# Unintrusive, Engaging, Simple Semantic Mapping over Time

A thesis submitted to the

**University of Dublin, Trinity College**

for the degree of

Doctor of Philosophy

Colm Conroy

Knowledge and Data Engineering Group (KDEG),

Department of Computer Science,

Trinity College, Dublin

Supervisor: Dr. Declan O'Sullivan

Co-Supervisor: Dr. David Lewis

Assistant Supervisor: Dr. Rob Brennan

2012

# Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Colm Conroy

April 2012

# Acknowledgements

Firstly, I would like to thank my supervisor Dr. Declan O'Sullivan for agreeing to supervise my research, for his guidance, support and insightful contributions to this research. I would also like to express my gratitude to Dr. Rob Brennan and Dr. Dave Lewis for their support and helpful contribution through the years. Additionally I would like to express my gratefulness to my colleagues in KDEG who have always been generous with their time. Furthermore I would like to acknowledge all the volunteers who have given their time to participate in the experiments of this thesis.

Finally I would like to express my gratitude to my family and friends for their encouragement, support and patience throughout this entire process.

# Abstract

Semantic models are used to annotate, define and classify data. The promise of these semantic models is in the sharing of an understanding of a domain that can be communicated between people and applications. However different semantic models can be used to express the same data due to human diversity and the differing of needs. This divergence becomes problematic when there is a need for interoperability. Semantic mapping is a solution to this interoperability problem. However generating mappings is a difficult process and requires input from an expert known as a knowledge engineer. Additionally these knowledge engineers may not have the sufficient domain knowledge for the task. Instead it would be more beneficial to have a domain expert complete the task. Unfortunately the task is too complex for these people, denoted as ordinary users, and can even be trying for knowledge engineers. Furthermore the mapping process is traditionally done over one or a couple of long sessions making the task more labour intensive. This thesis determines the parts of the semantic mapping process which are difficult for ordinary users, identifies techniques to reduce the difficulty of these parts, and develops a mapping process that will aid these users in the act of mapping over time. The thesis proposes an approach to allow semantic mapping to occur over time and become accessible, convenient, simple, effective, efficient and more satisfactory for ordinary users. The approach was implemented in a mapping tool called the tag-based mapping tool. An evaluation approach was devised and used to evaluate the tag-based mapping tool in a real world case study with ordinary users over time within their computing environment.

# Table of Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Motivation

Semantic mapping[1] [Kalfoglou 2003] is a solution to the semantic interoperability problem[2] [Heflin 2000], where information needs to be exchanged between two sources with different underlying semantic models[3]. Semantic mappings allow data to be consistently interpreted with different semantic models. Fully automatic mapping has been considered impracticable as indicated in [Noy 2004] which has led to the requirement for user input in the mapping process and has been asserted in [Falconer 2009] [Shvaiko 2008]. Traditionally, knowledge engineers have predominately been assumed to be responsible for bridging the semantic models by creating mappings. However, there has been a substantial increase in the number of semantic models published on the web of Linked Data[4]. Furthermore as the number of Internet users and the number of accessible Web pages grows, it is becoming increasingly more difficult for users to find information that are relevant to their particular needs. This has led to the requirement of personalised systems [Gauch 2003] with personal semantic models to support the user's needs, which has been stated in [Huhns 2002]. For a truly scalable solution to wide-scale semantic interoperability, it is argued in this thesis that knowledge engineers alone will be insufficient as the volume of mapping tasks will be too much for such a small cohort of skilled people alone. In addition, for data integration to occur between two companies a mapping will need to be developed and this task is typically carried out by a knowledge engineer or person who has received substantial training. Unfortunately these users generally have insufficient domain knowledge which is crucial for the task and will require the involvement of a domain expert from the company for each domain being mapped. This leads to one of the most time consuming data management problems, asserted in [Falconer 2009]. However this problem can be simplified if the domain experts themselves are empowered to carry out the task for their own domain. The motivation for this thesis was thus to identify a way for people who use mappings, to participate themselves in the creation of mappings between semantic models that are relevant to them. This creates a challenge because the people who use mappings may only

[1] Semantic mapping is the process of finding correspondences between semantic models, see section 2.1
[2] Semantic Interoperability requires that any two computer systems will derive the same inferences from the same information.
[3] A Semantic model is a conceptual data model used to describe the meaning of data
[4] As of September 2010, the web of Linked Data contains more than 200 interconnected data sources totalling in over 25 billion RDF triples, see http://lod-cloud.net

have basic computing skills and thus will not have the technical skills that are assumed in the majority of current mapping tools.

Current semantic mapping tools typically require a high level of technical expertise which limits the users of the tools to knowledge engineers or users who have received substantial training, stated in [Falconer 2009]. Recently, the need for more general user support has been highlighted as a key challenge in semantic mapping [Shvaiko 2008]. Nonetheless little research has been undertaken on enriching user support, asserted in [Jameson 2006]. In fact, any published ontology mapping research on enriching user support has been aimed at improving the experience for knowledge engineers with a high level of technical expertise, an example is [Falconer 2007]. To allow people who use mappings themselves to interact in the semantic mapping process, it is argued in this thesis that the design foci for semantic mapping tools needs to shift to people who only have basic computing experience, denoted as **ordinary users** in this thesis.

It is also argued in this thesis that semantic mapping is not a “one-shot” process and that mappings need to be developed and refined over time which would require regular involvement of users in the process. In the author’s opinion, the mapping process needs to be modified to allow the act of mapping to occur over time within the users own computing environment. A motivation for such a long term process and continuous user involvement is the use case of "personal information delivery". Personal information delivery involves the push of content from information streams on the web to users based on their interests. Mappings are used to personalise the content pushed to the user. For example if a person named John had an interest in football, modelled within his personal semantic model, then mappings would be constructed to semantic models representing sporting information streams, e.g. the BBC sport news RSS feed[5]. The resulting mappings would be used to personalise the sporting information from the streams by only delivering sporting news relating to football to John. The traditional mapping approach where knowledge engineers alone are involved in the mapping process is thus impracticable in this use case due to the volume of mapping tasks needed[6]. In addition, the user’s interests may change over time which would require the mapping to be continually refined.

A further scenario for such a process is biomedical science. The number of biomedical resources available to researchers is enormous. This causes a problem when the medical

[5] http://newsrss.bbc.co.uk/rss/sportonline_world_edition/front_page/rss.xml
[6] Over 3.4 million unique people visited the BBC sport news page during the world cup in 2006 which could lead to over 3.4 million mappings being required, see http://news.bbc.co.uk/2/hi/technology/5165574.stm, retrieved September 2011

literature is searched due to the large volume of documents retrieved. Using the system described would allow users to partially map documents together 'on the fly' as needed rather than the time-consuming method of mapping the entire documents together. Furthermore the process allows the mapping to be continually refined as the documents are modified.

As stated before, the majority of current mapping tools typically require a high level of technical expertise which limits the users of the tools to knowledge engineers or users who have received substantial training. Disruptive innovation allows services to be opened up to non-consumers by reducing the obstacles attached to the service [Christensen 1998]. In particular, often disruptive innovation makes the product much simpler, more convenient and more accessible as asserted by Anthony [Anthony 2008]. For examples, the Nintendo Wii allowed a new generation of gamers to be born through intuitive controls making playing games simpler and more accessible, whilst Google Ads made advertising on the web accessible to anyone[7]. In the author's opinion, to allow ordinary users to participate in the semantic mapping process the service itself, i.e. the act of mapping, must follow this approach and become **much simpler**, **more convenient** and **more accessible**.

The semantic mapping process can be extremely difficult, requiring tremendous patience and an expert understanding of the domain, terminology, and semantics as stated by Falconer [Falconer 2007]. The current state of the art mapping tools have interfaces that quickly become unmanageable due to the clutter of irrelevant and relevant information, even for knowledge engineers as asserted in [Robertson 2005] [Falconer 2009]. To **simplify** the semantic mapping process the time spent and the amount of effort the user expends will need to be dramatically reduced. The interfaces of semantic mapping tools will **also** need to be enhanced to allow ordinary users to use them. In particular, the **usability needs to be significantly improved**. Usability is:

> *The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.* (ISO 9241-11)

In the definition effectiveness means that the goal was accomplished by the user adequately, efficiency means that the goal was accomplished by the user with a minimum expenditure of time and effort, and satisfaction discovers what people think and feel about using a product, to assess the perceived quality of use.

[7] Scott Anthony talks about disruptive innovation, http://www.youtube.com/watch?v=9L66OH-x7a4, retrieved 16 March 2011

Current mapping tools are standalone applications and require users to open the tool and develop mappings whenever they are needed. Such a process will not be **accessible** or **convenient** for ordinary users as it requires users themselves to know when mappings are needed and to enter the tool of their own accord, which may not happen. Instead the mapping process **needs to be seamlessly integrated into the ordinary users computing environment** so that mapping tasks are **highlighted and displayed** to users **when required** as they occur naturally over time. This gives rise to the requirement that both the process and tool be *non-intrusive* and not disrupt users in their daily work. It also requires users to be constantly involved or *engaged* in the process over time.

In summary, contrary to the existing research trend which presupposes only knowledge engineers will be involved in the semantic mapping process, it is argued that the users of mappings need to be actively included in the semantic mapping process. This realisation comes from the need that in its current form, the semantic mapping process, which relies on the availability of knowledge engineers, will not scale to the World Wide Web as the propagation of semantic models continues to increase. This proliferation will lead to an increase in the volume of mappings needed. However the current mapping processes and state of the art mappings tools are extremely complex and can even be unmanageable for knowledge engineers. This leads to the question of how can the mapping process and tools be made accessible, convenient and usable for ordinary users?

## 1.2 Research Question

The research question posed in this thesis is:

*How can the act of semantic mapping become accessible, convenient, simple, effective, efficient and more satisfactory for ordinary users to achieve semantic mappings gradually and over time between semantic models of interest to the user*?

The act of semantic mapping is defined to be the generation of correspondences between portions of two different semantic models [Kalfoglou 2003]. Ordinary users in this thesis denote users who have only basic computing experience. Accessible is a general term used to describe the degree to which a product, device, service, or environment is available to as many people as possible. In the research question, accessible requires that the goal of constructing mappings is available to ordinary users over time. Convenient is a term used to describe a product or service being suited to one's purpose or needs. In the research question, convenient requires that the mapping process can be integrated into the ordinary user's computer environment. Simple is a term used to describe a task as not being complicated or easy to achieve. In the research question, simple requires that the goal of generating mapping will be easy to accomplish for ordinary users. Effective is a term used to explain that a purpose was accomplished adequately or produced the intended result. In the research question, effective requires that the goal of constructing mappings was accomplished by ordinary users sufficiently. Efficient is a term used to describe that a task was performed in the best possible manner with the least waste of time and effort. In this research question, efficient requires that generation of mappings was accomplished by ordinary users with a minimum expenditure of time and effort. Satisfactory is a term used to explain that the task was suitable or acceptable. In the research question, satisfactory requires that ordinary users perceived the task of constructing mapping to be acceptable.

The semantic models used in the thesis are OWL DL ontologies[8]. OWL ontologies were chosen as they are more complex than other types of semantic models and thus may be harder for ordinary users to understand which would make the act of mapping more complex for ordinary users. Examples of other semantic models are folksonomies [Vander der Wal 2007] and linked data [Bizer 2009]. In addition, many of the problems inherent in the ontology domain are consistent across other semantic model domains, see section 2.1.1.2. This allows for the approach proposed in this thesis to be applicable to mapping with other semantic model formats. The mapping use-case studied is restricted to one-to-one concept mappings as it is the most common use case used currently in the semantic mapping research community, see section 2.1.7.

## 1.3 Objectives and Goals

In order to investigate the research question the following objectives were derived:

> **RO1**: Determine whether ordinary users can actively participate in the construction of semantic mappings to a similar standard as knowledge engineers in areas of interest to the ordinary users;
>
> **RO2**: Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts;
>
> **RO3**: Develop a mapping process that will aid ordinary users in the act of mapping over time, called the incremental mapping process;

[8] http://www.w3.org/TR/owl-guide/

**RO4**: Develop a tool that is based on the incremental mapping process from **RO3** and the techniques identified in **RO2**;

**RO5**: Evaluate the tool with ordinary users over a long term period (at least 2 weeks and up to 5 weeks).

Although two to five weeks maybe perceived as a short term period, in the author's opinion, it will give a snapshot of the long term behaviour of the user with the developed mapping tool. Furthermore there is no formal definition in mapping for what a long term period is and there has been no user study conducted on mapping tools over time to give a basis on what to judge a long term period as.

## 1.4 Contribution

The major contribution of this thesis is the development of an approach which allows the act of mapping to become accessible, convenient, simple, effective, efficient and satisfactory for ordinary users. As mentioned in section 1.1 this type of approach is lacking in the state of the art. This approach includes discovering the mapping requirements needed for ordinary users to participate in the mapping process and identifying both the techniques and methods to solve these requirements. Such identification has not been carried out by the state of the art. Additionally an incremental mapping process is designed to allow the act of mapping to occur over time and an implementation of the process is used in the approach to reduce the workload of the mapping process. A mapping process tailored for ordinary users is lacking in the state of the art. The approach is implemented in a mapping tool called the tag-based mapping tool which is developed using the incremental mapping process and both the techniques and methods identified to solve the requirements. The tag-based mapping tool was evaluated and the results indicated that the approach allows the act of mapping to become accessible, convenient, simple, effective, efficient and satisfactory for ordinary users. The approach and the ideas for the approach have been published at [Conroy 2007] [Conroy 2008a] [Conroy 2008b] [Conroy 2008c] [Lanzenberger 2008] [Conroy 2009].

In addition there are two minor contributions of this thesis to the state of the art. The first minor contribution of the thesis is the design of an incremental mapping process itself. In the author's opinion, the current state of the art mapping tools rely on "one-shot" session for users to develop mappings which tend to be time-consuming. They lack the ability to intertwine the mapping process with the user's own work environment over time. The incremental mapping process, first presented at [Conroy 2008a], is designed to allow the mapping process to occur incrementally over time within the user's computing environment. This process is independent from the mapping approach taken in this thesis and can be used for different mapping approaches, e.g. supporting the participation of knowledge engineers over time. In addition the process is designed to allow the number of user-specific mapping tasks to be reduced. Furthermore it also is designed to support users during these user-specific mapping tasks and identifies key points of user interaction during the task. Such clear identification, separation and support of user interactions are currently lacking in the state of the art and are needed if the usability of mapping tools is to be improved. Finally the process also identifies and highlights where personalisation can occur in the mapping process.

