a means of populating the Semantic Web with publicly available structured knowledge.

Earlier I presented a shift in the publishing of personal information away from the private desktop to the public Internet. The decision to aggregate this information in a central KB is justified under the premise that having all of the personal metadata stored in a common knowledge representation language (RDF), will allow for more relationships to be discovered, taking advantage of knowledge from all the sources in the contextual log.

2.4 Memories for Life

Memories for Life[9] (M4L) has been championed as a grand challenge for UK computing. M4L focuses on the exploitation of technology to support human memory, and I believe that this research endeavour can be seen as a step in that direction. M4L draws together a number of recent technological and scientific advances from both life sciences and computer science in the area of human and artificial memory. M4L focuses on the use of technology alongside human memory in context, to provide support for memory management.

The focus of our work in relation to M4L is that of addressing the challenges of capturing, storing, and exploiting autobiographical metadata to support memory management. An evaluation of the metadata captured will be presented in terms of how useful the different forms of metadata are when considering the task of recalling personal experiences.

[9]The Memories for Life Network: http://www.memoriesforlife.org/

Given a system that could store a comprehensive collection of a lifetimes worth of acquired electronic media I detail a unified method of marking up this inherently heterogeneous data-set. This is studied to advance insight into the availability and utility of metadata as a way to sort through one's personal multimedia collection. Each of the sources of information modeled in our approach are presented as instances of events making up the user's digitally visible life. Our system presents a vocabulary of terms and relationships presented as a means of annotating these personal memory artifacts.

Human Memory is notoriously unreliable, and given the amount of people which are continuously trying to find out information about their past, through lineage websites, museums, photo collections, post-it notes, and so on, are examples of how humans are constantly looking at ways of documenting their lives for future reference. The research aims to develop a web accessible 'black box' of a users undertakings, allowing for future context aware services to be developed.

2.5 Context, Multimedia Asset Management, and Narratives

The goal of the proposed research is to identify readily available sources of information (see section 3.1.3), and combine them into a chronological autobiographical contextual log. The utility of the metadata gathered will be evaluated with regards to the application of searching and browsing personal photo-collections, and the ability to identify and piece together narrative structures within the knowledge base. The remainder of this section introduces the work completed and highlights the work to come, placing the related work section into context.

The Semantic Squirrels Special Interest Group[10] (SSSIG), based at the University of Southampton, aim to identify methods of logging available raw data , (or *'nuts'*), that describe aspects of a person's interactions with their computer and their physical environment(GPS track-logs and photos). The group is also concerned with identifying potential uses for the information gathered and this work is presented as just that, an exploitation of personal data.

A number of *squirrels* have been developed to gather the *nuts* available and it is these that have been propagated into our metadata logging infrastructure[11]. An ethos of the group is to preserve this raw data in order to retain any unforeseen potentials for exploiting the information gathered, this approach has also been adopted here. The group is keen to not process the information into a knowledge representation in order to transcend issues pertaining to platform, knowledge representation, and application restrictions, and this is where our research moves away from the work in the SSSIG.

[10]The Semantic Squirrels SIG: http://www.semantic-squirrel.org/
[11]The Semantic Logger Downloads Page: http://akt.ecs.soton.ac.uk:8080/downloads.php

This raw data forms the basis of the knowledge acquisition phase of our work, the sources of information utilised are enumerated in section 3.1.3. The infrastructure developed to store and interact with the captured raw data is detailed in section 3.1, and has been coined the Semantic Logger[12] (SL) (Tuffield et al., 2006a). The raw information identified by the interest group are parsed into RDF representations and then housed in the SL infrastructure.

The Semantic Logger is presented as a web-accessible means to capture, store, browse, and interrogate a user's contextual log. As mentioned earlier the contextual log acquired by the Semantic Logger is grounded in a multimedia asset management system, namely Photocopain (Tuffield et al., 2006b), an image annotation service. Photocopain integrates contextual information stored in the Semantic Logger, with content based information extracted from the images to generate metadata for one's personal photo-collection. The details describing the aggregation of the *context* and *content* based information are presented in section 3.2.

The generation of autobiographical narratives is presented as a task which requires a rich corpus of contextual information regarding a person. The ability to summarise a person's activities based on their contextual log is presented as a way of evaluating the quality and scope of the captured metadata. This is presented as a contribution to the utility of automatically generated metadata. This coupled with the proposed human centric evaluation of the metadata, given the task of browsing and searching personal photo collections, is presented as an investigation into the quality of surreptitiously gathered autobiographical context. The proposed method of evaluating the metadata acquired is that of a user based experiment facilitate through the mSpace faceted browser (section 3.2.6). This will allow for the RDF representation of a users log/photo-collection to be navigated through its various dimensions, identifying which metadata was most suitable for the task presented. This is seen as contribution to the study of the utility of metadata, and its use as a human memory management tool.

2.6 Related Work

2.6.1 Autobiographical Knowledge Management Environments

MyLifeBits (Gemmel et al., 2002) and SemanticLIFE (Ahmed et al., 2004) can be regarded as the modern seminal systems in this area building on the ideas put forward by Vannevar Bush (Bush, 1988) in his Memex device. While these have proven to be a valuable source of inspiration for this project, numerous others have been undertaken under the field of context based system (Cayzer and Castagna, 2005; Heath et al., 2005; Iofciu et al., 2005; Molle and Decker, 2005; Richter et al., 2005; Xiao and Cruz, 2005).

[12]The Semantic Logger: http://akt.ecs.soton.ac.uk:8080

The domain of interest of such systems is limited to the publishing, browsing, and sharing of information in tailored knowledge representations, where as this approach aims to employ as many existing models as possible. The infrastructure presented attempts to cater for any RDF data-structure uploaded to it, promoting flexibility.

The SemanticLIFE project uses a variety of techniques to extract metadata from personal communications, and considers issues relating to provenance and trust; these are presented in Weippl *et al* (Weippl et al., 2004). The SemanticLIFE project focuses on the capture and annotation of a user's 'work related life', where as our focus of interest is not confined to that particular subsection of a person's life.

MyLifeBits aims to organise and retell personal experiences using narratives as a structuring mechanism; (Gemmell et al., 2005) presents valuable insight on how stories can be generated based on information such as location and time. However, neither of these aforementioned systems strive to surreptitiously amalgamate as much contextual metadata as identified in this work. The metadata logging system presented below is presented as a low-cost metadata collection system, where as MyLifeBits is presented primarily as a method to to aid the human authoring and annotation of personal experiences.

SemanticLIFE is preoccupied with allowing users to set-up an information repository to provide enhanced querying capabilities, while MyLifeBits introduces the notion of automatically producing annotations by exploiting co-occurring events. In this section I set out to identify the principal differences between the Semantic Logger and such previously developed systems.

Such systems have engineered over-ranging knowledge representations to support the functionality they provide. The Semantic Logger makes no attempt to homogenise data that is heterogeneous by nature; this is left for applications that will use the system as a platform, as per their requirements. The rationale is that different mappings will be appropriate for different applications. A caveat worth mentioning is that data has to be in RDF.

The Semantic Logger aims to aggregate as much available personal information into a central knowledge base allowing for context-based systems (Falkovych and Nack, 2006) (Davis et al., 2005) (Adomavicius et al., 2005) (Tuffield et al., 2006a) (Tuffield et al., 2006b) to exploit it as needed. The Semantic Logger is currently being used as a platform for recommender system research, the work is elaborated upon in Antonis Loizou's work (Loizou and Dasmahapatra, 2006).

Another development worth mentioning is the NEPOMUK project[13], a EU FP6 funded collaboration of industrial and academic partners and industrial end-users. The project brings together various previous semantic desktop implementations, and focuses on

[13]NEPOMUK Project: http://nepomuk.semanticdesktop.org

knowledge integration in shared peer-to-peer environments, supporting automated community recognition. Detailed information on this has not been made available, however it seems that the focus is put once again in providing a solid platform for such sharing, rather than the ease of adding services to the system. For a full overview of the state of the art in developments in the field of semantic desktop research the interested reader is pointed to (Sauermann et al., 2005).

2.6.2 Multimedia Asset Management: Image Annotation

Photographs can be viewed as externalised additions to human memory; many human memory management tasks rely on collections of photographs, even if only shoe-boxes of photographic prints. Furthermore, as digital technology has dramatically increased the numbers of photographs taken (it has been estimated that up to 375 petabytes of information is created in photographic form annually), the problems associated with archiving and managing these photographs have become more pressing (describing their content, storing and retrieving them, and developing narratives to link them together). Image management is a highly labour-intensive task, and the Photocopain system has been designed with the intent of alleviating such burdens. Photocopain is presented as an example multimedia asset management system that utilises the contextual information captured by the Semantic Logger.

The generation of annotations from multimedia content can be seen as an application of image classification of the type used in content-based indexing and retrieval (Smeulders et al., 2000). Systems such as MAVIS2 (Joyce et al., 2000) and SIMPLIcity (Wang et al., 2001) use collections of image classifiers (embodied as agents, in the case of MAVIS2) to add semantic descriptions of the images' contents. Our system uses a combination of *content* and *context* based services to annotated personal photo-collections. The study of content based image retrieval techniques is still in its infancy, and are far from semantic, for a full discussion on the limitations of content based approach please see (Veltkamp and Tanase, 2000). Our approach combines annotations extraction from the content of an images, with the contextual information in the Semantic Logger, in order to annotate personal photo collections.