The second minor contribution of this thesis is the formulation of an approach to evaluate the user's participation with mapping tools within their own computing environment over time. This evaluation approach is used to analyse the behaviour and experiences of users using the tag-based mapping tool and can also be used on other mapping tools. Such an evaluation approach is currently lacking in the state of the art. The current state of the art research have conducted few user studies on mapping tools with [Lambrix 2003] and [Falconer 2009] being a couple of examples of user studies carried out. Furthermore, the focus of evaluation in these studies has been on the mappings made by the tools using the standard precision and recall measures rather than assessing the users' experience. The methodologies presented in this thesis, presented at [Conroy 2009], are centred on evaluating whether the act of mapping with a mapping tool is accessible, convenient, simple, effective, efficient and satisfactory for users. In addition the few user studies carried out by the state of the art have primarily been done in laboratory settings. However there is a need to evaluate mapping tools in real world case studies to analyse the user behaviour with the tool within their own computing environment. The evaluation approach included the design and implementation of a personalisation application, which uses the tag-based mapping tool, to replicate a real world situation. This allowed the tag-based mapping tool to be evaluated in a real world situation with the users.

## 1.5 Technical Overview

This section presents the approach taken in the investigation of this thesis and outlines how the rest of the thesis is structured. First the technical approach of the thesis is presented. Secondly the overview of how the experiments are conducted is described. Thirdly the analysis methods used to evaluate the results of the experiments is detailed. Finally the structure of the thesis is outlined.

### 1.5.1 Technical Approach

An initial study on the state of the art in the area of semantic mapping was conducted, which concentrated on the role of the user within the mapping process, in particular focussing on both the user interaction and visualisation of information. This literature review was supplemented with an initial user mapping experiment, due to the lack of user mapping studies carried out in the state of the art, which was designed to identify the practical user interaction issues. This involved contrasting the issues that two different mapping systems had on three different groups of users, each group having different levels of technology expertise reflecting the different types people can have. The initial experiment and literature investigation helped in the formation of an understanding of the issues involved in the user's role in the semantic mapping process. Based on the state of the art analysis and experience gained through the initial user mapping experiment, an incremental mapping process and software mapping tool were designed. The mapping tool developed was called the tag-based mapping tool. The incremental mapping process was designed to focus on reducing the work load and difficulty in the act of mapping for users.

Two further sequential experiments were used to evaluate various aspects of the incremental mapping process and tag-based mapping tool. Examples include how the ontological information should be displayed and when mapping tasks can be prompted. For these experiments the tag-based mapping tool was designed to be embedded within the user's own browser environment. The first experiment evaluated using a time-interval approach to display mapping tasks. The second experiment evaluated the impact of using a context-sensitive approach to display mapping tasks as well as several other changes made to the tool as a result of the evaluation of the previous experiment, for example only asking mapping tasks relating to the users interests. The modification to a context-sensitive approach was made to analyse if there was any change in the participation of ordinary users in the mapping process from using a time-interval approach.

### 1.5.2 Evaluation Overview

To address the research question, the experiments carried out in this thesis needed to evaluate the accessibility, convenience, simplicity, effectiveness, efficiency and user satisfaction of the tag-based mapping tool with ordinary users. Evaluating the tag-based mapping tool will assist in evaluating the incremental mapping process devised. The tag-based mapping tool was evaluated by analysing the user's performance with the tool and the user's feedback. Commonly in usability evaluation the efficiency is measured by the time taken to complete tasks with the tool while the effectiveness is measured by the accuracy of the tasks with the tool. The simplicity is measured by how easy it is for users to complete the act of mapping. For the act of mapping to be both accessible and convenient it needs: the process to be non-invasive to the user as they are carrying out their daily work as the work may be voluntary, making it important not to disrupt the user, and engaging enough so that the user will want to interact in the mapping process. Thus the evaluation categories used to measure accessible and convenience are *unintrusive* and *engaging*. The user satisfaction is evaluated through questionnaires and interviews with the users. The questions asked also supplement the evaluation of the unintrusive, engaging and simple categories. Table 1-1 lists how each experiment is evaluated.

**Table 1-1:** Thesis evaluation outline

| **Evaluation Category** | | **Experiment 1** | **Experiment 2** | **Experiment 3** |
|---|---|---|---|---|
| **Efficiency** | | Yes | Yes | Yes |
| **Effectiveness** | | Yes | Yes | Yes |
| **Accessibility & Convenience** | ***Unintrusive*** | No | Yes | Yes |
| | ***Engaging*** | No | Yes | Yes |
| **Simple** | | No | Yes | Yes |
| **User Satisfaction** | | Yes | Yes | Yes |

### 1.5.3 Analysis Approach

Both quantitative and qualitative analysis was undertaken in each experiment. The following statistical tests were chosen after consultation with a statistician:

- *Mann-Whitney U test* [Crichton 2000]: This is a non-parametric test used to assess whether two independent samples of observations have equally large values. This test was used to measure the responses of the same questionnaire from different experiments. This test is used in experiment three to evaluate if the questionnaire responses of users were different to the responses given by the users in experiment two.
- *Unpaired two-tailed t-test* [O'Connor 2003] [Ruxton 2006]: This is a two sample location test of the null hypothesis that the means of normally distributed populations are equal. The F-test was used to determine if variance in the data sets were equal or unequal. Based on the F-test result either an unpaired two-tailed t-Test with equal variance or unequal variance (Welch t-test [Welch 1947]) would be used. This test was

used to measure the differences in mapping results, e.g. mapping precision, between users with different technology experience. This test is used in each of the experiments to compare the mapping performances between the different user groups.

- *One-sample t-test* [O'Connor 2003]: This is a one-sample location test of whether the mean of a normally distributed population has specific value to the null hypothesis. This test was used to measure the mapping results against a standard. This test is used in experiment three to evaluate if the mapping performance of users was different to the standard set by the users in experiment two.
- *Paired two-tailed t-test* [O'Connor 2003]: This is a test of the null hypothesis that the difference between two responses measured on the same statistical unit has a mean value of zero. This test was used to evaluate whether the mapping results for users improved over time. This test is used in experiment two and three to evaluate if the mapping performance of users changed over time.
- *Slope t-test* [O'Connor 2003]: This test is used to evaluate whether the slope of a regression line differs significantly from 0. This test was again used to evaluate whether the mapping results improved over time. This test is used in experiment three to evaluate if the mapping performance of users changed over time.

For the tests the significant level was chosen as $\alpha = 0.05$, as it is conventionally used as the level of significance in statistical tests [Fisher 1925]. Also, any outliers from the data set that were more than two standard deviations away from the mean were eliminated to reduce the inflated error rates [Osborne 2004]. In particular, triangulation [Creswell 2003] is used to validate emergent themes, i.e. conclusions from the evaluations. This involves verifying that the themes are present in the multiple experiments carried out. This provides justification and evidence that the theme is a consistent usage pattern across a population of users. The analysis of the experiments supported the claim ordinary users can interact in the mapping process using the tag-based mapping tool.

### 1.5.4 Thesis Overview

Chapter Two provides a review and analysis of the state of the art in ontology mapping, with particular focus on the role of the user.

Chapter Three presents the first of our experiments which focussed on comparing two different approaches to semantic mapping on three groups of users with different expertise. The conclusions from this experiment as well as the review of the state of art helped in the design of our incremental mapping process.

Chapter Four presents the design of the incremental mapping process. A summary of the process is provided including detail description of each key interaction step within the process.

Chapter Five describes the implementation of the tag-based mapping tool based on the incremental mapping tool developed and a personal information delivery tool, SportsFlows, which uses the tag-based mapping tool to personalise sporting information for users.

Chapter Six presents the second of our experiments. The experiment required users to use the personal information delivery tool, SportsFlows, presented in the previous chapter. The experiment focused on evaluating the accessibility, convenience and usability of the mapping tool within the browser environment of three different groups of users with different expertise.

Chapter Seven describes the improvements made to the tag-based mapping tool and the personal information tool SportsFlows based on the conclusions from the previous experiment.

Chapter Eight presents the final experiment. The experiment was a final evaluation on the accessibility, convenience and usability of the process and tool constructed in this thesis.

Chapter Nine gives an overview of the entire evaluation carried out in this thesis and describes if the objective of the research question in this thesis was achieved. In particular it describes how the objectives of each experiment were evaluated and discusses the conclusions from the experiments.

Chapter Ten discusses the contributions made, presents suggestions for future work, and concludes with some final remarks.

Appendix A contains the complete results for the statistical tests of the experiments.

Appendix B displays the interview questionnaire used in the first experiment, see section 3.4.

Appendix C gives the list of added labels for each concept in the sports news ontology used in the SportsFlows applications, see section 5.2.

Appendix D presents the questionnaires used in the second experiment, see section 6.4.

Appendix E contains the email used in recruitment of the participants for the third experiment, see section 8.4.

Appendix F gives the questionnaire used in the third experiment, see section 8.4.

Appendix G displays the table of contents for the accompanying DVD media.

# 2 State of the Art – Ontology Mapping

This chapter presents a summary of the state of the art in ontology mapping. In particular the emphasis is examining research carried out regarding the involvement by users in mapping activities. First an overview of the ontology mapping problem is described together with a motivating example and a discussion of the difficulties in mapping. Then an overview of ontology matching tools is presented focussing on the different types of algorithms that can be used. Next, the literature on mapping processes is discussed with an emphasis on evaluating where user input is required. Subsequently, current state of the art mapping tools are detailed highlighting the different types of user interfaces supported. Afterwards, recent research on adding user support to mapping tools is outlined. After that, a summary of the different types of interactions mapping tools can use is given. This is then followed by an overview of the techniques and approaches for user evaluation of mapping tools. Finally, requirements needed for ordinary users to use mapping process and tools are presented.

## 2.1 Overview of Ontology Mapping Problem

A mapping problem occurs when there are different representations of similar information. To allow for the information to be shared a mapping must be constructed in order to transform one representation into another. The mapping will allow information to be understood with each representation. Ontology mapping is one solution to this semantic heterogeneity problem [Shvaiko 2008]. Matching is the generation of a set of candidate correspondences between semantic models. These candidate correspondences need to be validated before they become correspondences. A mapping between ontologies consists of a set of correspondences between semantically related entities. Each correspondence is formally defined as a 5-tuple $<id, e_1, e_2, n, r>$ [Euzenat 2006]: where *id* is the unique identifier of the correspondence, $e_1$ is an entity in the source ontology, $e_2$ is an entity in the target ontology, *r* is the relation between entities, and *n* is the confidence measure that the correspondence holds for $e_1$ and $e_2$.and is generally a value between 0 and 1. The relationship is typically one of equivalence, more general, less general, disjointed, or overlapping [Shvaiko 2008]. However the exact relationship specified is often application dependent.

### 2.1.1 Motivating Example

Figure 2-1 shows a snapshot of a partial mapping between two ontologies. On the left is the source ontology which is a partial branch of the university ontology for the University of Aberdeen [Aberdeen 2003]. On the right is the target ontology which is a partial branch of the university ontology for the University of Manchester [Manchester 2003]. The ontologies are provided in the accompanying DVD media under SOA.

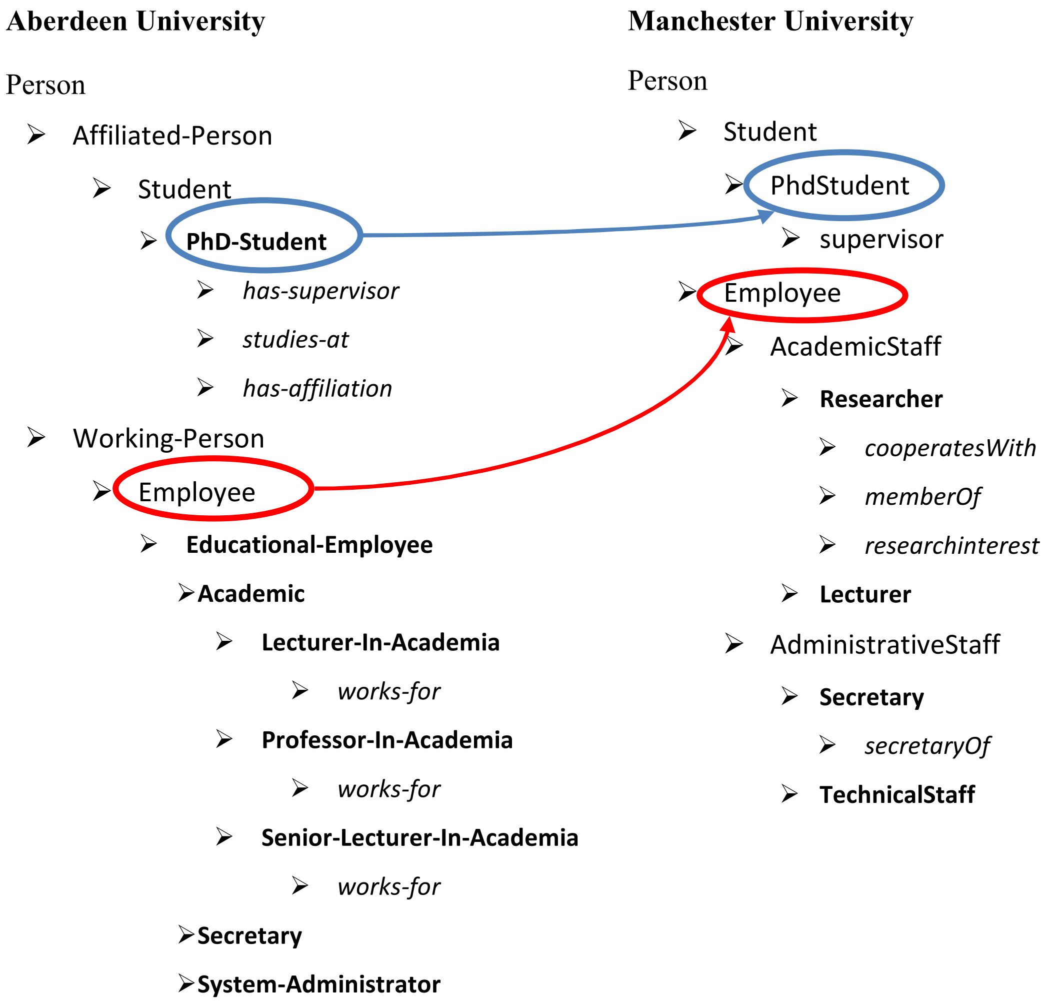



***Figure 2-1:*** *Example of partial portion of a mapping between Aberdeen University ontology and Manchester University ontology. Entities from both source and target ontology involved in the mapping are bounded by ellipses and correspondences are represented by the solid curved arcs.*

The ontologies are both modelled as a different structure and uses a different vocabulary to represent similar information. This difference can be due to both the differing needs of each ontology and the fact that different people, who would have differing views [Lewotin 1982], were used to build each ontology. An example of two entities which use different vocabulary but model similar information is 'System-Administrator' from the Aberdeen ontology and 'TechnicalStaff' from the Manchester ontology. The two ontologies also contain many of the same concepts but the concepts are sometimes represented differently, for example 'PhD Student' is represented differently in both ontologies. The Aberdeen ontology represents 'PhD Student' as a sub class of 'Student' with the properties *has-supervisor*, *studies-at* and *has-affiliation*. While the Manchester ontology represents 'PhD Student' as a sub class of 'Student' with the property *supervisor*. Both concepts have different properties and the different class based structure, 'Student' is a sub class of 'Affiliated-Person' in the Aberdeen ontology but is not in the Manchester ontology. This heterogeneity causes problems when information needs to be shared or exchanged. To resolve this semantic interoperability problem, a mapping needs to be constructed between the two ontologies. However, developing a mapping is a difficult process.

### 2.1.2 Difficulty in Mapping

The study of mapping problems is not new in computer science and has been persistent through various different areas from theoretical computing [Hertling 1999] to the database community [Konstantinou 2008]. In fact people deal with mapping problems everyday with examples being both reading and interpreting our surroundings. Ontology mapping is closely related to the mapping problems we encounter every day, i.e. mapping one person's view of the world to another person's view of the world. However, ontologies are quite limited when expressing representations as they are based on the classical view of categories [Murphy 2002]. This classical view has three main claims: first that concepts are mentally represented definitions which provide the necessary conditions for membership in this category, second that every object is either in or not in a category, and finally that each member of a category is equally good, that is, a member cannot be a more typical member than another member.

One of the main problems using this approach is that it is very difficult to define concepts through necessary and sufficient conditions. For example "dog" can be defined to be an animal that has four legs, barks, has fur, eats meat but this is not a valid definition as there are dogs with less than for legs and dogs that are furless [Murphy 2002]. Also studies have shown that people have difficulty assessing category membership by not being able to segregate items into clear members and non-members [Hampton 1979]. These issues can lead to similar problems in mapping where the same term is structured and verbalised differently. Also the underlying data language used for specifying the ontology (e.g. OWL, RDF) introduces added problems as they constrain the expressiveness. For example, many formats lack information relating to the context of use [Bernstein 2007]. Without knowing the context of use for a concept it can be difficult to understand what is the meaning of the concept. For example 'Football' can have different meanings based on the perception of different cultures, e.g. in Europe against in USA. This is especially problematic in mapping as the context of both terms needs to be understood.

Mismatches between ontologies are the key type of problem that hinders the combined use of independently developed ontologies [Klein 2001]. Figure 2-2 displays Klein's classification for ontology mismatches.

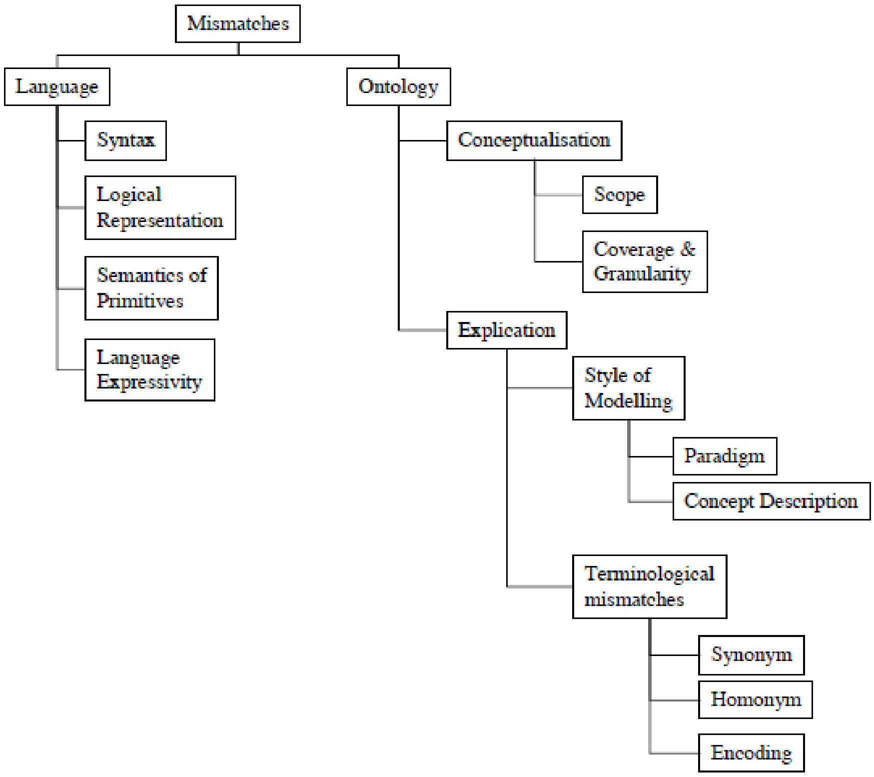



***Figure 2-2:*** *Classification of Ontology Mismatches [Klein 2001]*

The "language level" distinguishes mismatches of the language primitives that are used to specify the ontology. The "ontology level" distinguishes the differences in the way the representations of ontologies are modelled and is subdivided into conceptualisation and explication mismatches. "Conceptualisation Mismatches" relate to the difference in the way a representation is interpreted and is subdivided into Scope, two classes may seem to be the same but are modelled differently, and Coverage & Granularity, appear to be modelling the part of the same domain but at different levels and detail. "Explication Mismatches" represent a difference in the way the conceptualisation is specified, covering style of modelling, terminological and encoding mismatches. As recognised by Klein [Klein 2001], it relies on human expertise to recognise ontology mismatches.

All these problems make mapping generation a very challenging problem. Despite significant research on the development of matching algorithms, it has become obvious that the user must accept a degree of imperfection [Gal 2005]. A couple of prime reasons for this is the enormous ambiguity and heterogeneity of concept descriptions of data [Gal 2006] and the description of a concept in a schema can be semantically misleading [Miller 2000]. This has lead to automatic mapping being seen as impracticable [Noy 2004] and there remaining a need for "a human to be in the loop" [Hull 1997] [Shvaiko 2008] [Falconer 2009].

## 2.2 Ontology Matching Tools

This section presents a survey of candidate correspondence generation tools. There have been a variety of approaches used to automatically generate candidate correspondences between ontologies. For example, Shvaiko and Euzenat [Shvaiko 2004] discuss over 20 different matching algorithms and tools, see Figure 2-3.

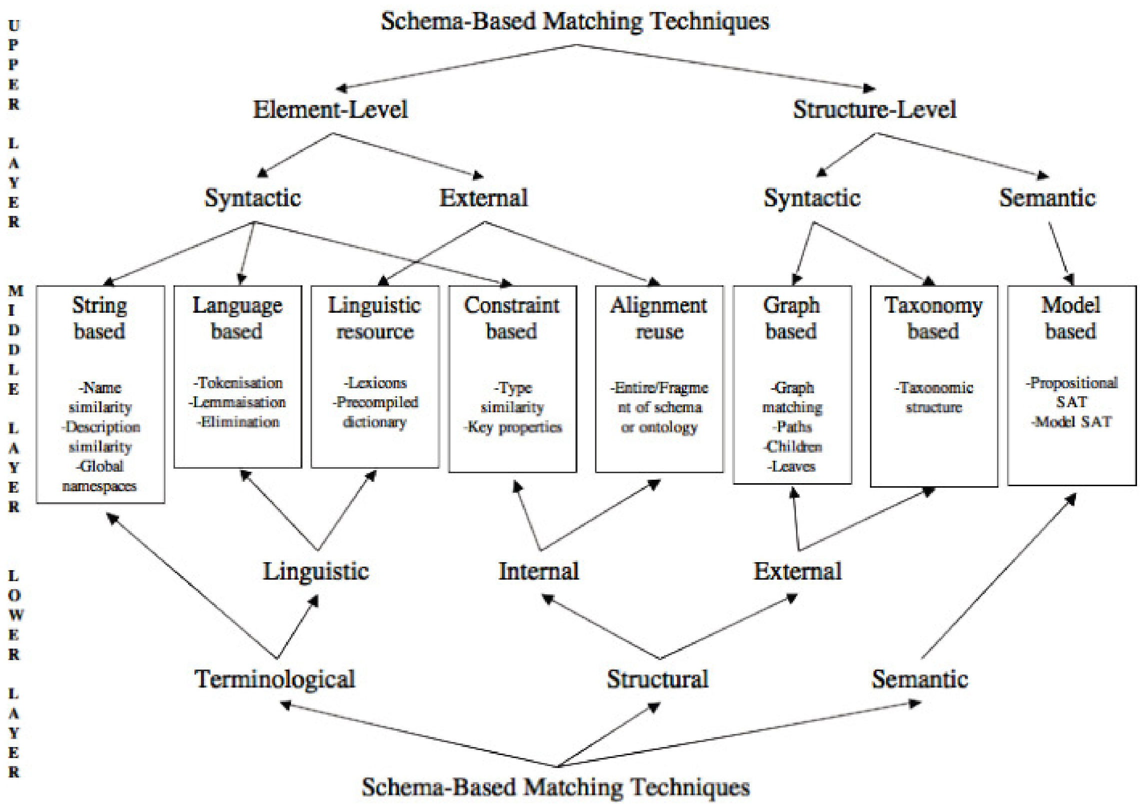



***Figure 2-3:*** *List of schema based matching approaches [Shvaiko 2004]*

One of the most widely used methods for computing similarities is heuristic techniques applied to the schema or ontological description. Heuristics are generally applied in two different ways. First, they are applied to the labels in the ontologies to compute lexical similarity, for example similar concept names, and second, they are applied to the structure of the ontology to measure structural similarity between terms, for example similar hierarchical relationships. The majority of mapping algorithms are based on heuristic techniques and will often apply a hybrid approach of lexical similarity and structure similarity.

Another, less widely used approach is the instance-based or instance-level approach [Doan 2003]. Here, the instances of concepts are compared rather than their representation. An instance is an actual value of a concept, for example, an instance of a concept “Footballer”, would be an actual footballer, such as Lionel Messi. Concept similarity can then be measured by comparing shared instances. Typically an instances-based approach will require the concepts to have overlapping instances.

An additional matching approach is based on mapping ontologies to a standard data dictionary such as WordNet[9] or UMLS (Unified Medical Language System)[10]. Using this technique, the data dictionary acts as a central ontology and ontologies will need to be mapped to this dictionary. The ontology can then be compared to the data dictionary and the most similar term in the data dictionary becomes the canonical representation of the ontology term. This approach has been used to extend ontology matching to the medical domain using UMLS [Shamdasani 2009].

A further matching approach is probabilistic mapping [Moulton 1993]. With this approach the concepts are mapped together with an associated probability and this probability is used to infer other correspondences between related concepts. An example of this approach is OMEN [Mitra 2005] that is a probabilistic ontology mapping tool which is dependent upon a set of meta-rules that express the influences of correspondences on the existence of other candidate correspondences across concepts in source ontologies that are located in the proximity of the correspondence.

There has been much research on examining and evaluating different types of matching algorithms across different domains [OAEI 2011]. Mapping tools generally included multiple matching algorithms to generate candidate correspondences.

## 2.3 Mapping Process

This section presents an examination of the mapping processes represented in current state of the art ontology mapping systems. Four ontology mapping systems were selected: the MAFRA system [Maedche 2002], the COMA system [Do 2002], PROMPT [Noy 2003] and CoGZ (Cognitive Support and Visualisation for Human-Guided Mapping Systems) [Falconer 2009]. The primary focus of this inspection was to identify the role of the user in the mapping process within each system. Each of the mapping systems is described in the sections below.

[9] http://wordnet.princeton.edu/
[10] http://www.nlm.nih.gov/research/umls/

**Table 2-1:** Overview of Selected Mapping Systems

| **Mapping System** | **Reason for Selection** |
|---|---|
| *The MAFRA system* | This system has influenced the design of other mapping frameworks, such as the ontology mediation management framework of the IST SEKT project [Bruijn 2005]. |
| *The COMA system* | Represents a typical database schema matching approach and has been extended to mapping ontologies with COMA++ [Aumuller 2005]. This system has features which allow parts of the framework to be adaptable |
| *PROMPT* | Is a plug-in mapping tool for the popular ontology editing tool Protégé [Noy 2000]. PROMPT is extensible and allows other mapping tools to be built on top of PROMPT |
| *CoGZ* | Is a mapping tool built on top of PROMPT. CoGZ is a recently developed mapping tool designed to improve the user support during the mapping process |

### 2.3.1 The MAFRA System

The main goal of the Ontology Mapping Framework Toolkit (MAFRA) system is to transform instances of the source ontology into instances of the target ontology. Semantic Bridges specify how to perform these transformations and are classified into concept bridges and property bridges. The MAFRA conceptual framework shown in Figure 2-4 consists of five horizontal and four vertical modules.

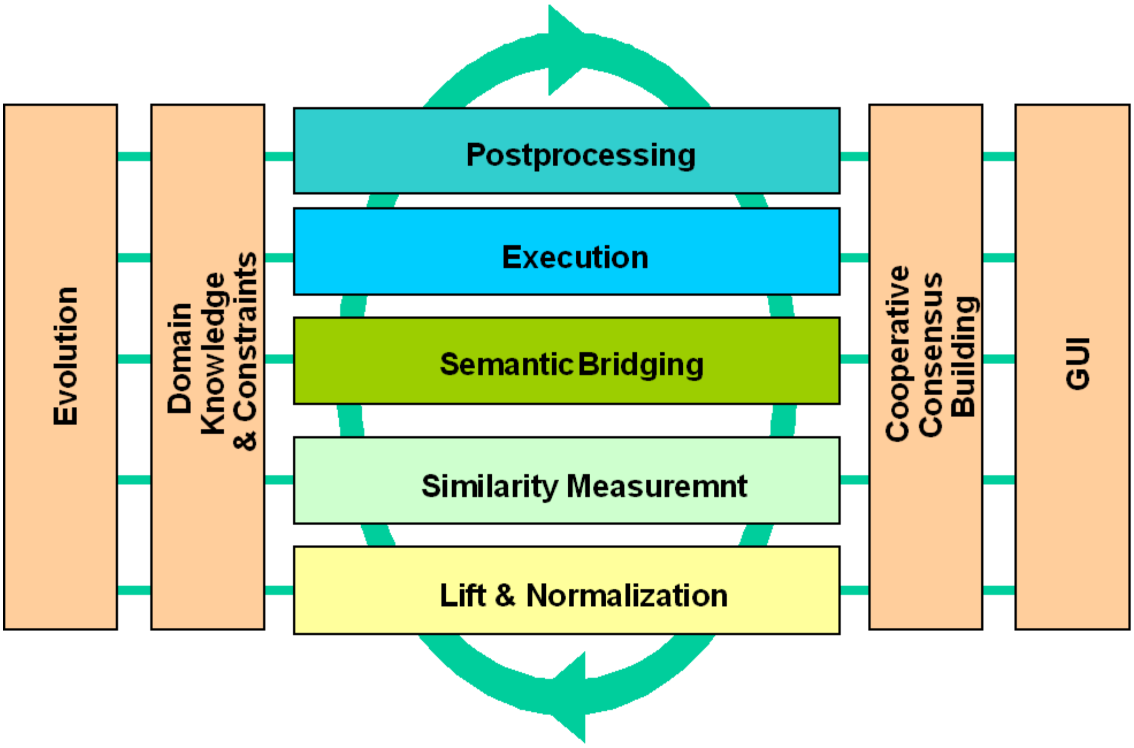



***Figure 2-4:*** *Overview of MAFRA framework [Maedche 2002]*

The horizontal modules correspond to five ontology mapping activities that have been defined by MAFRA, namely: Lift & Normalization, Similarity, Semantic Bridging, Execution and Post-processing. The first two modules are mapping system specific and do not involve user participation. The aim of these modules is to develop candidate correspondences which the user will validate in the Semantic Bridging module. As well as validating candidate correspondences the user can also generate any missed correspondences using the graphical user interface, see Figure 2-5. In the execution phase the actual instances of the source and target ontologies are transformed by evaluating the semantic bridges. In the final post-processing phase, the results of execution are examined by the system in order to check and improve the quality of the transformation.