A series of experiments have been performed looking to identify the usefulness of various forms of metadata given the task of searching and browsing through personal photo collections. In (Naaman et al., 2004) an evaluation of a number of different types of metadata is presented with regard to the personal photo search and browsing. Namann's research aims to provide insight into which metadata categories are useful for mentally recalling and finding photographs, and is presented as a multimedia asset management system. Personal photo libraries usually have little discernible structure, and enriching the individual photos with metadata is presented as a means of organisation. The work captured the following metadata categories to enrich the photo collections: year, time

of the day, location, elevation, season, light status, weather status, and temperature. In section 3.1.3 I outline the sources of information captured, and present them as a super-set of the work undertaken by Mor Namann.

Work undertaken by Professor Marc Davis (Davis et al., 2004a) proposes three main categories of contextual information: spatial, temporal, and social. The Semantic Logger also attempts capture and archive all three of these forms of information, the infrastructure has been designed to allow for annotations and knowledge to be shared between peers. Work has also been published presenting the various aspects of digital photography (House and Davis, 2005) that of self-presentation, self-expression, memory archival, and that of social documentary. These findings are mimicked in our motivation of supporting human memory management.

2.6.3 Interfacing

User interface issues arise when designing how users may interrogate their contextual log and their annotated photo collections. Narratives and SPARQL are not the only ways envisaged for users to access the knowledge gathered, the incorporation of a faceted browser such as mSpace (m. c. schraefel et al., 2003a) (see 3.1), taking heed of the lessons learned from Haystack (Huynh et al., 2002), and Lloyd Rutledge's work on Making RDF presentable (Rutledge et al., 2005), have been applied to help the user navigate through the generated knowledge space. Here I present work that allows for mSpaces to be generated for a given SPARQL endpoint. The Photocopain system also allows for the incorporation with the AKTive Media (Ajay Chakravarthy, 2006) user interface to ease the burden of manual annotation. Our initial investigation has suggested that avoiding the use of *deep tree-lists*, as employed by MindSwap's PhotoStuff system (Halaschek-Wiener et al., 2005), and the adoption of a simple web interface similar to that of flickr, is seen as a step in the right direction.

The work undertaken to incorporate the mSpace faceted browser is envisaged to be more than just a user interface for the knowledge bases generated by the Semantic Logger and the Photocopain system. It is presented as a means to evaluate the personal metadata generated by the two aforementioned systems. Users will be asked to perform a task of searching and browsing through their personal photo-collections through the mSpace browser. The evaluation will be based on which aspects of the meta-data the users found most useful given the aforementioned task. This proposed experiment is presented as a follow up to the experiments undertaken at Stanford by Mor Namman (Naaman et al., 2004) which employed the use of the Flamenco faceted browser (Yee et al., 2003) and presented as a contribution to the utility of metadata dimensions as a form of memory support.

2.6.4 Metadata and Narrative Generation

The metadata captured by the Semantic Logger and Photocopain is described as a surreptitiously gathered personal information chronology. The Knowledge Base generated from the techniques described in this report is presented as a contribution towards the identification of how much personal metadata can be automatically generated with out imposing a cumbersome and intrusive to the end user.

A vast amount of work has been undertaken to design methods to generate narratives from bespoken knowledge bases (Alani et al., 2003b; Geurts et al., 2003; Mulholland et al., 2004). The proposed evaluation (see section 3.2.6) of the metadata identified in this work is presented as a contribution to the field of computer based narrative generation. The contribution being, the applicability of previously developed narrative generation methodologies to automatically captured, i.e. not *hand-crafted* metadata. In all of the aforementioned narrative generation systems the design of the knowledge bases have been tightly coupled with the generation technique. This research aims to identify which, if any, of already identified narrative generation systems are applicable to surreptitiously gathered metadata.

Chapter 3

Towards Autobiographical Context

This chapter presents a summary of the work undertaken thus far. The Semantic Logger (Tuffield et al., 2006a) is presented describing the sources of information harnessed, the infrastructure developed, and the resulting functionality available.

The Photocopain image annotator (Tuffield et al., 2006b) is then described as a multimedia asset management service that utilises the context gathered by the Semantic Logger. The discussion is followed with a brief evaluation of the system's annotation performance.

These systems are presented as the backbone for the evaluation of surreptitiously gathered metadata, and are presented as the foundation for testing the proposed hypothesis.

3.1 The Semantic Logger

3.1.1 Introduction: Semantic Logger

The Semantic Logger, currently live on the web, is presented as a system for the importing, housing, and harnessing personal information. The Semantic Logger's utility is grounded in two context-based applications, namely a photo-annotation tool (see section 3.2), and a recommender system. A description of the recommender system designed is beyond the scope of this report, more information can be found in (Tuffield et al., 2006a) and (Loizou and Dasmahapatra, 2006).

Upon registration of a Semantic Log, a user is presented with client-side tools that allow for the surreptitious recording of personal information, and server-side functionality allowing users to interact with their KBs. The presented list of information sources is far

from an exhaustive one, and is not intended to limit the functionality of the system. The Semantic Logger has been designed in a manner to allow information, in various forms of RDF to be posted to the knowledge base (KB). The sources of information identified and implemented are rationalised by the nature of the services currently provided by the system, and are merely presented as inspiration for future development.

One of the Semantic Logger's long-term visions is to provide a solid platform for evaluating the research agenda of the two developers, Mischa M Tuffield's and Antonis Loizou's, that of autobiographical metadata acquisition and narrative generation, and recommender systems. I argue that by virtue of knowledge integration alone, added value emerges. The principal support for this argument stems from the power of enabling the application of SPARQL queries on the available information, to answer questions that would be unfeasible under representations of singular domains. This along with the inferential capabilities of the knowledge representation are key to our approach of information integration.

For example:

- Queries

  How many users of the system attended the same events as me between time X and Y?

  This can be achieved by first selecting all events attended by the user between X and Y, using the iCal data, and then selecting all users with similar entries. If geo data is also available, it can be used extend and target the query.

  How many hyperlinks did I receive in email correspondence that I have yet to visit?

  A single query can be used to tackle this, by querying the email and browser history representations.

  What document was I reading on the way to event X?

  What was the name of the band I discovered while on holiday in Y ?

  This can be seen as the first step in the development of a queryable personal memory store.

- Inference

  The system does not require the user to produce hand crafted annotations. The existence of various domains in the knowledge base supports the automatic creation of such metadata. For example, iCal entries referring to the same time period as GPS location data can be used to provide suggestions for the name of the place with the specified coordinates, with respect to the current user.

The importance of exposing the system as SPARQL end-points is rooted in the motivation to allow applications to exploit the knowledge in unforeseen ways. Our attempts

to comply with as many W3C recommendations as possible is seen as a promotion of interoperability and is proposed as a contribution to the Semantic Web Vision.

3.1.2 Infrastructure: Semantic Logger

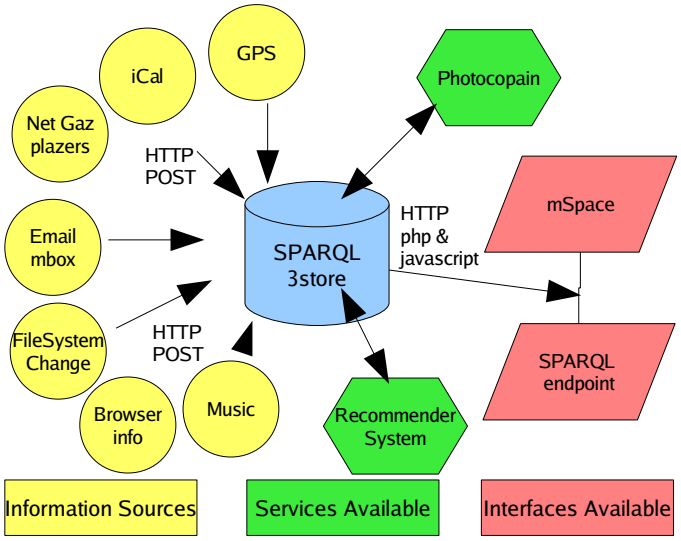

FIGURE 3.1: Overview of the Semantic Logger Architecture

The Semantic Logger is based on a service-based architecture, as shown in Figure 3.1, and has been designed so that new services may join on in an ad-hoc manner. The interactions between Web Services have been implemented using HTTP requests, namely HTTP POST, while the interactions with the central RDF triplestore make use of the SPARQL RDF query language (World Wide Web Consortium, 2005).

At the heart of the system is the AKT Project's SPARQL-compliant RDF triplestore 3store (Harris, 2005). The key role of the triplestore is to act as a persistent store for the system, and to mediate the interactions between the other system components. The main requirements in selecting an appropriate RDF Knowledge Base implementation were efficiency and consistency. 3store is a system bench-marked against other RDF storage and query engines such as Jena (McBride, 2001), Sesame (Broekstra et al., 2002) and Parka (Stoffel et al., 1997) and shown to outperform in terms of both efficiency and scalability (Streatfield, 2005) (Lee, 2004).