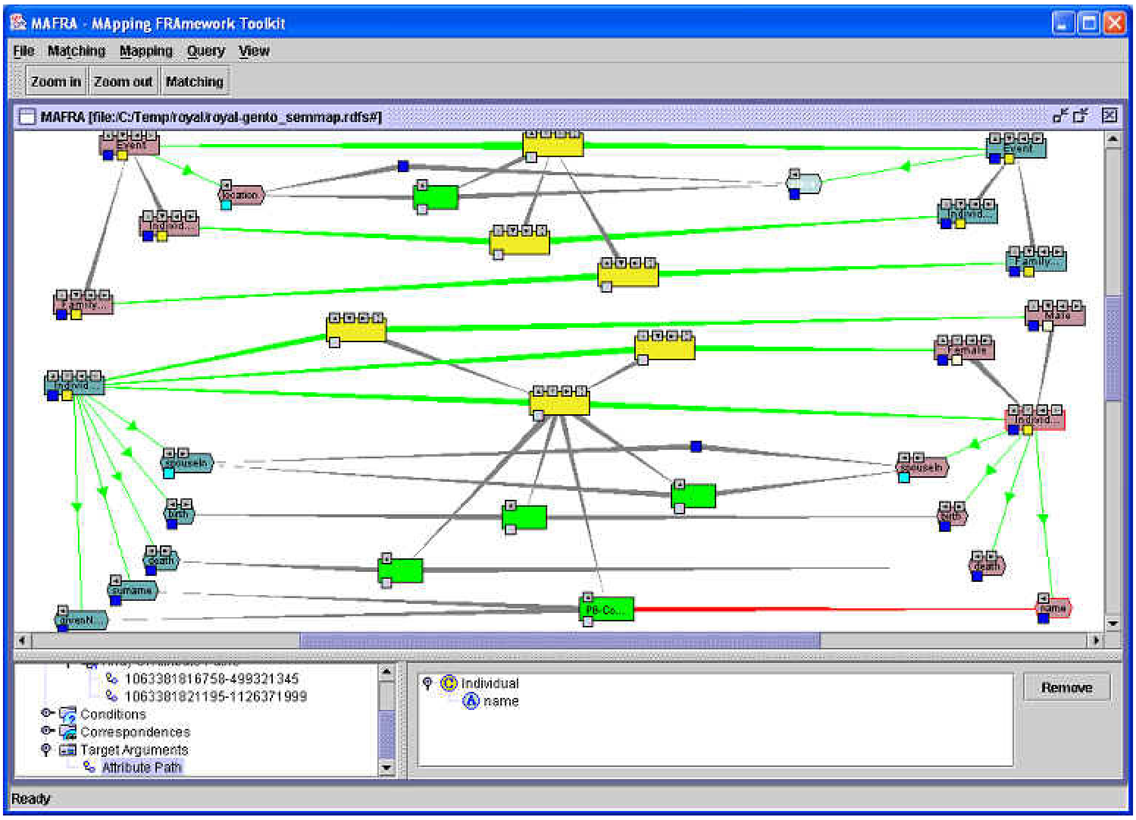



***Figure 2-5:*** *The graphical user interface of MAFRA [Maedche 2002]*

The role of the user in the MAFRA system is to develop a mapping by validating the candidate correspondences suggested and appending any missed correspondence. There is no default navigation between candidate correspondences and it is left to the user to navigate this list themselves. The candidate correspondences are generated by a custom made matching algorithm. The MAFRA system does not allow for modification or replacement of the matching algorithm

### 2.3.2 The COMA System

The original COMA system was designed to undertake schema mapping between elements of two schemas, supporting the mapping of relational database schemas and XML schemas [Do 2002]. COMA takes a composite match approach in that it combines the results of several independently executed match algorithms. The user selects which matching specific algorithms to use. COMA++ [Aumuller 2005] extends the original COMA system (for XML and database schema mapping) to include ontology mapping, as well as adding new tool and matching features. New tool features include a graphical user interface, a repository for ontologies and mappings, and a mechanism to compose, merge, and compare different mappings. These new features also include the addition of ontology matchers and a mechanism to reuse existing mapping results. The COMA++ architecture is shown in Figure 2-6.

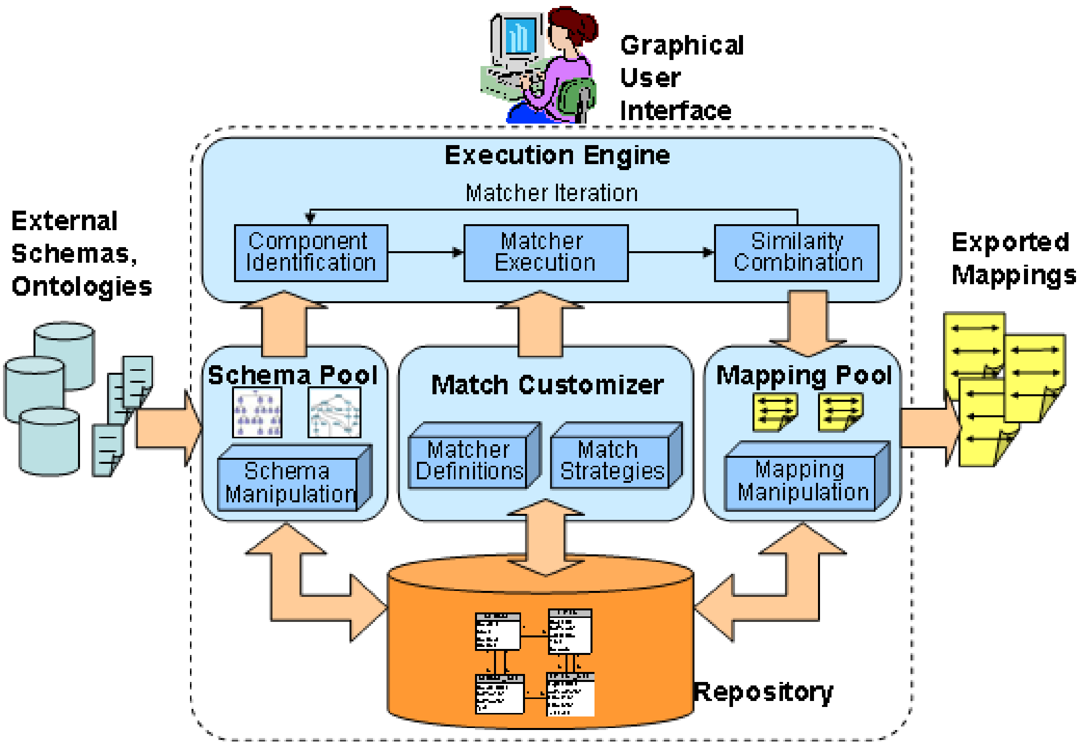



***Figure 2-6:*** *Overview of COMA++ Architecture [Aumuller 2005]*

The role of the users in COMA++ is to develop a mapping by validating candidate correspondences and appending any missed correspondences. There is no default navigation between candidate correspondences and it is left to the user to navigate this list themselves. The user can select the matching specific algorithms to use. The following is the process of user interaction with COMA++: (1) the user opens the tool, ontologies to be mapped and any partial mapping already constructed if available, (2) the user selects which matching algorithm(s) to use and modifies the parameters if required, (3) the user executes the matching algorithm and generates a set of candidate correspondences, (4) the user validates the candidate correspondences, (5) the user appends any missed correspondences, and (6) the user saves the mapping and closes the tool.

### 2.3.3 PROMPT

PROMPT is a plug-in for Protégé [Noy 2000] that allows the management for the merging and mapping of multiple ontologies. The PROMPT algorithm for the merging and mapping of ontologies, shown in Figure 2-7, creates an initial list of candidate correspondences based on class names and then loops through a series of steps: first the system takes one of the PROMPT suggestions and prompts it to the user, next the user validates the candidate correspondences, after which the system undertakes the operation involved in the validation, then the system automatically makes appropriate additional changes and finally updates the suggestions list accordingly. The role of the user is to validate the candidate correspondences suggested. The navigation of the suggestions is decided by the system, although users can choose to navigate the list of suggestions themselves. The series of suggestions continues until the user decides the mapping process has finished. Users can also append any missed correspondences with the tool.

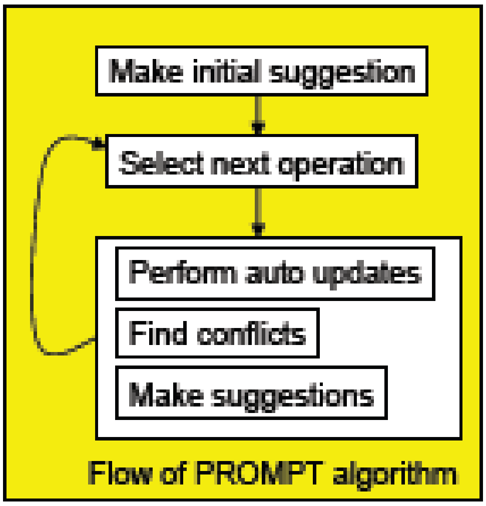



***Figure 2-7:*** *PROMPT Algorithm overview [Noy 2003]*

### 2.3.4 CoGZ

The Cognitive Support and Visualization for Human-Guided Mapping Systems (CoGZ) [Falconer 2009] is a mapping tool built on top of PROMPT. The mapping tool was designed to improve user support during the mapping process. Figure 2-8 displays the mapping process. The process however does not make a distinction between states and activities. The role of the user is to validate the candidate correspondences suggested and append any missed correspondences. The navigation between candidate correspondences is guided by suggestions from the tool but ultimately the user decides how to navigate.

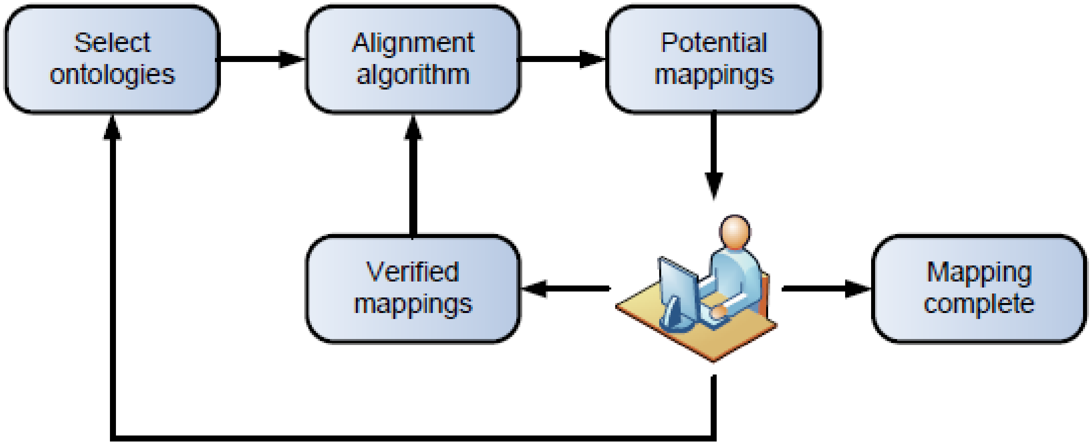



***Figure 2-8:*** *CoGZ mapping process [Falconer 2009]*

### 2.3.5 Summary

This section presents a comparison of the four mapping systems analysed. In particular, the comparison focuses on the involvement of the user in the mapping process and is displayed in table 2-2. The criteria for comparison focus on assessing the volume of the workload for the user in the mapping process and how the user interacts in the mapping process.

**Table 2-2:** Comparison between Selected Mapping Systems

| **Mapping System** | **Role of user** | **Generation of candidate correspondences** | **Approach to mapping** |
|---|---|---|---|
| *The MAFRA system* | Validate the candidate correspondences<br>Append any missed correspondence | The system uses its own matching API to generate the candidate correspondences | The user is required to devise their own mapping approach |
| *The COMA system* | Validate the candidate correspondences<br>Append any missed correspondence | The system allows the user to choose which matching algorithm and parameters to generate the candidate correspondences | The user is required to devise their own mapping approach |
| *PROMPT* | Validate the candidate correspondences<br>Append any missed correspondence | The system uses its own matching API to generate the candidate correspondences | The candidate correspondences are suggested in an iterative loop |
| *CoGZ* | Validate the candidate correspondences<br>Append any missed correspondence | The system allows the user to choose which matching algorithm and parameters to generate the candidate correspondences | The candidate correspondences are suggested in an iterative loop |

In each system the role of the user is the same, requiring the user to create a mapping by validating the candidate correspondences suggested and appending any missed correspondences. Two of the systems, the COMA system and CoGZ, allow the user to modify the matching approach used to generate the candidate correspondences. While another two of the systems, the MAFRA system and the COMA system, require the user to devise their own approach for mapping. The other two systems, PROMPT and CoGZ, guide the user through the mapping process by presenting suggested candidate correspondences in an iterative loop. However users are free to take their own mapping approach with these systems. In fact in a recent study with the ontology mapping community it was shown that less than 15% of the people surveyed relied on the order in which the suggestion were shown by the mapping system [Falconer 2009]. All of the systems rely on the user to find and append any missed correspondences.

The mapping systems which were reviewed are representative of the current mapping process of the state of the art mapping systems. The systems have emphasised that the level of user involvement in the mapping process is high. The user-specific mapping tasks can be time-consuming and/or highly cognitive. The full list of user-specified mapping tasks were derived from analysis of the state of the art mapping systems and are displayed in table 2-3. The "Required" column indicates if the task needs to be accomplished by the user or if the task is optional and dependent on the implementation of the mapping system.