The cornerstone in designing this architecture has been to develop an open and accessible system, so that third parties can exploit the knowledge stored. I have chosen to expose two distinct methods of interacting with the system, namely in a public and private fashion. A number of distinct knowledge bases are maintained: A system-wide shared one - *public*, and one for each user - *private*, which is created automatically for the user upon registration. When information is imported into the system, users are able to

specify whether or not it should be publicly accessible. If this is the case, the information is added to both the shared and private knowledge bases. Both are exposed through web-based user interfaces to allow SPARQL queries on the data and the import, and removal of new knowledge. Furthermore, user interfaces have been designed to support browsing of the knowledge space, as it can not be expected of system end users to be fluent in SPARQL.

When data is represented in an RDF graph, by virtue of the representation there exists multiple dimensions in which the data may indexed and viewed. The mSpace interface (m. c. schraefel et al., 2003b) has the ability to organise such data, in multipane browsers. In addition, the edges of the graph are allowed to be reordered, using dimensional sorting independent of the hierarchical nature of the representation, allowing for a number of such trees to be visualised and browsed.

The software has been considered a good opportunity for visualising the data gathered, and a good method of supporting user studies to evaluate the metadata generated from the Semantic Logger and Photocopain. The experiments proposed is the searching and browsing of annotated personal photo collections, what is has previously been to eluded to, an experiment into *multimedia asset management*.

mSpace requires the definition of a *default column* and a *target column* along with the path through the ontological relationships (edges in the graph), between them to create a multi-columned re-arrangeable browser. While in the current implementation of mSpace[1] (0.6.2.3) these have to be made explicit by a knowledge engineer, an algorithm has been implemented to automate this procedure allowing users to choose their target column, i.e. the class they wish to navigate to. Furthermore, it is important to stress that this browsing ability is greater than that achieved through representations of singular domains, since the all the information logged by the system will be interconnected in automatically inferred or hard-coded ontologies. An example mSpace real estate model is shown in Figure 3.2

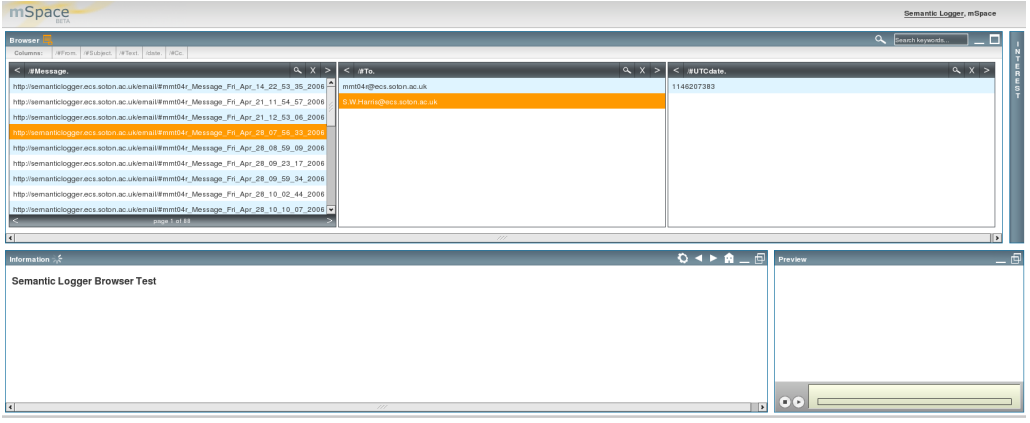

FIGURE 3.2: An example mSpace interface

[1]mSpace API definition: http://www.mspace.fm/

It is crucial for the Semantic Logger to impose the minimum burden on a user joining the system. Focus has been placed on allowing the import of knowledge described in heterogeneous, widely used vocabularies, to avoid the need for prior semantic agreement. The lack of an overall representation however, introduces the need for alternative means of knowledge integration. Where it is possible, this is to be achieved via automated means, such as the S-MATCH algorithm, developed by the University of Trento (Giunchiglia et al., 2004). Alternatively, where disagreement is too complex to be resolved in an automated fashion, mappings will be hard-coded into applications that use the Semantic Logger as a knowledge source, in *ad-hoc* fashion as per their requirements.

The richness of the metadata acquired, enables the system to be used as a platform for Community of Practice identification (Alani et al., 2003a). For example, named entity recognition can be applied to email correspondence to identify closely related groups while co-authorship and co-reference of scholarly articles can be analysed as shown in (Alani et al., 2003a). Co-location at various events can be inferred from geo-data and calendar entries, while the latter, in combination with the analysis of locally stored multimedia files (e.g. music and video files) can aid in identifying common interests.

A final feature of the Semantic Logger worth mentioning is the way the logger makes use of the FOAF model. A user's FOAF file, is used to allow a user to publish data about themselves, using a URI, allowing for the user's data to be referred to from any dataset, or from within any context. Another advantage of the adoption of personal FOAF files is the ability for a user to define his/her friends, allowing for further connections to be made when using the system to identify communities of practice. This feature supports the incorporation of social annotations as described in Marc Davis's work (Davis et al., 2006) (Davis et al., 2004b).

3.1.3 Knowledge Acquisition

The information sources presented in the knowledge acquisition phase, are a result of the work undertaken in the Photocopain project and the discussion brought forward from the Semantic Squirrels SIG. Our motivations and interests are similar to that of the Smilie Project at MIT[2], where they host a number of RDFizers[3], tools for converting information into RDF. This work was the foundation for the Piggy Bank project (Huynh et al., 2005) , this methodology differs from MIT's work insofar as it is not focused not supporting web-browser based harvesting of RDF. This work presents an easy to use, scalable, SPARQL compliant, accessible framework for the housing and query of RDF data. This infrastructure is presented along side a number of services to capture and upload contextual information.

[2]Simile Project: http://simile.mit.edu
[3]RDFizers: http://simile.mit.edu/RDFizers/

The SL collects, and propagates the following types of *'nuts'* into RDF representations. The metadata sources listed below are the foundations for the autobiographical contextual log captured:

- Calendar entries

  The SL has adopted the W3C recommendation in representing calendar entries in RDF [4]. A client-side application is available for download from the Semantic Logger site to automate the export of iCal (Dawson and Stenerson, 1998) files (commonly used and platform independent) into this representation. In addition to querying capabilities as before, calendar entries can serve as context indicators for geographical locations (described below), enabling to an extent the resolution of co-location.

- Geo-Data

  In an attempt to build up a log of a user's geographical data, a two pronged approach is taken. For research purposes we have been carrying around GPS units to log our data, this information is extracted and parsed into an RDF representation, taken from `http://www.hackdiary.com/`. The RDF model builds on-top of the dublin core namespace[5], and W3C's recommendation for geographical data[6].

  GPS information is being used to track a user's change of location, but is not always a suitable method of tracking, for tall buildings, and movement between buildings within close proximity is hard to track, so a decision was taken to start employing a network gazetteer. The network gazetteer Plazes is currently being employed by the Semantic Logger. Plazes supplies the end user with client side applications that pick up a laptop's current network connection and provides information about the location if information has been entered for that WiFi[7] hotspot. Plazes provides a comprehensive API, and RSS 1.0 feeds, that export parsable RDF, of a users activity. A decision has been taken to adopt their namespace for the purposes of logging network activity.

  The combination of the GPS information, a user's network gazetteer (given that the user has a laptop computer), and a his/her iCal file, along with the Getty Geographical Name gazetteer, allow us to infer a user's geographical context.

- Music playcount statistics

  Audioscrobbler[8], is a music search engine based on a large collection of music profiles. These profiles are built through allowing the users to download and install plugins to their respective media players that propagate the information

[4] iCal RDF representation: `http://www.w3.org/2002/12/cal/ical`
[5] Dublin Core namespace: `http://dublincore.org/documents/dces`
[6] W3C Geo-Data Namespace: `http://www.w3.org/2003/01/geo/`
[7] Wireless Internet Network
[8] Audioscrobbler: `http://www.audioscrobbler.net`

to the system. The representation used to describe artists is Musicbrainz[9], a freely accessible knowledge base for the music domain, that publishes the data in their ontology in an attempt to provide a comprehensive music information service. These systems are currently in the process of developing their metadata vocabulary to be published in an ontology. In this interim phase, a local version is being developed to describe the data made available through their web-service API.

- Firebox bookmarks, downloads and navigation history

  By virtue of its cross-platform nature, the Mozilla Firefox has been selected as our web-browser of choice. Firefox exposes the download information in RDF form[10] and thus can be easily imported to the system. Scripts have been developed to parse the bookmarks and history data into RDF. The RDF model uses two namespaces taken from the Mozilla developers centre [11].

- Email

  A simple ontology has been constructed to describe email correspondence [12] as one of satisfactory quality has not been found to be readily available. A client-side application has been developed to parse and convert the widely used MBOX representation into the local format. The intended use of this information, in addition to the ability to query one's records, is to support the identification of communities of practice, under a predefined temporal context.

- File System Information

  Beagle[13] search indexes every file found on a user's computer. This is achieved by combining specialised analysis tools for extracting content from different file types. This creates a personal information space describing a computer at the file-system level. The information is parsed into a simple ontology and can be loaded into a user's SL. This enables services to detect the presence and usage of files, giving an indication to a user's interests.

- Weather Information

  Weather information is proposed as a means of putting an event into context. The weather service that captures weather information given a time and location is still under development. Work is underway to harvest data from `www.weather.com`.

- News Headlines

[9] Musicbrainz: `http://musicbrainz.org/mm/mm-2.1`
[10] Netscape namespace: `http://home.netscape.com/NC-rdf#`
[11] Mozilla namespace: `http://developer.mozilla.org/en/docs/XUL_Tutorial:_RDF_Datasources`
[12] Semantic Logger email namespace: http://semanticlogger.ecs.soton.ac.uk/email/#
[13] The Beagle Project: `http://beagle-project.org`

The capturing of News Headlines is also presented as another method of enriching personal context. Work is underway to harvest New Headlines from the BBC website[14]. This is presented as another means of placing events into context.

3.1.4 Services: Semantic Logger

The following section presents some services that integrate the information stored in the Semantic Logger. These services are proposed as a method to enrich the knowledge found in both the public knowledge base and the individual private knowledge bases.

It is assumed that the human selection process is better modeled through a dynamic function that operates on some weighted subset of an artifact's physical and contextual attributes. Defining this subset statically at the outset is expected to have a negative effect on the quality of the knowledge base. The Semantic Logger architecture employs a variety of components, each capable of performing a subroutine of the knowledge integration. These are combined at run-time to find correlations in the knowledge housed in our infrastructure.

- Clustering algorithms

  Clustering algorithms will be used to partition the dataset into groups of similar items and users. For users, this is achieved through exploiting subsets of the information available in their profile, while in the case of items the clustering is carried out by considering a subset of the descriptive features available for them. A wide variety of such algorithms is needed to facilitate the architecture and instances are chosen based on their past performance under similar contexts, as logged by the system. The Semantic Logger infrastructure allows for a number of novel approaches to such clustering, such named entity recognition in email correspondence, co-authorship, co-location inferred from GPS data, event attendance from calendar entries, file system similarity and so forth. This clustering algorithms will be used to highlight the salient/recurring events logged in a user's Semantic Log.

- Ontology aligners/Aggregators

  Heterogeneity exists between the representations of different types of resources. In order to assess similarity the system will need to acquire the relevant partial translations from those representations to a temporary shared one, which will be discarded after the process is facilitated. Since it would be unfeasible to define a representation to which any user-defined ontology can be translated to, a variety of such components will be implemented to enable different modes of generalisation or specialisation. Work has been undertaken to identify same place as, and same day, week, month as relationships in the Knowledge Bases.

[14]BBC News website: http://www.bbc.co.uk/news/

- Recommenders

  These are the components responsible for evaluating the context of a recommendation need and for selecting the components that will be used to produce that recommendation. Recommenders will also receive predicted ratings computed by aggregators and augment them according to the recommendation context. Different recommenders may use other component selection, and ranking strategies to improve performance in specific contexts. The bias in choosing a particular instance is again determined by its past performance. These are elaborated upon in (Loizou and Dasmahapatra, 2006).

3.1.5 Conclusions: Semantic Logger

The Semantic Logger, a system for logging personal information has been presented as the foundation for our proposed future work and evaluation. The design and resulting infrastructure have been grounded in two systems, and the web service aspect of the architecture has been stressed to encourage future developments. It goes without saying that the Semantic Logger will require users to supply data for it to be a success, for it to be a sucess the system has been released to the public for testing purposes and to help locate any bugs.

OpenKnowledge (OK) [15] is an open peer to peer semantics based system, that accommodates knowledge sharing through interaction. Peers are able to carry out tasks and collaborate through executing interaction models associated with the task, provided they have downloaded appropriate software capable of satisfying any constraints posed in such models. These interaction models are expressed in the Lightweight Coordination Calculus (Robertson, 2005) and peers assume the roles described in them and carry out message passing to accomplish a successful execution. There is an intention to define the interaction between the components of the Semantic Logger in such models, such that peers that join the OK system are able to automatically use the functionality made available by the Semantic Logger.

I indent to set up a Wiki, to allow people to present new sources of information, or new services as they see fit. It can be envisaged that software designed to log a user's video viewing habits will be developed, to aid richness of our contextual log, allowing for further context. Integration with the Google Maps API [16] will allow for the information to be displayed and browsed through a geographical landscape, complimenting our faceted browser. The benefits of using both of the aforementioned interfaces, have been motivated in (Naaman et al., 2004).

[15] Open Knowledge Project: http://openk.org
[16] Google Maps API: http://www.google.com/apis/maps/

3.2 Photocopain

3.2.1 Introduction: Photocopain

Photo annotation is a resource-intensive task, yet is increasingly essential as image archives and personal photo collections grow in size. There is an inherent conflict in the process of describing and archiving personal experiences, because casual users are generally unwilling to spend large amounts of effort on creating the annotations which are required to organise their collections. This section describes the Photocopain system, a semi-automatic image annotation system which combines information about the context in which a photograph was captured, from the Semantic Logger, along with information from other readily available sources in order to generate outline annotations for that photograph that the user may further extend or amend.

3.2.2 Photocopain: System Architecture

In Tuffield *et al* (Tuffield et al., 2006b) the Photocopain photo annotation system is presented as a stand alone system that utilises context and content based methods to generate metadata to enrich one's personal photo-collection. The semi-automatic nature of the service is stressed, identifying the need to allow a user to author any proposed annotations, highlighting the *'Gold Standard'* of any manual annotations. The integration of a number of sources of highly heterogeneous data, along with the combination of low-level content based feature vectors, allows us to suggest annotations to the user. Ultimately, our plan is to discover relationships between photos, based on the meta-data assembled by Photocopain and stored in the Semantic Logger to generate an autobiographical narratives.

The remainder of this discussion is focused around the Photocopain system, this will include a list of information sources utilised, and will be followed by an insight into the advances made to the system incorporating it with the Semantic Logger. As with the Semantic Logger, Photocopain performs best when presented with many sources of information. The utility of Photocopain running off a user's Semantic Log is proportionate to how much knowledge is stored in the personal Knowledge Base.

3.2.3 Photocopain: Annotation Sources

Work has been undertaken to examine the potential sources of information that can be used to produce annotations, with a particular emphasis placed on ambient contextual information which can be applied to photographs with minimal effort on the part of the user. At this stage, the annotation task was considered in abstract terms, aiding the

making of the conscious decision to examine the inputs to the annotation task, rather than the structure of the task itself, for two main reasons:

First, with this research, the intention is to examine mechanisms and architectures for acquiring and organising the available information that are cheap, in that they do not require a great deal of human intervention. Our initial research milestone was to discover which information sources *are* cheap, as opposed to appearing cheap; easily acquired information may have hidden costs associated with its representation, or its integration with other information.

Secondly, it is clear that a user carrying out the image annotation task has a number of different priorities. He/She will be interested in different aspects of the photography depending on the specific task they are performing: annotating a photo library for the press; organising a personal collection of photographs; or selecting CCTV stills during a security operation. Rather than developing a special-purpose system that supports a specific task, a flexible system has been developed; it may be that the readily-available information that have been identified here is more useful for some tasks than others, the intention was not to introduce biases by making assumptions about the specific goals of the end user.

The following sources of information are harnessed alongside the metadata stored in the Semantic Logger:

- Camera metadata

  Exchangeable Image File (EXIF) (EXIF, 2002) metadata records camera parameters at the time that the photograph was taken. These parameters include: aperture setting; focal length of the lens; exposure time; time of photo; flash information; camera orientation (portrait/landscape); and focal distance. We can also derive other information from these metadata, such as the average scene brightness of an image. The EXIF is extracted from the images, presented to the Photocopain system, and then uploaded to the 3store, in a RDF representation.

- Global positioning data

  GPS data can be recorded in EXIF if the camera is equipped with the required hardware, or alternatively a GPS tracklog matched with a photos timestamp can be used to determine location accurately. This is primarily of use when the camera is used outdoors. As described in *'Geo-Data'*, section 3.1.3 the GPS, the Network Gazetteer, the Getty Gazetteer, and the iCal information can be used to piece together a geographical log of a given user. All of aforementioned sources of information can be found in a user's Semantic Log.

- Image analysis, Classification, and Flickr

A selection of image analysis techniques, such as the CIELab colour-map, *Hue, Intensity, Texture (HIT) Map*, and the edge direction coherence vector have been used to propose annotations for image content (see section 3.2.4.2 for further details). A number of classifiers have been trained using flickr's image pool as our source of training data, these are elaborated upon in (Tuffield et al., 2006b). Flickr users may associate images with a number of free text tags (e.g. Tim BL, WWW2006, Edinburgh); I use the photographs associated with certain tags as training sets for our image analysis algorithms. For example one hundred and fifty images of the tag 'flower' where taken from flickr via its API[17], and then images that are not flowers were also downloaded from flickr, in order to train a classifier.

This process was automated by first identifying what words have been clustered together inside flickr (getRelated function via flickr api), and then listing the words have been clustered in conjunction with the list of flower's related tags. This list was then combined with the terms related to flower extracted from Wikipedia's Categories[18], and was used as a filter when randomly collecting a set of one hundred and fifty images I assume to be the class of 'not flower'. The RDF representation of Wikipedia categories can be accessed through the Semantic Logger using the username and password *wiki*.

I selected a handful of flickr's most popular tags to be our initial content-based annotations (vocabulary). These include: landscape, cityscape, portrait, group-photo, architecture, seascape, and flower. The decision to use this dataset has ensured that any proposed annotations are grounded within flickr's shared conceptualisation of these terms. For example, if Photocopain proposes the annotation 'landscape', what it actually means is 'this image is similar to images tagged landscape by the flickr community' as opposed to the developers' understanding of the word.

The web service based architecture developed for Photocopain, allows easy integration of new image analysis algorithms, and/or new clustering algorithms as needed.