**Table 2-3:** List of User-Specific Mapping Tasks

| Step | Mapping Task | Description | Required |
|---|---|---|---|
| 1 | Decide when to map | Involves the user deciding when it is time to generate or modify a mapping | Yes |
| 2 | Open mapping tool | Involves the user opening the mapping tool in their computing environment | Yes |
| 3 | Open ontologies | Involves the user locating and opening the ontologies for mapping within the mapping tool | Yes |
| 4 | Open mapping | Involves the user opening any previously saved mapping or collaborative shared mapping which is to be modified | No |
| 5 | Select matching algorithm | Involves the user selecting a matching algorithm which will be used to generate candidate correspondences, the algorithm may be specific to the domain of the ontologies | No |
| 6 | Configure matching algorithm | Involves the user configuring the parameters of the matching algorithm to best suit the ontologies to be mapped within the mapping tool | No |
| 7 | Execute matching algorithm | Involves the user executing the matching algorithm to generate the candidate correspondences | No |
| 8 | Decide mapping approach | Involves the user deciding the approach to take to map the ontologies | Yes |
| 9 | Validate candidate correspondence | Involves the user validating a candidate correspondence generated within the mapping tool by analysing the structure of each concept and making an informed choice based on this analysis | Yes |
| 10 | Append missing correspondence | Involves the user searching through each ontology looking for any potential correspondence the matching algorithm missed | No |
| 11 | Decide when mapping finished | Involves the user deciding when the work done is enough to end the mapping session | Yes |
| 12 | Save mapping generated | Involves the user saving the mapping constructed | Yes |
| 13 | Deploy mapping | Involves the user deploying the mapping generated | Yes |

The first step requires the user to decide when there is a need to develop or alter a mapping. This is a high cognitive task requiring that the user to understand when there is a need for mapping. Steps two to four necessitate that the user opens the mapping tool, ontologies and mapping file if available. Steps five to seven require that the user needs to understand the matching algorithm provided by the mapping tool and the circumstances of when the algorithm should be used. These steps require the user to understand the benefits and downsides of using algorithms in different domains and are high cognitive tasks. The eighth step involves the user deciding which mapping approach to take. This is non-trivial and is a high cognitive task. In fact a study carried out on the mapping community showed that there is no consensus on which mapping approach to take and is usually down to personal preferences [Falconer 2009]. Steps nine and ten require the user to map the two ontologies by validating candidate correspondences and appending any missing correspondences. These steps involve reading the information of the candidate correspondences to make an informed decision and to search through the ontologies to locate any missing correspondences. These are high cognitive tasks and necessitate that the user have knowledge and understanding of the domain the ontologies represent. Step eleven is a high cognitive task and requires the user to understand when the mapping session has concluded. The user must understand when the level of work done is sufficient. The final two steps require the user to save the mapping constructed and know how to deploy the mapping.

The steps above reveal that the mapping process is both labour intensive and requires that the user has strong knowledge engineering skills. This indicates that the user in this process would need to be a person with a knowledge engineering background or a person who has been trained by or is working hand in hand with a knowledge engineer.

## 2.4 Ontology Mapping Tool User Interfaces

As identified earlier, fully automatic derivation and generation of mappings from correspondence matches is considered impracticable [Noy 2004]. In general, the mapping process requires the need for a "human in the loop". A large variety of mapping tools exist to help users develop mappings. Ontology mapping tools typically use graphical user interfaces to allow users to validate candidate correspondences and append missing correspondences with the tool. The rest of this section reviews a list of mapping tools with different types of graphical user interfaces. As there has been no formal and systematic review of such interfaces from the perspective of interest to this thesis (i.e. how to best support a user generate mappings), the commentary below is primarily based on the opinion of this author. The questions asked by the author were: 'would the interface allow ordinary users to use the tool?', 'what type of technical skills would users need to use the tool effectively?' and 'would ordinary users be able to navigate the correspondences efficiently?'.

### 2.4.1 Chimaera

Chimaera [McGuinness 2000] is a software system that supports ontology merging. The system has a web-based interface where the user interacts with web forms to upload ontologies, select algorithm parameters, and merge similar ontology entities, see Figure 2-9 for an example. The merge algorithm produces a list of candidate correspondences based on term name similarity, term definitions, possible acronyms and expanded forms, and suffix matching [Euzenat 2004a]. Chimaera supports OWL ontologies and produces the mapping results in OWL descriptions. The tool is accessed from the web which allows users to easily enter the tool from within their internet browser. The algorithm provides a list of suggestions which can guide the user through the mapping process. The interface is awkward looking due to the volume of information displayed and the appropriate information for each correspondence not being significantly emphasised and highlighted. In the authors' opinion, this type of interface could be unappealing and hard to understand for ordinary users.

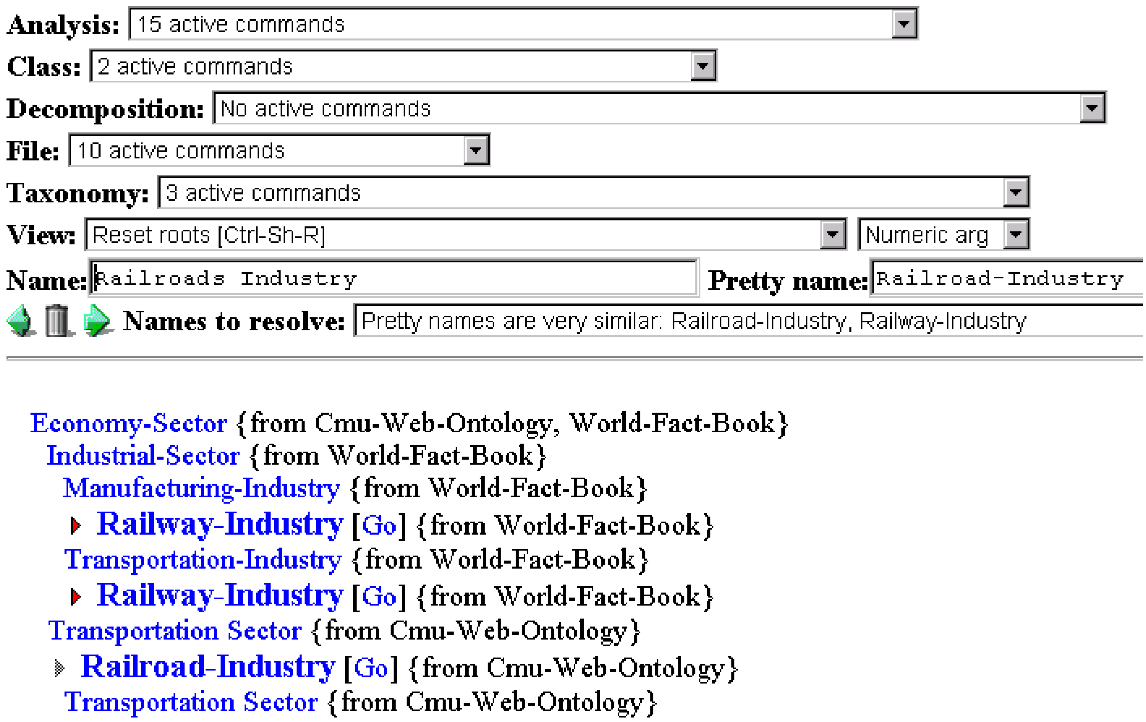



***Figure 2-9:*** *Screenshot of Chimaera interface for merging two classes*

### 2.4.2 COMA++

COMA++ [Aumuller 2005] is a standalone mapping tool that uses a graph based graphical user interface. Both the source and target ontologies are displayed with a tree type structure and users can scroll through the ontologies using scrollbars. COMA++ automatically generates candidate correspondences between source and target schemas (XML or OWL), and constructs lines to represent these correspondences visually, see Figure 2-10. Users can also append missing correspondences by interacting with the schema trees and constructing lines between entities. Hovering over a line displays a confidence measure of the correspondence as a numerical value between zero and one. The navigation of the candidate correspondences is left to the user. In the author's opinion, users would need to have ontology modelling knowledge to understand the information displayed and use the tool, otherwise they would just be label matching, i.e. just matching based on the name of the concept and not taking into account the structure of the concept.

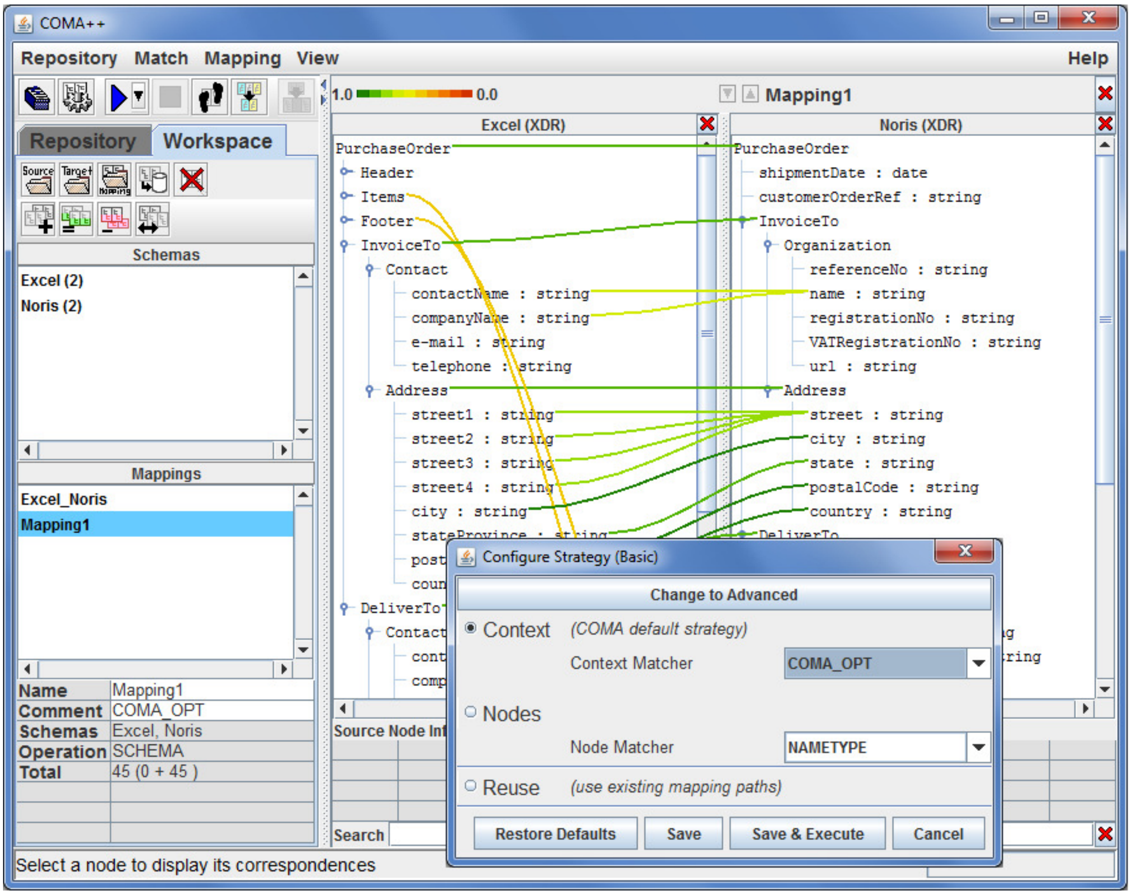



***Figure 2-10:** Screenshot of COMA++ interface*

### 2.4.3 PROMPT

PROMPT [Noy 2003], see Figure 2-11, is a plug-in for the popular ontology editor Protégé [Noy 2000] which supports tasks for managing multiple ontologies including ontology differencing, extraction, merging, and mapping. It represents the ontologies in a tree type structure which users can scroll through using scrollbars. PROMPT begins the mapping procedure by requiring the user to specify a source and target ontology. It then computes an initial set of candidate correspondences based largely on lexical similarity between the terms used in the ontologies the user selected. PROMPT then presents a candidate correspondence suggestion to the user based on this list, using an algorithm, and the user validates the candidate correspondence. Once a user has validated the candidate correspondence, PROMPT's algorithm uses this to perform structural analysis based on the graph structure of the ontologies. This analysis usually results in further candidate correspondence suggestions. This process is repeated until the user determines that the mapping is complete. Users can append any correspondences that were missed by the algorithm. The interface represents the ontologies in a graph type display similar to COMA++. A simple proof is provided for each suggested candidate correspondence. However for users to be able to use the tool effectively, in the author's opinion, they will need to have modelling knowledge, to understand the concept structures and their significance.

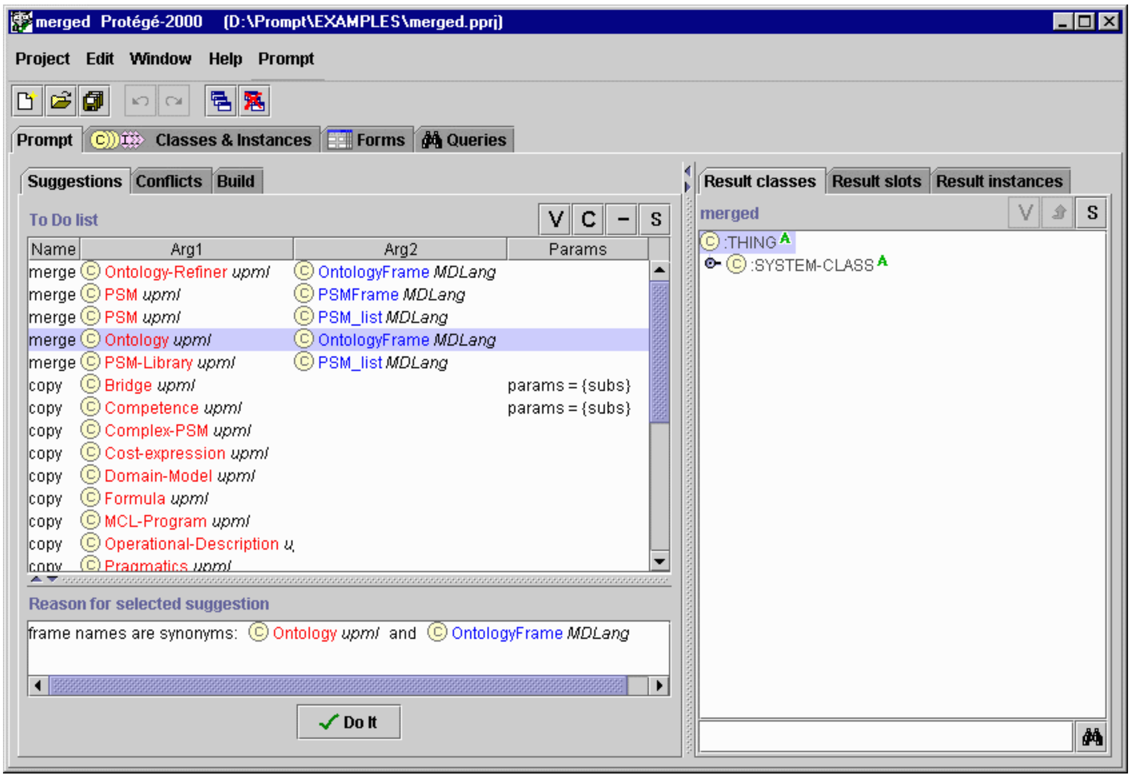



*Figure 2-11:* *Screenshot of list of candidate correspondences suggestions by PROMPT plug-in*

## 2.4.4 AlViz

AlViz [Lanzenberger 2006] is a plug-in for Protégé which was specifically designed to support the mapping of ontologies visually. It applies multiple-views via cluster graph visualization along with synchronized navigation within standard graph tree type, see Figure 2-12. The tool attempts to facilitate user understanding of the candidate correspondences generated by providing an overview of the ontologies in the form of clusters. The clusters represent an abstraction of the original ontology graph and moreover the clusters are coloured based on their potential concept similarity with the other ontology. The tool highlights possible correspondences through visualisation although the navigation is left to the user. In the author's opinion, users would still need to have modelling knowledge to be able to use the tool effectively.