Work is currently underway to build automatic classifiers for commonly occurring human annotations. If a user persistently annotates his/her photos with a tag not recognised, the system will attempt to collect some test and training data from flickr, and automatically try different combinations of the different image analysis algorithms, along side the different classifiers. Another initiative on this front is the notion of having the 'gold standard' annotations to feed-back into the system, retraining the classifiers in an incremental fashion.

- Community Annotations from the Semantic Logger

[17]The Flickr API: http://www.flickr.com/services/api/
[18]Wikipedia Categories: http://en.wikipedia.org/wiki/Wikipedia:Browse

> Given that Photocopain has been repurposed to work with the Semantic Logger, a piece of social software, the scope now allows for annotation's to be shared within communities. Friend information can be exploited in order for annotations to be shared with communities. Given that a friend of your took a picture at the same time and place will allow for annotations to be proposed by the system. This notion of using other peoples annotations was inspired by the undertakings of the ZoneTag project [19] at Yahoo!.

Photocopain shows how the information found inside the Semantic Logger can be used to enrich one's personal media library. And as mentioned before Photocopain is presented as a *multimedia asset management system*. Photocopain uses *content*, *context*, and *community* based knowledge in order to generate as much metadata as possible. Photocopain presents the user with a number of annotations for each image submitted, while these are in turn corrected by the user, and then uploaded back to the users Semantic Log. This process of importing this photo-specific information back to the Semantic Logger, adds another dimension to be exploited by the aforementioned recommender system.

These information sources are quite basic, and will not allow the creation of sophisticated annotations such as 'Yet another shot of Briony falling into the pool', they can be combined to derive other relevant information. Work undertaken my Marc Davis at Berkeley provides insight into how context can be combined with content to aid the identification of faces inside photographs (Davis et al., 2006). For example, the focal length, focal distance and sensor size could be used to calculate the size of an object in the focal plane, which might enable us to differentiate between a leaf and a tree, even though both are in the same location and have a similar colour. The aim of our system is not to create annotations per se, but to generate hints which guide the user as they create the annotations. As a further example, consider how colour information can be used combined with the technical information available from the camera and other devices. If I know that the picture was taken with a wide angle lens, the GPS log shows that the photo was taken in a city, and the picture is largely grey, I may be able to infer that the picture is of a building or a piece of architecture. Given such an inference, and possibly some user-provided confirmation, I might then deploy an image analysis algorithm to spot sharp edges.

Understanding of photographic genres is certainly relevant here. For instance, portraits are typically taken at focal lengths of around 100mm, with wide aperture (f/4 or above), fill flash (reduced power), and the subject of the photo, a head/shoulders, will be of a standard size. The background at top left and top right will be out of focus (low sharpness), and the photo will have a portrait aspect ratio. These are fairly standard settings which are used by the majority of photographers; they are also the settings that

[19] ZoneTag API: http://zonetag.research.yahoo.com/zonetag/

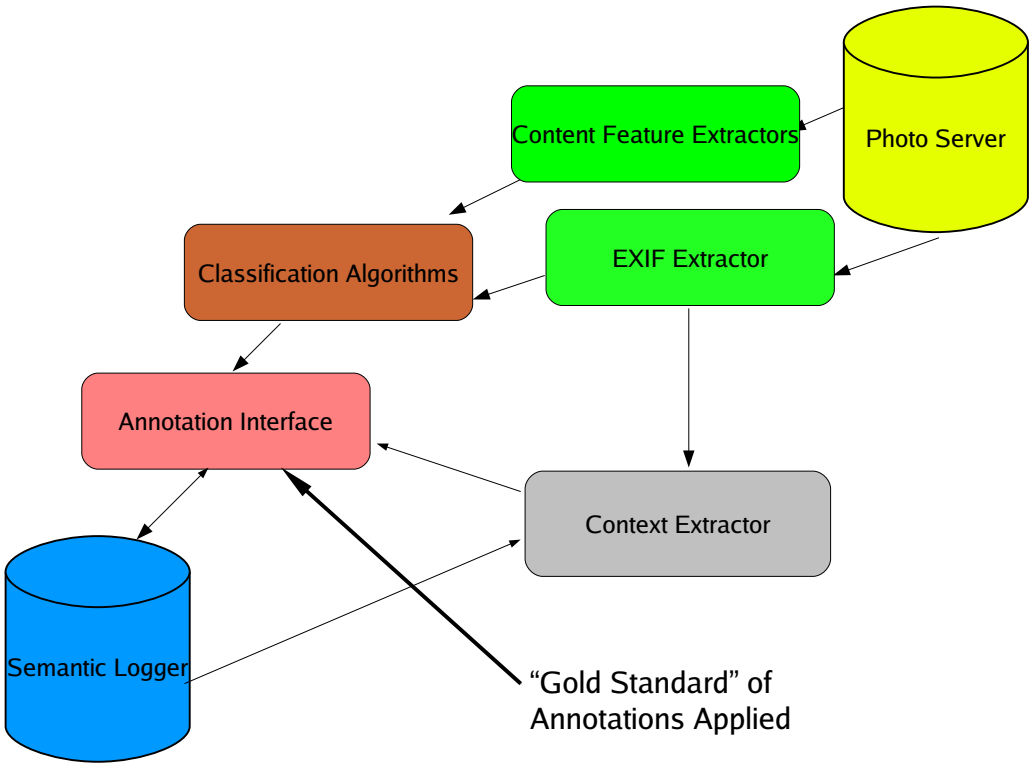

FIGURE 3.3: The Photocopain Architecture

an automatic camera set to 'portrait' will usually use. In the latter case, the photograph may record the automatic setting of the camera, in which case the annotation system may not need to search for the individual settings of the camera.

3.2.4 Photocopain: Architecture

The Photocopain system has a service-based architecture, as shown in Figure 3.3, and has been designed so that new annotation or classification services may be added on in an ad-hoc basis. The connections between system components have been implemented as simple Web Services (HTTP or SOAP) wherever possible, and interactions with the central RDF triplestore make use of the SPARQL RDF query language (World Wide Web Consortium, 2005). Below is a description of the components of the Photocopain system, illustrating their interactions with a simple workflow.

3.2.4.1 Annotation Server

The central component of the system is a SPARQL-compliant RDF triplestore (Harris and Gibbins, 2003), the Semantic Logger, which is used to store the image annotations,

both candidate and user-vetted. The key role of the triplestore is to act as a persistent store for the system, and to mediate the interactions between the other system components.

All of the contextual information harnessed by Photocopain is stored in the Semantic Logger, and subsequently after the annotation process all of the extra metadata produced by Photocopain is imported back into the user's Semantic Log, enriching the user's contextual log.

3.2.4.2 Image Feature Extraction

Most of the current work on semantic annotation of images focuses on attempting to solve the computer vision problem. Our guiding principle with Photocopain has been simplicity; study into the use of simple, well-understood techniques with additional metadata has been core to the development. The system is extensible so that new feature extraction algorithms can be included to extend its breadth of recognition, or its classification accuracy. For prototyping purposes four feature extraction techniques have been implemented. These were selected from the literature to see how they perform with the extra contextual information that our approach is able to provide.

A simple CIELab colour-map is used alongside other feature extraction methods for increasing the quality of a match. The colour map lays a 10x10 grid over the image and takes the average colour in the perceptually-uniform CIELab colour-space for each grid element. This spatial-colour map can be used to qualify classifications (such as providing more evidence towards a 'landscape' annotation due to the top of the image being the colour of sky, for example).

A face detection module provides an estimate of the number of faces that appear within an image. This uses the Hue, Intensity, Texture (HIT) Map (Mateos and Chicote, 2001) technique to detect face-like areas. The method performs a combination of thresholded hue, and intensity channels to obtain a 'skin' image. Clearly this is quite a simplification of the description of a face, but it works to identify skin regions well. After connected-component labelling and segmentation, regions are filtered based on shape and size allowing the module to return only those images which are likely to be faces. However, this process often leads to false positives, so combining this with the EXIF data from the camera lens, an approximate physical size for the possible face region can also be calcuated. As faces can only ever be certain sizes, this provides more evidence for identifying whether a region within the image is a face. Combined with other EXIF-based metadata it is possible to annotate images as being portraits, or group photos.

The artificial vs. natural feature extractor uses classification of the edge direction coherence vector (Vailaya et al., 1998) to classify image content, based on the assumption

that artificial structures tend to be created with straight-edges, whereas natural structures do not. A Canny edge-detector is used to filter the image and short edges are binned into a histogram. An edge-tracing algorithm is used to find those edges that are coherent, and these are also binned into another histogram. The two histograms make up the edge direction coherence vector, which can be classified against, or is able to provide a measure of 'artificial-ness'. In our early prototype photocopain exclusively classified an image as artificial or natural, however, it became clear that an image can be both artificial and natural simultaneously (for example a photo of a cityscape behind a park). Creating two classifiers, rather than one, allows images to be annotated with both labels if necessary.

A focus-map performs Fourier transforms over the image in a grid. These transforms are filtered to obtain a measure of high and low frequency components within each grid element. The measures are then linearly combined at each location to provide an overall map for the image, of where focus lies. In images that have a shallow depth of field, and therefore contain unfocused regions, the map is able to provide both evidence of this, and also information on where the feature extractions should be targeted, if possible. Combined with other techniques, this extraction may be able to provide more evidence towards portrait annotations (photos of which tend to have shallow depth of fields at a medium focal length), or landscape annotations (photos of which tend to have a large depth of field, focused near the hyperfocal distance).

3.2.4.3 AKTiveMedia Annotation Tool

The annotations produced by the image classifiers and the context-based annotators as treated as candidate annotations; while human-authored annotations are expensive, they are seen as a 'gold standard', and have considerable value by virtue of their provenance. The AKTive Media Tool (see Figure 3.4) is used as an interface to easy the burden of annotating the images at hand.

3.2.4.4 Metadata Output

Once a set of annotations have been generated for an image, they may be exported as RDF, and put back into the Semantic Logger. These metadata may also be sent back to to a web gallery (flickr) ; the RDF is used to add annotations to the image, allowing other systems to search and process the images. The mSpace interface has been chosen as our primary method of interaction with the images and their metadata.

One of our long term goals for this work, is the creation of narrative structures which present photographs in context, in order to better archive personal experiences. The annotations form a foundation for these structures, and enable chronological and narrative based queries over the data such as: "Where did we go in the last week of our

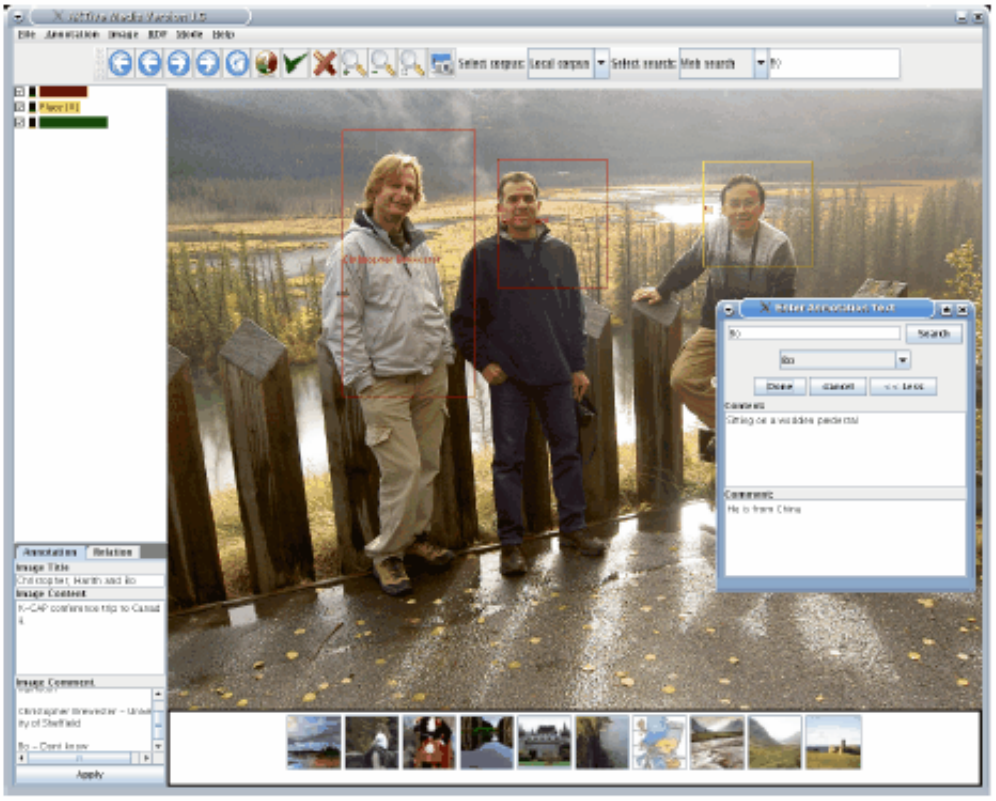

FIGURE 3.4: The AKTive Media Interface

holiday in the Peloponnese?" or "Who pulled me out of the crevasse I fell into in the mountains?" or even "What happened next?". This form of meta-level query is a valuable way of unlocking memories; we use the narrative structure to facilitate the most effective presentation of the data to the user whether that be as a slide show, themed collection (such as an online album), a structured tree (Spinelli, 1999), or multimedia presentation (Alani et al., 2003b).

3.2.4.5 Workflow

Broadly speaking, there are two parallel workflows shown in Photocopain: one concentrates on annotations which can be gleaned from the content of the image, while the other creates annotations based on its context, see Figure 3.3.

The content-based annotation workflow, contains an EXIF extraction service which retrieves the camera metadata from the image, and a set of feature extraction services which analyse the content of the image. These metadata and features are passed to the image classifiers, which perform simple classification on the images (e.g. natural vs. artificial, indoors vs. outdoors, and so on). Finally, these classifications are passed as candidate annotations to the user annotation interface.

The context-based workflow is executed along-side the content based services. The time-stamping information found in the EXIF metadata is used to identify contemporary events in the user's Semantic Logger account. The public Semantic Logger Knowledge Base is then queried to find any possible annotations used by other user's at the same place and time. The spatial, temporal, social, and content-based metadata is then passed to the annotation interface in order to be verified by the user. Upon submission of the 'Gold Standard' annotations, the knowledge is then uploaded to the user's Semantic Logger account.

3.2.5 Evaluation

Our initial evaluation has focused on the performance of the individual feature classifiers. I have evaluated our classifiers for indoor/outdoor images, and for artificial/natural environments. In both cases, I use a training set of 150 images, with 75 instances representing each classification, and a test set of 30 images.

For the classification of images as indoor or outdoor, I have found that a combination of the EXIF data and the CIELab colour map date without any dimension reduction performed the best, yielding only 4 errors in the 30 tests. Our classification of natural and artificial environments yielded its best results with a combination of the edge direction coherence vector, the CIELab colour map and the nearest neighbour clustering algorithm, giving 3 errors in the 30 test.

I am currently creating more automatic classifiers (flowers, seascape, cars), and changing the decision process from a binary one (i.e. either indoor or outdoor) to one containing certainty factors, such as indoor 90%, outdoor 5%, artificial 70%, flower 2% and so on, and will revisit the evaluation of our classifiers when this is completed.

3.2.6 Future Work

Once annotations giving a spatiotemporal context for a photograph have been established, it becomes possible to use narrative structures to organise image collections and exploit this context. Our current research is concentrating on the construction of narratives and stories to allow users to make better sense of the continuous stream of digital data they are acquiring (Zarri, 2001). These narratives could link a collection of photos taken when on holiday in a foreign country, or a collection of photos taken at conferences during the previous year. Assuming a story is the presentation of a succession of events based around a given topic, I aim to generate such narratives based on the resources available.

To this end, I propose a two-layered approach to allow maximum automation and full flexibility. The first layer is an automatically generated time line, based on the timestamps of the images, and using a default chronological granularity generated from the calendar spread of the data. At this level, user intervention is limited to altering the granularity of the time segments should the user feel this necessary. I intend to combine this with location information, which provides useful cues for generating stories to refresh existing memories (Gemmell et al., 2005).

The second layer organises photos into a narrative structure based on educated guesses made from the metadata, allowing users to correct or adjust the narrative; it is a core research issue in this project to determine the appropriate balance between user intervention and automatically generated narrative structures. In essence, the annotated images function as instances in an underlying narrative ontology (or narrative script), which will allow for sequential organisation of the instance events into a cognitively coherent narrative structure (Furtado, 1997; Hobbs, 2003). These scripts contain stereotypical narrative constructs (Uhlír and Falc, 2004) (based on the memory cue humans rely on: namely *who*, *what*, and *where* thought to be predominate in memory aid (Wagenaar, 1986)) from which the user specifies sequential and causal relations between images, and have been influenced by the definition of narrative concepts in our earlier work on the OntoMedia ontology (Jewell et al., 2005).

3.2.7 Conclusion

In this section, I have presented the Photocopain system, which integrates a number of existing and well-understood technologies that address the task of photo annotation, both those which rely on the content of the images, and those which examine their context.

Organising non-textual artifacts such as photographs in such a way as to facilitate retrieval is a difficult problem; the usual retrieval strategy of exploiting keywords is of course unavailable without annotation. However, annotation requires significant effort on the part of the user, and so I have focused on ways of automating or semi-automating the image annotation task. Photocopain takes an unusual route of using information that is readily and, in many cases, cheaply available. This information is extremely heterogeneous, and the much of it has low value (certainly compared to manually-generated annotations). Although the benefits may be small, the low costs of acquiring it may mean that some of these sources are still worth exploiting. It is argued that it is possible to process and integrate this sort of information, and that basic processing might yield further information. Such information, and inferences over it, could be used to help suggest annotations or annotative strategies for users, particularly in sharply circumscribed domains, and in relatively routine annotation tasks.

It is clear that much work has still to be done, and that the value of the various information sources needs to be evaluated in genuine task contexts. Nevertheless, given the cost of annotation, experiments on the value of found information is producing interesting results, and is presented as the foundations for our future research agenda.

Chapter 4

Conclusions and Future Work

In this final section of the report remaining work needed to complete the implementation process is identified. This will then be followed by a detailed description of the evaluation work to be performed, while identifying any foreseen technical challenges, this is used to scope the work proposed, and motivate the work plan presented. These experiments aim to support the validity of the contributions of this proposed research endeavour.

4.1 Future Implementation

In order to complete the implemention of the meta-data gathering infrastructure presented in this report, namely the Semantic Logger and Photocopain, the following features need to be finished.