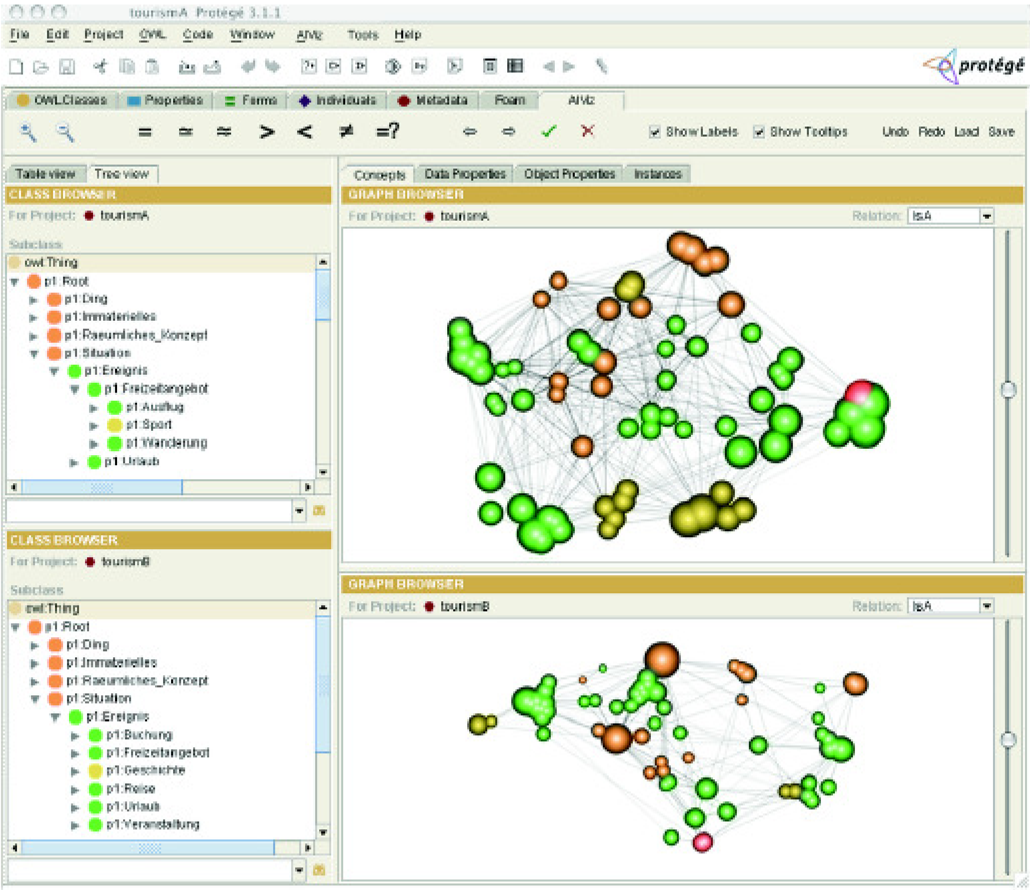



*Figure 2-12: Screenshot of AlViz plug-in while mapping two ontologies*

## 2.4.5 OLA (OWL Lite Alignment)

OLA (OWL Lite Alignment) [Euzenat 2004b] is a tool for mapping construction as well as an environment for manipulating mappings. The tool supports parsing and visualization of ontologies, automated computing of similarities between ontology entities, manual construction of mappings, visualization of mappings, and comparison of mappings, see Figure 2-13. OLA only supports OWL Lite ontologies[11] and uses the Alignment API specified in [Euzenat 2006] to generate candidate correspondences. The matching algorithm finds candidate correspondences by analyzing the structural similarity between the ontologies using graph-based similarity techniques. This information is combined with label similarity measures to produce a list of candidate correspondences. Graphically, ontology terms have different visual representations which could potentially help the user identify the information better. However no user evaluation has been conducted on the tool to evaluate this. In the author's opinion, using different representations for terms offer a good way of highlighting and identifying their differences. However as the ontology grows the visualisation will become unmanageable due to

[11] http://www.w3.org/TR/owl-features/

the volume needed to be represented. Although the visualisation separates and highlights different terms of the ontology, the information is still represented in a complex way and ordinary users still may not be able to interpret the information.

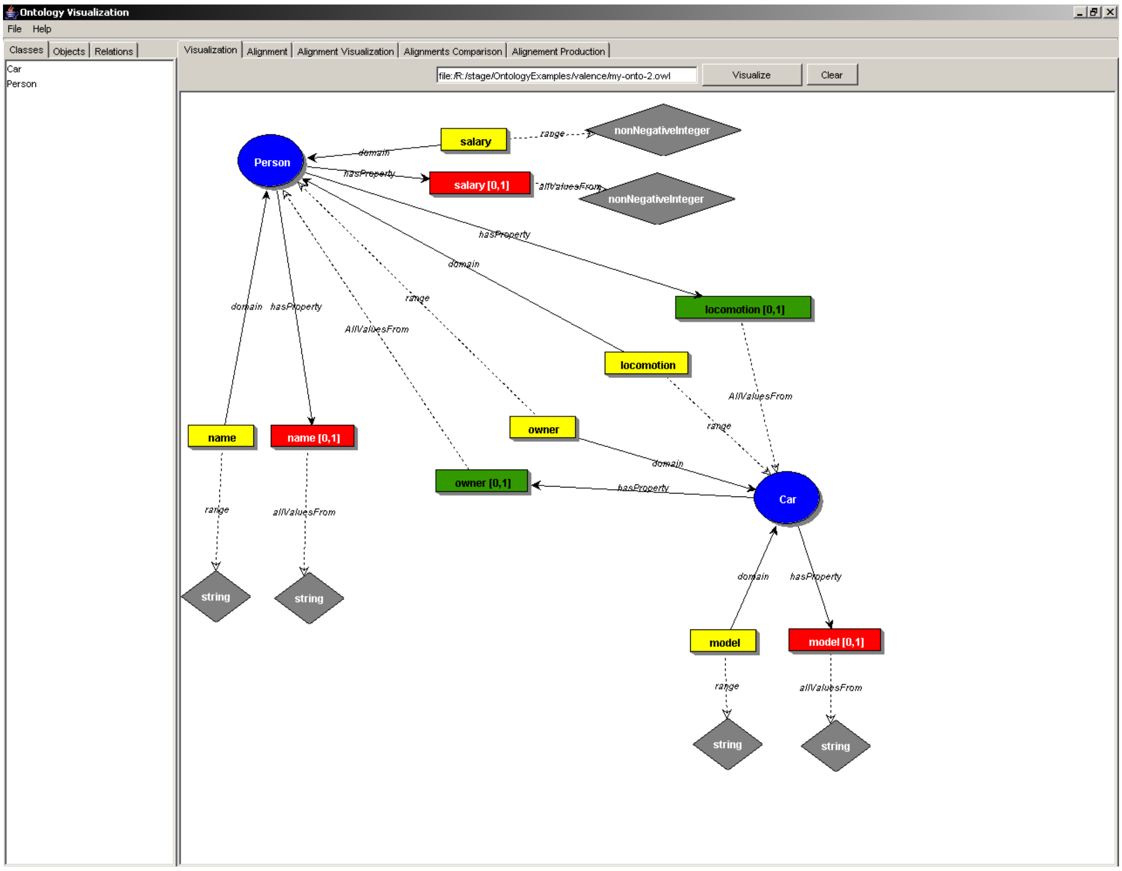



***Figure 2-13:*** *Screenshot of OLA visualisation of OWL ontology*

## 2.4.6 SMART

SMART [Morishima 2005] is a tool for semantic driven creation of complex XML mappings. It represents the source and target schema using a UML format, see Figure 2-14. The correspondences between the two schemas are represented using lines and symbols to represent the relationship. For example '=' symbol is used for equivalence. The symbols and UML structure could help non-knowledge engineers understand the information better. However as mappings grow the interface could become unmanageable due to the volume of information being too large to display clearly to the user. In the author's opinion, the symbol offers a good way of representation the relation to ordinary users. However the user would still need modelling skills to understand the information and structure of the concepts.

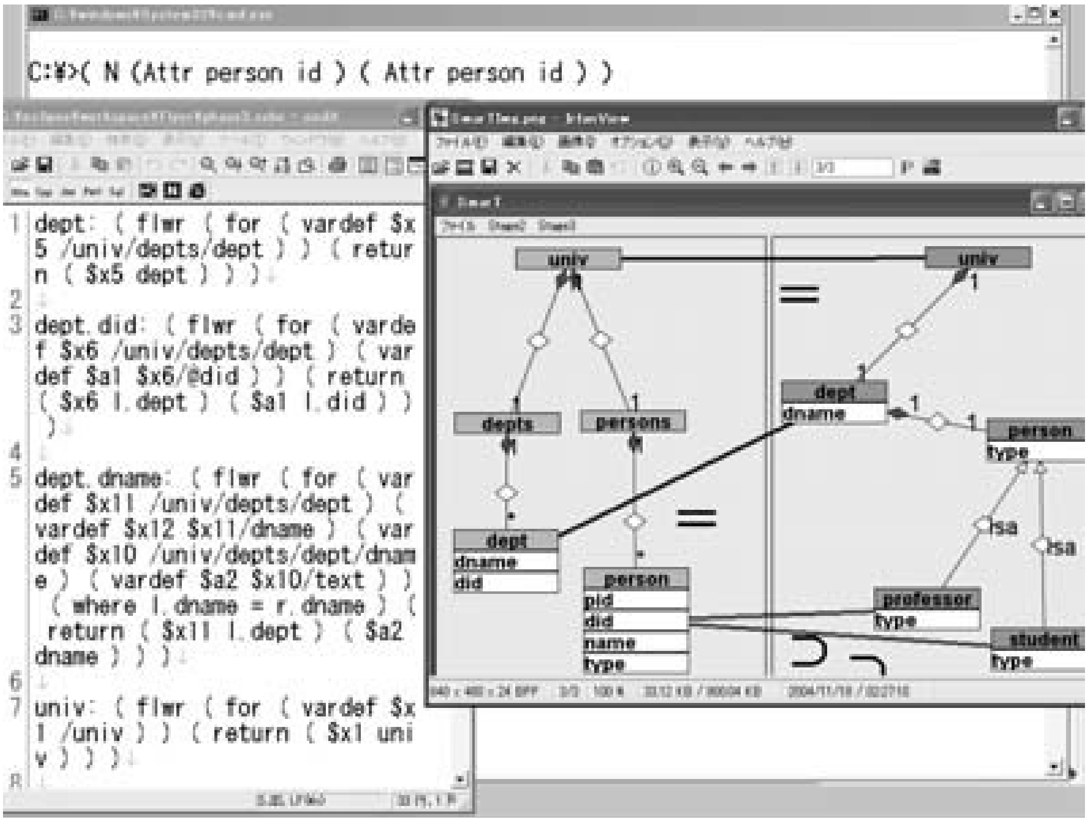



*Figure 2-14:* *Screenshot of SMART interface*

## 2.4.7 Schema Mapper

Schema Mapper [Raghavan 2005] is a hyperbolic interface for schema mapping where both the source and target schemas are represented by hyperbolic graphs. A mapping table is used to hold the correspondences while each term is coloured to identify if it has been mapped. When a term is selected by the user a table displays the suggested candidate correspondences for the term. The candidate correspondences are generated with a simple string comparison between the term names. The colour gives a good indicator of the state of each concept. However the hyperbolic graph will lead to an awkward and hard to read interface when a schema becomes too big. There is no option to zoom in and out on different parts of the graph which may have partially solved this issue. In the author's opinion, the visualisation may offer a way for ordinary users to view the local neighbourhood of each concept but the distinction and identification of the information would be practically impossible for ordinary users to understand. In addition, the navigation of the ontology may lead to ordinary users becoming confused and disoriented due to parts of the graph becoming impossible to read and identify.

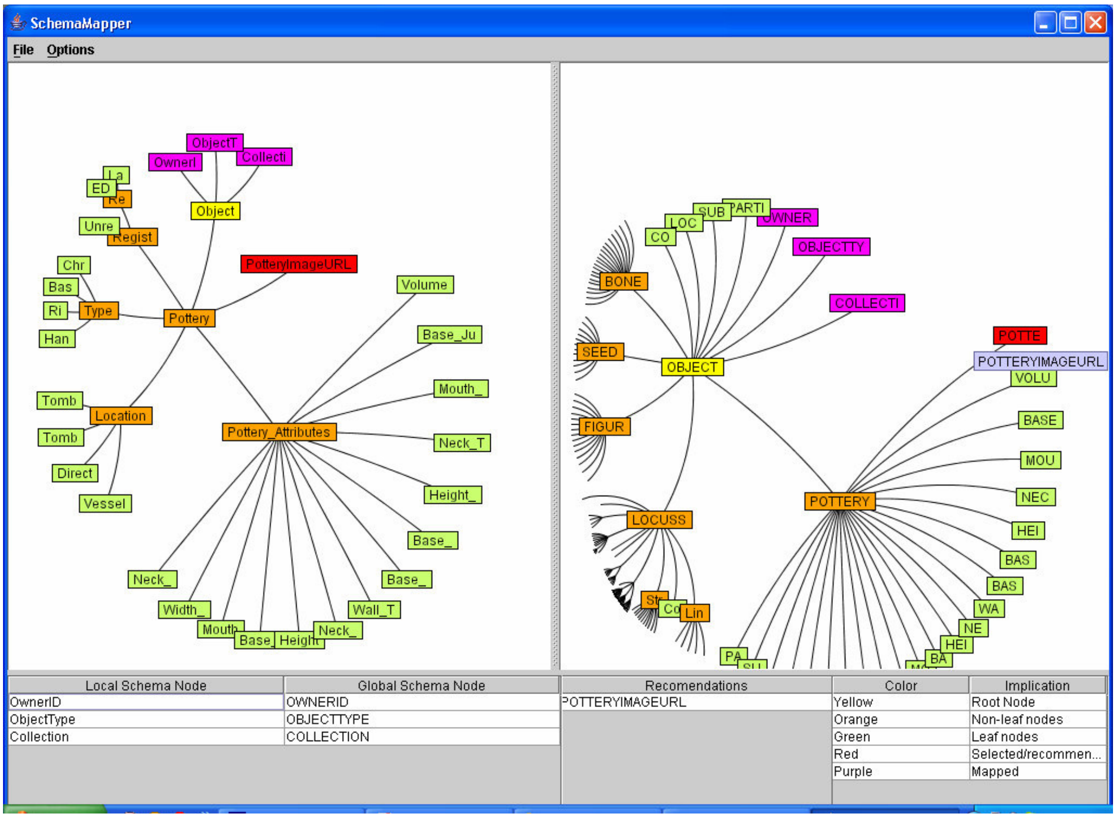



***Figure 2-15:*** *A Screenshot of Schema Mapper interface*

## 2.4.8 Roundtrip Ontology Authoring

Roundtrip ontology authoring [Davis 2008] is an ontology authoring tool and is not an ontology mapping tool. However, traditionally ontology mapping tools have closely followed the steps of ontology authoring tools. In fact, most mapping tools have even matched the interface design of ontology authoring tools. Recently, authoring tools with natural language interfaces (NLIs) have been designed to allow domain experts to engage in the authoring process. Kaufmann and Bernstein [Kaufmann 2007] presented an evaluation study on the usability of natural language interfaces and concluded that they are indeed useful for ordinary end-users. NLIs allow users to formulate their knowledge in familiar natural language rather than learn an unfamiliar formal and complex system. The natural language interfaces used in ontology authoring tools allow domain experts to edit and query ontologies using a control natural language, a type of English. But the ontology is still represented and visualised using a graph-type display, similar to the interface shown in section 2.4.3. Examples of these types of authoring tools include GINO [Bernstein 2006a] and ROO [Dimitrova 2008].

Roundtrip ontology authoring in contrast also uses a natural language interfaces to display the ontology to the user, see Figure 2-16. This allows feedback of the actions to be visualised to users in a way they are accustomed to. The user evaluation study found favourable results among domain experts with the Roundtrip ontology authoring tool and might indicate that this type of tool as ontology editor might be a preferred option for ontology development within industry. In the author's opinion, representing the ontology using natural language will allow ordinary users to comprehend and understand the information.

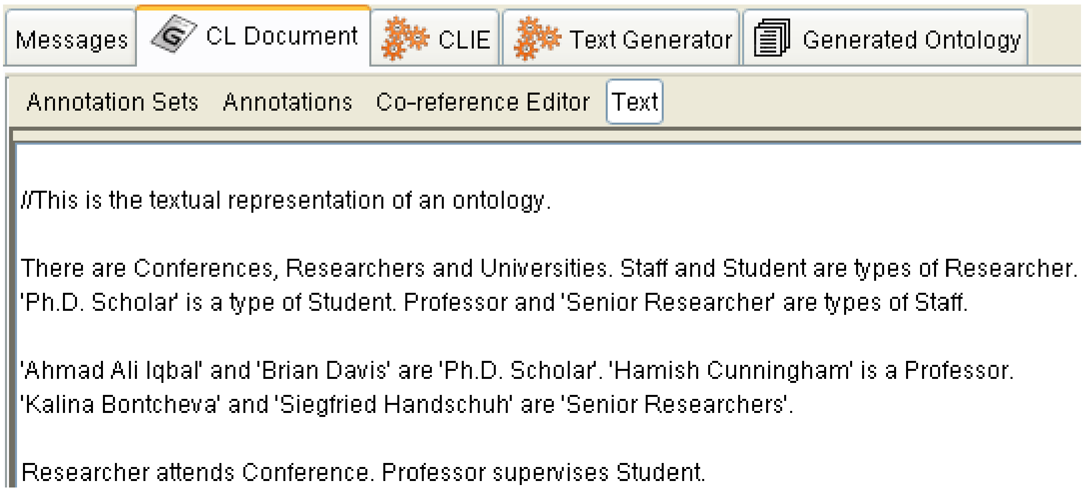



***Figure 2-16:*** *Screenshot of Ontology represented in Natural Language by Roundtrip Ontology Authoring*

### 2.4.9 Summary

Table 2-4 gives a comparison of the user interfaces used by the mapping tools reviewed in the previous sections. The table compares how each tool represents mapping correspondences to the user. The criteria assess how the task is visualised, navigated and what information is supported. The Roundtrip ontology authoring tool is not included in the table as it is not a mapping tool.

**Table 2-4:** Comparison of User Mapping Interfaces

| Mapping Tool | Interface | Navigation | Global View | Local Neighbourhood |
|---|---|---|---|---|
| *Chimaera* | Web form with Item List | Suggestions | No | Yes, shows the local structure of each concept |
| *COMA++* | Tree Graph | Left to the user | Yes | No, the local neighbourhood is meshed into the global view |
| *PROMPT* | Tree Graph | Suggestions | Yes | No, the local neighbourhood is meshed into the global view |
| *Alviz* | Cluster Graph | Left to the user | Yes | Yes, shows local structure of both concepts in a panel view |
| *OLA* | Direct Graph | Left to the user | Yes | No, the local neighbourhood is meshed into the global view |
| *SMART* | UML Graph | Unknown | Yes | No, the local neighbourhood is meshed into the global view |
| *Schema Mapper* | Hyperbolic Graph | Suggestions | Yes | Yes, the hyperbolic graph centres around each concept showing the local neighbourhood |

The comparison shows that a wide variety of visual representations have been used to represent ontological information to users. However the interfaces of the mapping tools require the user to have modelling skills to understand the information being represented. Three of the mapping tools have features which may help ordinary users understand the information being represented. OLA displays the ontology terms in different colours and representations which allows users to see the distinction between the information. SMART displays the relationship of a correspondence with a symbol which can help ordinary users understand the relationship. Schema Mapper only shows the local neighbourhood of each concept being considered for mapping. This highlights only the relevant information for the correspondence being considered. Furthermore, the Roundtrip ontology authoring tool allows ordinary users to understand the information of ontologies by displaying the information in natural language.

## 2.5 Ontology Mapping Tools with User Support

There has been an emerging trend to add user support features into ontology mapping tools. The research and tools in this section have been designed to enrich user support within the mapping process. In this section six mapping tools and four current mapping approaches which aim to improve user support are reviewed.

### 2.5.1 Webscripter

Webscripter [Yan 2003] uses a spreadsheet interface to allow users to quickly assemble, fuse and extract data from multiple DAML sources into reports. Users can then publish their reports on the web and have them automatically refresh to contain the most up-to-date information. The published reports can be output in DAML, thus allowing reports themselves to become the source of further reports, see Figure 2-17. Hence users can map between ontologies by graphically corresponding their data from multiple sources in a simple spreadsheet-like view without having to know anything about ontologies or even taxonomies. This approach is primarily focussed on allowing ordinary users to be the "human in the loop". However this approach has two main issues. Firstly it requires the users to manage their own data which would be non-trivial and problematic for ordinary user. Secondly it requires instances to be available within each of the ontologies to be mapped. No user study has been carried out with ordinary users to evaluate the potential of WebScripter.

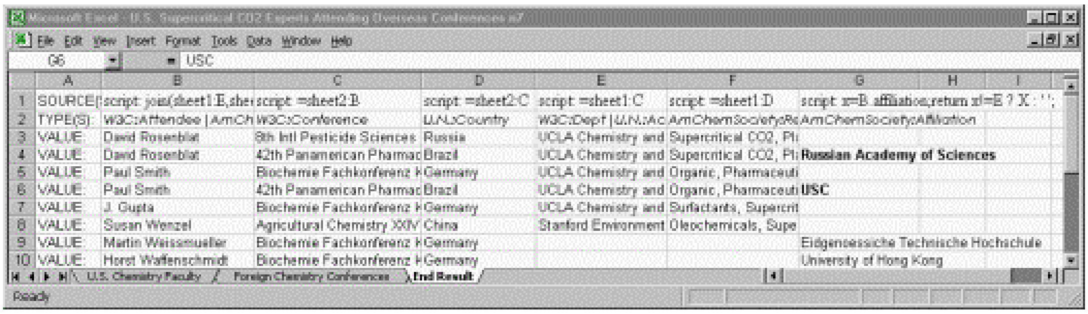


***Figure 2-17:*** *A sample report of WebScripter [Yan 2003]*

### 2.5.2 Visualisation of Mappings

Robertson et al. [Robertson 2005] have presented a novel approach to the visualization of the mappings between schemas. They modify BizTalk [BizTalk 2005], which is a graph-based schema mapping tool similar to COMA++, with several standard visualisation techniques to address the problems of scale experienced by current users of mapping tools for navigation tasks, see Figure 2-18. The list of the visualisation techniques used is the following: *highlight propagation* which highlights all links of a schema's element when it is selected, *auto-scrolling* correspondences are automatically centred when selected, *coalescing trees* filtering away any unnecessary information for the term selected, *multi-select* displays how multiple schema elements interact with multiple selection capability, *incremental search* allows search of the tree with a text box predictive search function, *bendable links* to make the correspondences more clearly identifiable, and *focus on linked elements* allow shortcut keys. The approach focuses on improving the mapping interface for knowledge engineers. However in the author's opinion, the interface would still be too complex for ordinary users.

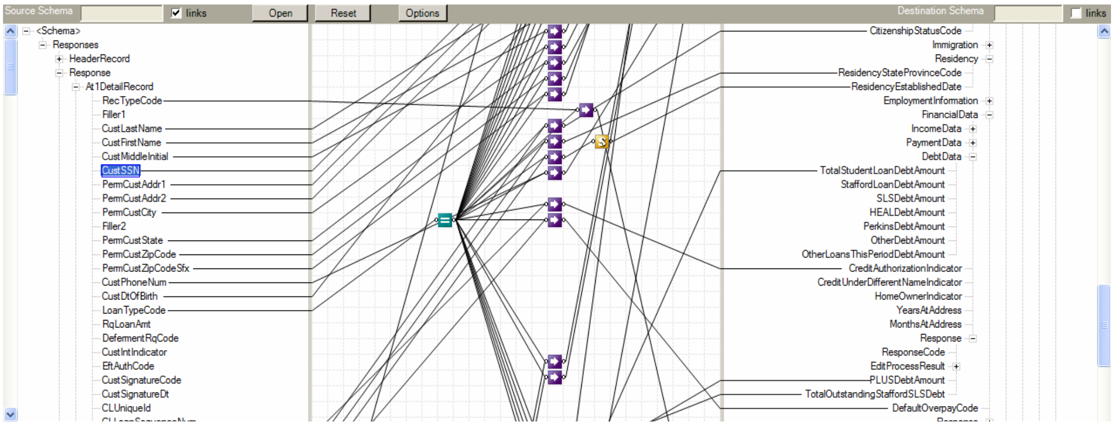



***Figure 2-18:*** *Screenshot of modified BizTalk interface [Robertson 2005]*

## 2.5.3 Incremental Schema Matching

Bernstein et al. [Bernstein 2006b] have presented an incremental schema matching technique which displays only suggested candidate correspondences for a selected element rather than for the global set of schema elements. They modified the BizTalk schema mapping tool to make their prototype, see Figure 2-19. The history of the user's mapping actions is also taken as a heuristic in selecting candidate correspondences for each schema element. The focus of this work is again on making the mapping process more usable, navigation in particular, for the current users of mapping tools. The approach reduces the complexity of the mapping process by focussing on only one source schema element at a time. However, no user evaluation has been published.

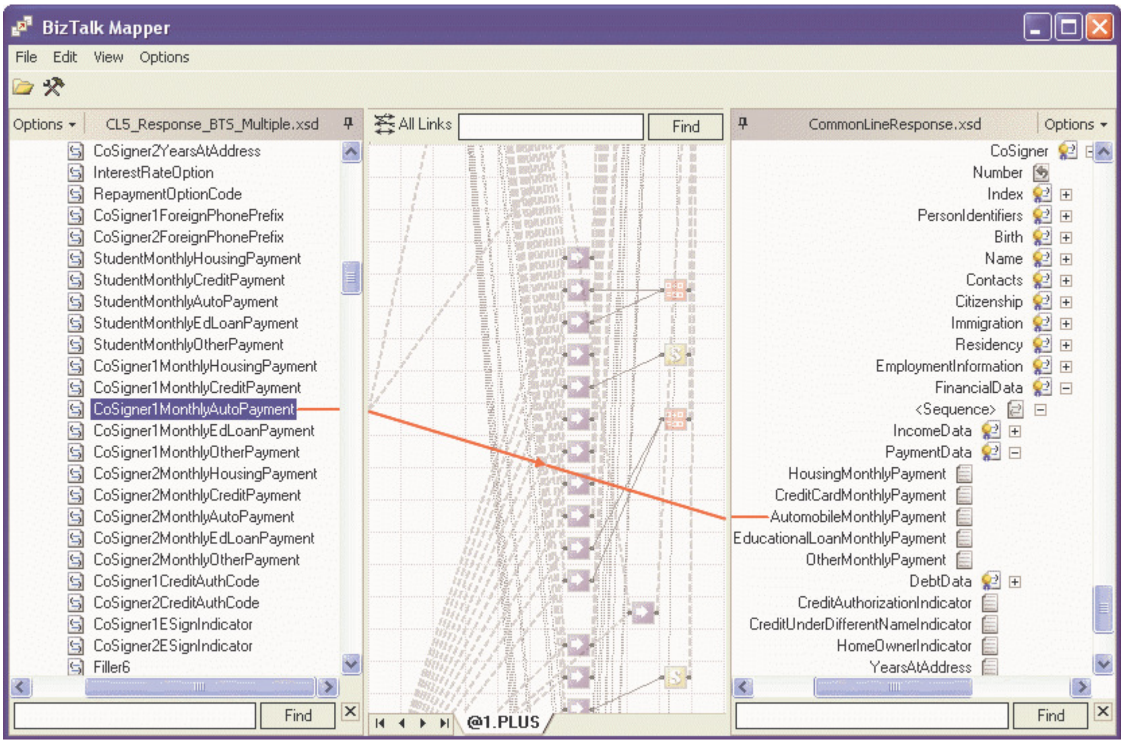



***Figure 2-19:*** *Screenshot of modified BizTalk mapper interface [Bernstein 2006b]*

## 2.5.4 CoGZ

CoGZ [Falconer 2007] is a plug-in for PROMPT that is a Cognitive Support and Visualization for Human-Guided Mapping Systems. The tool is derived from their framework for cognitive support in ontology mapping. This framework is specifically designed to improve the mapping experience for the current users of mapping tools. Figure 2-20 displays this theoretical cognitive framework.

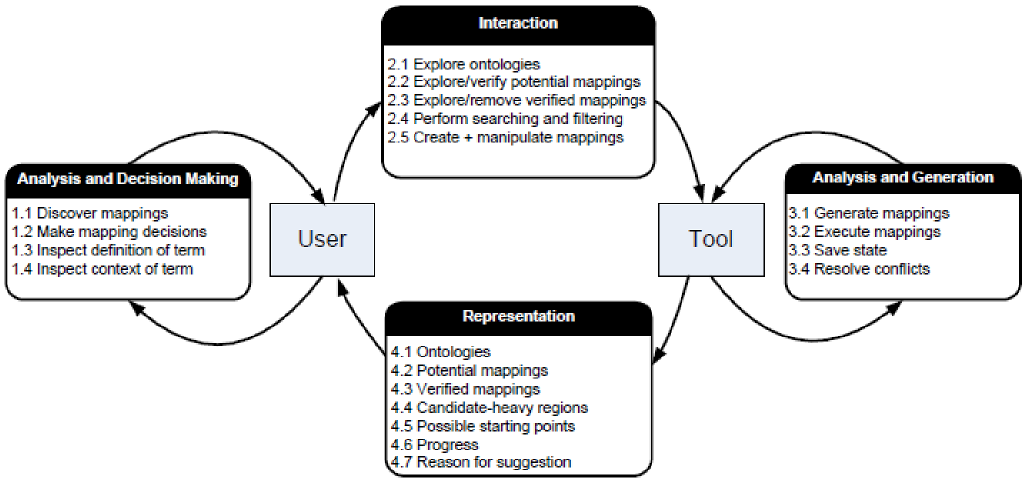



***Figure 2-20:*** *Theoretical framework for cognitive support in ontology mapping [Falconer 2009]*

This framework gives a guide on aspects that need to be included in mapping tools. In the design and implementation of the tag-based mapping tool this framework will be referenced with emphasis on how each aspect was supported, see section 5.1.6. Figure 2-21 gives a screenshot of CoGZ which is the mapping tool built on this framework. The tool provides candidate correspondence suggestions with proofs which help navigate the user through the mapping process. Both the global and local neighbourhood of each concept selected is represented visually to the user. The evaluation results of CoGZ indicated that there was an improved performance than using PROMPT for knowledge engineers. However in the author's opinion, the interface is too complex for ordinary users and although the cognitive framework gives a guide on aspects that need to be included for mapping tools the focus on the design was for knowledge engineers and not ordinary users. It still remains to be evaluated whether all the aspects need to be supported for ordinary users. This is addressed by the author in the design and implementation of the tag-based mapping tool, see section 5.1.6.