- Named Entity Recognition

  In order to aid the discovery of ontological matches in the sources of information we have identified in section 3.1.3, a *named-entity recognition* algorithm needs to be implemented. This will allow for correlations to be found in the knowledge bases, for example we would like to recognise the fact that when my calendar file states that i have a meeting with Nigel Shadbolt, this is referring to the same Nigel Shadbolt found in my email correspondence, hence making an explicit connection in the KB. The SimMetrics[1] library has been selected to aid this task.

- Sources of Contextual Information

  The *News* and *Weather* information sources still need to be integrated. The process of capturing news and weather information given a specific time and local still requires some further coding. The parsing of the Getty Gazetteer[2] into an RDF

[1] SimMetrics Sourceforge page: http://sourceforge.net/projects/simmetrics/

[2] The Getty Gazetteer: http://www.getty.edu/research/conducting_research/vocabularies/tgn/

representation to be incorporated into the public knowledge base is still subject to some future work. This will allow for more meaningful, human friendly annotations to be made from the geographic information harnessed in this work, as well as allowing named entity recognition to be performed on the information captured via the GPS tracklogs and the network gazetteers, with text found in a person's calendar and emails.

- mSpace

  The mSpace model generation algorithm, need further optimisation to improve performance. The mSpace browser will also require further customisation to help visualise personal photo-collections. Identifying matches between different data sources will allow for multi-domain mSpace visualisation of the personal meta-data. This will then we used for the proposed human centric experimentation.

- Photocopain

  The Photocopain prototype (Tuffield et al., 2006b) was developed as a stand alone application that captured contextual knowledge on a *need to know* basis from a web-server. The system is currently being redesigned to run off the Semantic Logger infrastructure. Further development and evaluation into the performance the content based techniques and clustering algorithms, used in conjunction with flickr as a data set are at creating classifiers for images tags.

- Data-Sets

  In order to justify the validity of any future thesis, a substantial corpus of data is required. Antonis Loizou and I have been using the Semantic Logger to build up a contextual log for a few months, and I have asked and been offered Hugh Glasers last year's worth of data to experiment upon. A request is being presented to the Semantic Squirrels SIG for any further data available. Scripts are currently under development to parse a person's contextual log as a batch process.

The implementation proposed in this section, along with some final testing and debugging of the Semantic Logger is presented as at most a months worth of programming.

4.2 Future Research Work

The future research plan attempts to present two complementary experiments to test and ground the utility of the semi-automatic autobiographical metadata discussed in this report. These are presented below:

- Human-Centric Evaluation

As eluded to earlier in this report, I am proposing the design of a human centric personal image search and browsing task, similar to that undertaken by Mor Naaman (Naaman et al., 2004). This is presented as a manner of evaluating the utility of the various metadata domain we are using to support *personal photo collection asset management*. The results of this experiment is intended as a contribution to the identification of useful automatically captured metadata to aid memory recall.

The task will be supported by logging a user's interactions with the aforementioned mSpace interface, where each column will present a different dimension in the contextual knowledge. This will allow for the most used information sources to be exposed. The users will be asked to evaluate which contextual cues were the most useful for searching and retrieving images. This will provide some insight into the quality of surreptitiously gathered autobiographical metadata.

- Algorithmic Evaluation

  In order to accommodate any claim of being able to generate autobiographical narratives from an automatically captured contextual log, a number of previously identified narrative generation techniques will be implemented. Autobiographical narrative generation, or the summarisation of the activities of a person's life for given a time period, is presented as a highly context dependent task. Experiments into this task will involve the design of a timeline generator that will summarise the key activities of a person's life by mining the salient, recurrent events given a time granularity, placing them in order (Harada et al., 2004). The validity of these algorithms will be evaluated through inspection of the end narrative produced.

  Following on from the simple timeline generator, work is proposed to populate templates to summarise a persons actions. Templates similar to those in the ArtEquAKT project (Alani et al., 2003b) are proposed as a starting point to this evaluation. Work from the CWI research lab in Amsterdam (Falkovych and Bocconi, 2005) highlighting the issues with template approaches, and the importance of understanding genre are presented as valuable insight to this task. Lessons headed from the Digital Tapestries work (Rother et al., 2005) are presented as insightful with regards to techniques in how to juxtapose personal photos to tell the story of a given time interval. The results of the above experiments aim to provide insight into which forms of narrative generation are applicable given the unobtrusive nature of our knowledge acquisition process, hence grounding the research into automatically captured autobiographical metadata. The results of the future work will be described in context of the work I presented in my 9 month Report[3] into computer based narrative generation systems.

[3] 9 month Report: http://eprints.ecs.soton.ac.uk/12896/

4.3 Conclusions

This upgrade report aims to motivate the study of serrepititiously gathered metadata by presenting narratives as a natural way of evaluating the utility of the captured contextual. My future research will present whether semantically annotated photo collection metadata is useful to an end user, and whether the relationships explicit in the metadata could be harnessed to generate narratives. This paper has identified the shortcomings of existing systems, and will be followed by the development proposed in this future work section. This will also demonstrate whether ontology driven metadata and reasoning can bring the expressive power of the Semantic Web to user-centric context based systems.

Bibliography

G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *j-TOIS*, 23(1):103–145, January 2005. ISSN 1046-8188.

M. Ahmed, H. H. Hanh, S. Karim, S. Khusro, M. Lanzeberger, K. Latif, M. Elke, and Kha. Semanticlife: A framework for managing information integration and web-based applications and services. In *Proceedings of the 6th International Workshop on Information Integration and Web-based Applications and Services (IIWAS 2004)*, 2004.

Fabio Ciravegna Ajay Chakravarthy, Vitaveska Lanfranchi. Requirements for multimedia document enrichment, 2006.

H Alani, S. Dasmahapatra, K. O'Hara, and N. R. Shadbolt. Identifying communities of practice through ontology network analysis. *IEEE Intelligent Systems*, 18(2):18–25, 2003a.

H. Alani, S. Kim, D. Millard, M. Weal, W. Hall, P. Lewis, and N. Shadbolt. Automatic ontologybased knowledge extraction from web documents, 2003b.

T. Berners-Lee. Keynote speech: Semantic web status and direction. In *International Semantic Web Conference 2003(IWSC-03)*. Springer, Berlin, Heidelberg, 2003.

T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5), May 2001.

J. Broekstra, A. Kampman, and F. van Harmelen. Sesame: A generic architecture for storing and querying rdf and rdf schema, 2002.

Vannevar Bush. As we may think. pages 17–34, 1988.

S. Cayzer and P. Castagna. How to build a snippet manager. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

Marc Davis, Simon King, Nathan Good, and Risto Sarvas. From context to content: leveraging context to infer media metadata. In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 188–195, New York, NY, USA, 2004a. ACM Press.

Marc Davis, Simon King, Nathan Good, and Risto Sarvas. From context to content: leveraging context to infer media metadata. In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 188–195, New York, NY, USA, 2004b.

Marc Davis, Michael Smith, John Canny, Nathan Good, Simon King, and Rajkumar Janakiraman. Towards context-aware face recognition. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 483–486, New York, NY, USA, 2005. ACM Press.

Marc Davis, Michael Smith, Fred Stentiford, Adetokunbo Bambidele, John Canny, Nathan Good, Simon King, and Rajkumar Janakiraman. Using context and similarity for face and location identification. In *Proceedings of the IS&T/SPIE 18th Annual Symposium on Electronic Imaging Science and Technology Internet Imaging VII*. IS&T/SPIE Press, 2006.

F. Dawson and D. Stenerson. Internet calendaring and scheduling core object specification (icalendar), 1998.

EXIF. Exchangebale image file format for digital still cameras: EXIF version 2.2. Technical report, Japan Electronics and Information Technology Industries Association, 2002.

Kateryna Falkovych and Stefano Bocconi. Creating a Semantic-based Discourse Model for Hypermedia Presentations: (Un)discovered Problems. In *Workshop on Narrative, Musical, Cinematic and Gaming Hyperstructure*, Salzburg, Austria, September 2005.

Kateryna Falkovych and Frank Nack. Context aware guidance for multimedia authoring: harmonizing domain and discourse knowledge. *Multimedia Syst.*, 11(3):226–235, 2006.

Antonio L. Furtado. Narratives and temporal databases: An interdisciplinary perspective. In Peter P. Chen, Jacky Akoka, Hannu Kangassalo, and Bernhard Thalheim, editors, *Conceptual Modeling*, volume 1565 of *Lecture Notes in Computer Science*, pages 73–86. Springer, 1997. ISBN 3-540-65926-9.

J. Gemmel, G. Bell, R. Lueder, S. Drucker, and C. Wong. MyLifeBits: fulfilling the memex vision. In *MULTIMEDIA '02: Proceedings of the 10th ACM international conference in Multimedia*, pages 235–238, 2002.

J. Gemmell, A. Aris, and R. Lueder. Telling stories with MyLifeBits. *ICME 2005*, 8: 6–9, July 2005.

J. Geurts, S. Bocconi, J. van Ossenbruggen, and L. Hardman. Towards ontological-driven discourse: From semantic graphs to multimedia presentations. CWI Technical Report INS-R0305, 2003.

F. Giunchiglia, P. Shvaiko, and M. Yatskevich. S-match: An algorithm and an implementation of semantic matching. In *European Semantic Web Symposium*, pages 61–75, 2004.

Christian Halaschek-Wiener, Andrew Schain, Michael Grove, Bijan Parsia, and Jim Hendler. Management of digital images on the semantic web. In *Proceedings of the International Semantic Web Conference ISWC05*, Galloway, Ireland, 2005.

Susumu Harada, Mor Naaman, Yee Jiun Song, QianYing Wang, and Andreas Paepcke. Lost in memories: interacting with photo collections on pdas. In *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pages 325–333, New York, NY, USA, 2004. ACM Press.

S. Harris and N. Gibbins. 3store: Efficient bulk RDF storage. In *PSSS'03*, pages 1–15, Sanibel, Fl, 2003.

Stephen Harris. SPARQL query processing with conventional relational database systems. In *WISE Workshops*, pages 235–244, 2005.

T. Heath, E. Motta, and M Dzbor. Context as a foundation for a Semantic Desktop. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

Reginald L. Hobbs. Using XML to support military decision-making. In Lauren Wood, editor, *XML Conference*, Philadelphia, PA, 2003. 7-12 December 2003.

Nancy A. Van House and Marc Davis. The social life of cameraphone images. In *Proceedings of the Pervasive Image Capture and Sharing: New Social Practices and Implications for Technology Workshop (PICS 2005) at the Seventh International Conference on Ubiquitous Computing (UbiComp 2005)*. UbiComp 2005, 2005.

D. Huynh, D. Karger, and D. Quan. Haystack: A platform for creating, organising, and visualising information using rdf. In *Semantic Web Workshop, The Eleventh World Wide Web Conference 2002.*, 2002.

David Huynh, Stefano Mazzocchi, and David R. Karger. Piggy bank: Experience the semantic web inside your web browser. In Yolanda Gil, Enrico Motta, V. Richard Benjamins, and Mark A. Musen, editors, *International Semantic Web Conference*, volume 3729 of *Lecture Notes in Computer Science*, pages 413–430. Springer, 2005.

T. Iofciu, C. Kohlshutter, W. Nejdl, and R. Paiu. Keywords and rdf fragments: Integrating metadat and destktop search in beagle++. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

M O Jewell, Faith Lawrence, M M Tuffield, M S Nixon, A Prügel-Bennett, N R Shadbolt, D E Millard, and m c schraefel. OntoMedia: An ontology for the representation of heterogeneous media. In *Proceedings of MultiMedia Information Retrieval (MMIR) Workshop at SIGIR*. SIGIR, August 2005.

D. Joyce, P. lewis, R. Tansley, M. Dobie, and W. Hall. Semiotics and agents for integrating and navigating through multimedia representations of concepts. *Proceedings of Storage and Retrieval for Media Databases*, (3972):120–131, 2000.

Ryan Lee. Scalability report on triple store applications. Technical report, MIT, 2004.

A. Loizou and S. Dasmahapatra. Recommender systems and ontology alignment in the semantic web. In *ECAI2006 Workshop on Recommender Systems*, Riva del Garda, Italy, 2006.

m. c. schraefel, M. Karam, and S. Zhao. mSpace: interaction design for user-determined, adaptable domain exploration in hypermedia. In P. De Bra, editor, *Proceedings of AH 2003: Workshop on Adaptive Hypermedia and Adaptive Web Based Systems.*, pages 217–235, Nottingham, UK, 2003a. Springer.

m. c. schraefel, M. Karam, and S. Zhao. mSpace: interaction design for user-determined, adaptable domain exploration in hypermedia. In P. De Bra, editor, *Proceedings of AH 2003: Workshop on Adaptive Hypermedia and Adaptive Web Based Systems.*, pages 217–235, Nottingham, UK, 2003b. Springer.

F. Manola and E. Miller. RDF primer: W3c recommendation. http://www.w3.org/TR/rdf-primer/, 2004.

Cameron Marlow, Mor Naaman, Danah Boyd, and Marc Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *HYPERTEXT '06: Proceedings of the seventeenth conference on Hypertext and hypermedia*, pages 31–40, New York, NY, USA, 2006. ACM Press. ISBN 1-59593-417-0.

G. G. Mateos and C. V. Chicote. A unied approach to face detection, segmentation and location using hit maps. In *In Syposium Nacional de Re-conocimiento de Formas y Anlisis de Imgenes*, Benicasim, Castelln, 2001.

B. McBride. Jena: Implementing the RDF model and syntax specification. In *Proceedings of the 2nd International Workshop on the Semantic Web, SemWeb*, 2001.

K. Molle and S. Decker. Harvesting dekstop data for semantic blogging. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

P. Mulholland, T. Collins, and Z Zdrahal. Story fountain: intelligent support for story research and exploration. In *In Proceedings of 9th international conference on Intelligence User Interfaces*, pages 62–69. International Conference on Intelligent User Interfaces, 2004.

Mor Naaman, Susumu Harada, QianYing Wang, and Andreas Paepcke. Adventures in space and time: Browsing personal collections of geo-referenced digital photographs. Technical report, Stanford University, April 2004. Available at http://dbpubs.stanford.edu/pub/2004-26.

J. Richter, M. Volkel, and H. Haller. Deepamenta - a semantic desktop. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

D. Robertson. A lightweight coordination calculus for agent systems. *Lecture Notes in Computer Science*, 3476:183–197, 2005.

Carsten Rother, Sanjiv Kumar, Vladimir Kolmogorov, and Andrew Blake. Digital tapestry. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1*, pages 589–596, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2372-2.

Lloyd Rutledge, Jacco van Ossenbruggen, and Lynda Hardman. Making RDF presentable: integrated global and local semantic web browsing. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 199–206, New York, NY, USA, 2005. ACM Press.

L. Sauermann, A. Bernandi, and A. Dengel. Overview and outlook on the semantic desktop. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

N. Shadbolt and K. O'Hara. AKTuality: An overview of the aims, ambitions and assumptions of the advanced knowledge technologies interdisciplinary research collaboration. *AKT Selected Papers 03*, pages 1–11, 2003.

Nigel R. Shadbolt, Wendy Hall, and Tim Berners-Lee. The Semantic Web: Revisted. *IEE-Intelligent System*, 21(3):96–101, May 2006.

A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380, 2000.

Gabriella Spinelli. A multimedia system to collect, manage and search in narrative productions. In Hans-Jörg Bullinger and Jürgen Ziegler, editors, *HCI (1)*, pages 451–455. Lawrence Erlbaum, 1999. ISBN 0-8058-3391-9.

Kilian Stoffel, Merwyn G. Taylor, and James A. Hendler. Efficient management of very large ontologies. In *AAAI/IAAI*, pages 442–447, 1997.

Michael Streatfield. Report on Summer Internship Work For the AKT Project: Benchmarking RDF Triplestores. Technical report, University of Southampton, 2005.

Mischa M. Tuffield, Stephen Harris, David P. Dupplaw, Ajay Chakravarthy, Christopher Brewster, Nicholas Gibbins, Kieron O'Hara, Fabio Ciravegna, Derek Sleeman, Nigel R. Shadbolt, , and Yorick Wilks. Image annotation with photocopain. In *Proceedings of the First International Workshop on Semantic Web Annotations for Multimedia (SWAMM)*, Edinburgh, May 2006a. held as part of 15th World Wide Web Conference (22-26 May, 2006).

Mischa M Tuffield, Antonis Loizou, David Dupplaw, Sri Dasmahapatra, Paul H Lewis, David E Millard, and Nigel R Shadbolt. Semantic logger: Supporting service building from personal context. In *Proceedings of Capture, Archival and Retrieval of Personal Experiences (CARPE) Workshop at ACM MM*. ACM MultiMedia, October 2006b.

Jan Uhlír and Martin Falc. Annotating narratives using ontologies and conceptual graphs. In *DEXA Workshops*, pages 84–88. IEEE Computer Society, 2004. ISBN 0-7695-2195-9.

A. Vailaya, A. Jain, and H. Zhang. On image classication: City images vs. landscapes. In *Pattern Recognition.*, page 19211935, 1998.

R. Veltkamp and M. Tanase. Content-based image retrieval systems: A survey, 2000.

Willem A. Wagenaar. My memory: a study of autobiographical memory over six years, 1986.

James Ze Wang, Jia Li, and Gio Wiederhold. SIMPLIcity: Semantics-sensitive integrated matching for picture LIbraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9):947–963, 2001.

Edgar Weippl, Alexander Schatten, Shuaib Karim, and A. Min Tjoa. SemanticLIFE collaboration: Security requirements and solutions - security aspects of semantic knowledge management. In *PAKM*, pages 365–377, 2004.

World Wide Web Consortium. SPARQL query language for RDF, working draft. Technical report, World Wide Web Consortium, 2005.

H. Xiao and I. F. Cruz. A multi-ontology approach for perssonal information management. In *Proc. of Semantic Desktop Workshop at the ISWC*, 2005.

Ka Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted metadata for image search and browsing. In *Proceedings of the conference on Human factors in computing systems*, pages 401–408. ACM Press, 2003. ISBN 1-58113-630-7.

Gian Piero Zarri. A knowledge engineering approach to deal with 'narrative' multimedia documents. In Dines Bjørner, Manfred Broy, and Alexandre V. Zamulin, editors, *Ershov Memorial Conference*, volume 2244 of *Lecture Notes in Computer Science*, pages 363–377. Springer, 2001. ISBN 3-540-43075-X.