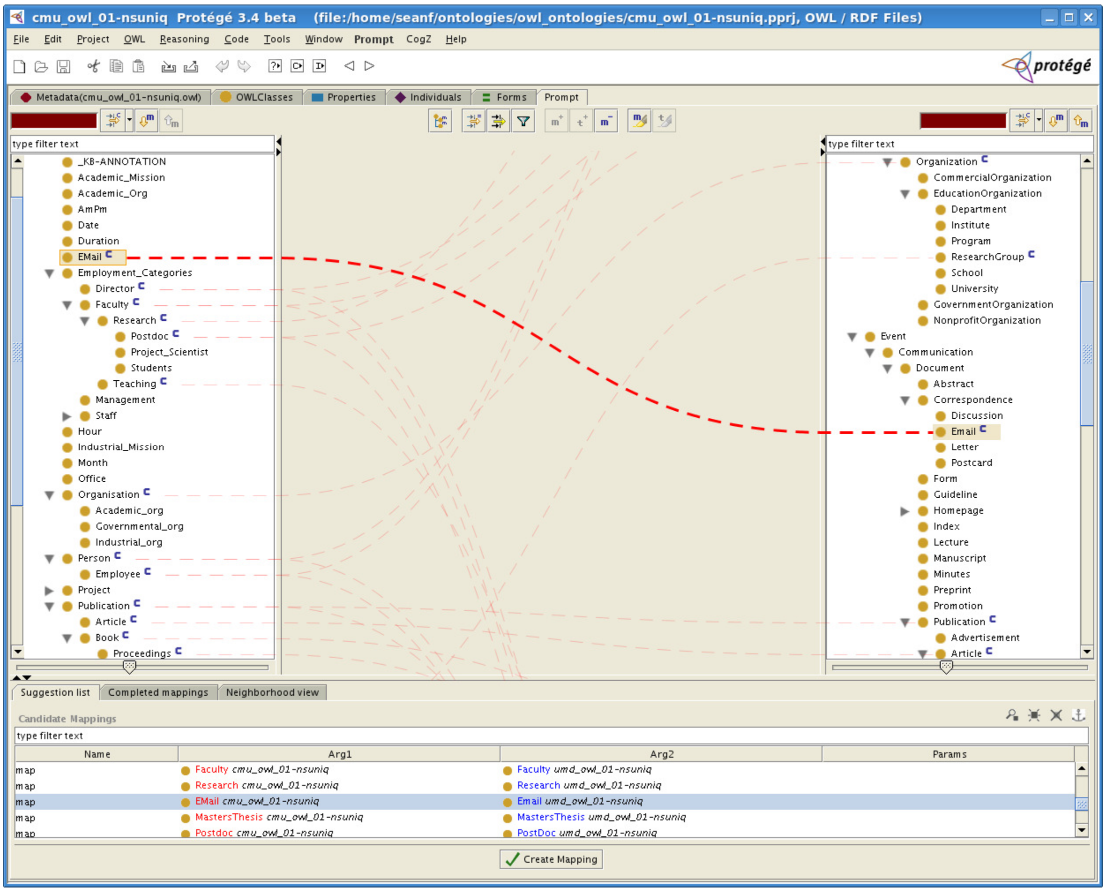



***Figure 2-21:*** *Screenshot of CoGZ interface [Falconer 2009]*

### 2.5.5 Community-Driven Ontology Mapping

Zhdanova et al. [Zhdanova 2006] have extended the idea of ontology mapping to community-driven ontology mapping. The approach is to partition people into multiple different communities (groups) where mappings can be shared, see Figure 2-22. The approach allows people to be members of multiple groups. It also extends and preserves the advantages given to communities on the web. A prototype was developed based on their approach and showed the feasibility of acquisition and sharing of ontology mappings among the web communities. However, the emphasis of their approach was to improve the recall of matching tools via repository of ontology mappings for different domains rather than investigating the user experience with the tool. Recently, Noy et al. [Noy 2008] have taken the collaborative mapping approach to the medical domain. However there has been no published user evaluation of their approach.

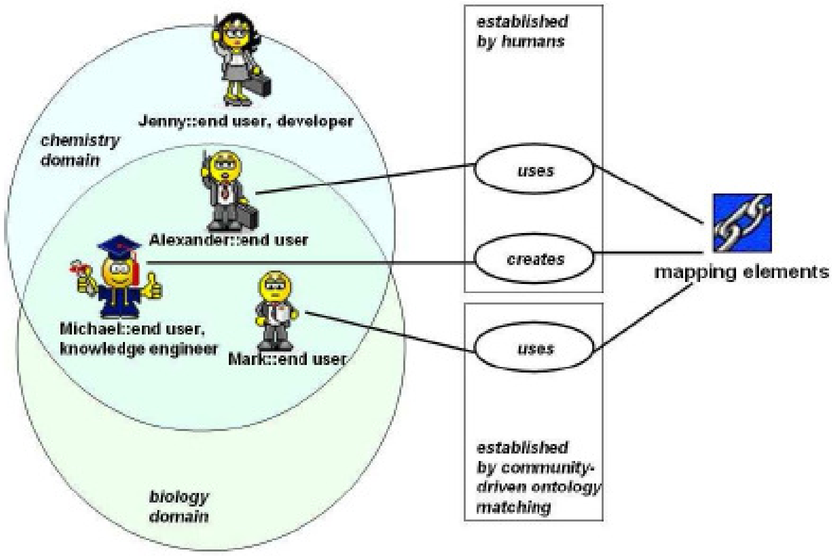



***Figure 2-22:*** *Collaborative mapping process presented in [Zhdanova 2006]*

### 2.5.6 Web Explanation for Semantic Heterogeneity Discovery

Shvaiko and colleagues [Shvaiko 2005] have provided a semantic mapping system which provides proofs and explanations in natural language without technical detail. The system extends S-Match [Giunchiglia 2004] to use the Inference Web infrastructure [McGuinness 2003] and its Proof Mark-up Language (PML) [Pinherio 2008] to expose meaningful fragments of S-Match proofs in which the user can browse, see Figure 2-23. The proofs are designed to use short, natural language, high level explanations without any technical details. These are designed to be intuitive and understandable by ordinary users. No user study has been carried out on the system. Also no tool has been developed using this system.

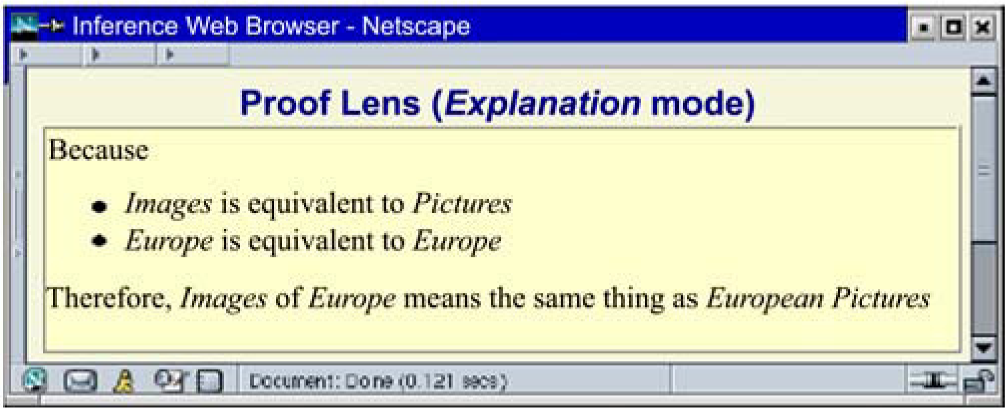



***Figure 2-23:*** *Example of proof of a candidate correspondence [Shvaiko 2005]*

### 2.5.7 Formal Model for Ontology Mapping Creation

Mocan et al. [Mocan 2006] have proposed a formal model for ontology mapping creation which provides the use of connected perspectives to deemphasize non relevant aspects of the ontologies being mapped while displaying the relevant entries for each candidate correspondence. This model hides the complexity of the underlying logic and gives a better understanding of the mapping actions. Once again the focus is to improve the usability of mapping tools for the current users of mapping tools. There has also been no publication of any user evaluation. In the author's opinion, although the views may hide some of the complexity, the user will still need modelling experience to understand the information.

### 2.5.8 Actively Learning Ontology Mapping via User Interaction

Recently, a system of learning mapping through user interaction has been proposed [Shi 2009]. The approach uses active learning, which is widely used in machine learning, on the user feedback for validating one candidate correspondence to validate several other correspondences. The system uses a correct propagation algorithm to find the most informative candidate correspondences, which will affect other candidate correspondences, to present to the user. This approach limits the number of user interactions while maximising the accuracy through the propagation of interactions. The focus of the research is to reduce the volume of mapping actions for current users of mapping tools and not on reducing the complexity of the interfaces for mapping tools.

### 2.5.9 Ontogame

Ontogame [Siorpaes 2009] aims to stimulate user participation in semantic content creation by designing a series of entertaining multi-player games that hide the tasks related to the content creation. It is based on the model of 'games with a purpose' [Von Ahn 2006] where people can collectively solve large scale problems. One of the suggested games was for ontology mapping. However this approach has only been theoretically suggested and has not been implemented or seen any user evaluation. The main focus of this approach is to expand the audience of semantic content creation away from technical experts to other users who are interested in the games. However, this will only appeal to a niche group and not all ordinary users. Also in the author's opinion, this group of users may not have the sufficient background knowledge to answer mapping tasks that are not related to their knowledge.

### 2.5.10 Dialogue-Based Interactive Semantic Meditation

Soantag [Soantag 2008] has proposed an approach for a dialogue-based interaction mapping process where medical domain experts will validate candidate correspondences via verbal communications. Only a theoretical poster paper has been published on this approach and no user study has been carried out. The suggested approach does not address the issue of the interface being too complex for ordinary users.

### 2.5.11 Summary

The research presented in this section is improving the mapping experience for users. However, they are predominantly focused on the current users of mapping tools, knowledge engineers, and any research focussed on expanding the audience for mapping tools is either theoretical, or only aimed at a subset of ordinary users or has not seen any meaningful user evaluation.

## 2.6 User Interactions in Ontology Mapping Tools

This section details the different types of interactions mapping tools use to allow users to develop mappings. How the user interacts with mapping tools and develops mappings is important when considering the usability of the tools. The interaction needs to be simple and even familiar to use to allow ordinary users to develop mappings with the tool. Table 2-5 outlines six different types of interactions that have or could potentially be used by mapping tools. Details are given on how each interaction modifies the correspondences of mappings, see section 2.1. The first four are interactions which have been implemented in current state of the art mapping tools. The final two are interactions which have not yet been implemented in mapping tools. The first of these interactions is the structure sentence interaction which is the proposed approach of ontogame, see section 2.5.9, and is also an approach used by ontology authoring tools, see section 2.4.8. The second interaction is a tagging approach which was identified from the popularity of tagging in web 2.0 websites and applications.

**Table 2-5:** Types of Interaction that are/can be used in Mapping Tools

| Type | Understandable[a] | Flexibility[b] | Action[c] |
|---|---|---|---|
| Button Press | High | None | Modification of the confidence measure |
| Construct Lines | Medium | Low | Modification of the confidence measure & Generation of new correspondences |
| Enter Formulas | Low | High | Modification of the relationship & Generation of new correspondences |
| Data Grid | Low | Low | Generation of new correspondences |
| Structure Sentences | High | None | Modification of the relationship |
| Tagging | High | High | Modification of the relationship |

[a] How simple is the interaction to understand
[b] Does the interaction allow user specific input
[c] What action on the correspondence does the interaction allow

## 2.6.1 Button Press

The button press interaction is widely used in mapping tools to verify candidate correspondences. The interaction is simple and requires the user to verify a candidate correspondence via a button press. If the user verifies the correspondence, the confidence measure is modified to 1 to signify that the correspondence is valid. If the user rejects correspondence, the confidence measure is modified to 0 to signify that the correspondence is invalid. The difficulty of the task is then just reliant on the question and information being asked. There is no flexibility in the interaction and the user can only select one of the options the button(s) present. PROMPT, see section 2.4.2, and OLA, see section 2.4.5, are examples of mapping tools that use the button press interaction.

## 2.6.2 Construct Lines

The construct lines interaction is also widely used in mapping tools. This interaction is used to verify candidate correspondences and construct new correspondences. The candidate correspondences are presented as correspondences initially displayed with lines between the concepts. The user can remove the line to invalidate a correspondence, i.e. rejects the correspondence. The confidence measure is then modified to 0 to signify that the correspondence is invalid. The user can also append a missed correspondence by joining two concepts with a line. This action generates a new correspondence with the equivalent relation between the two concepts the user has selected with a confidence measure of 1. Certain mapping tools allow the line to be weighted which modifies the confidence measure of the correspondence. Furthermore some of the mapping tools combine the construct lines with the enter formulas interaction, see section 2.6.3, to allow the user to input the relationship of the correspondence. This interaction is not as simple as pressing a button and in particular scanning the global structures may cause issues for ordinary users. The interaction offers slight flexibility by allowing the user to weight the confidence measure. COMA++, see section 2.4.2, and CoGZ, see section 2.5.4, are examples of mapping tools that use the drawing lines interaction.

### 2.6.3 Enter Formulas

The enter formulas interaction is not as commonly used in mapping tools as the previous two interaction types. This interaction allows the users to modify the relationship of the correspondences. Additionally users can also generate new correspondences between concepts with different relationships. The confidence measure will be set to 1 but the mapping tool can uses another interaction to allow the measure to be modified, see section 2.6.2. Furthermore highly specific relations can be generated for the correspondences, such as the uncle relation, if the formulas are provided by the tool. In the author's opinion, this type of interaction would be too complex for ordinary users to understand. The interaction is highly flexible with even certain mapping tools allowing new formulas to be developed and added by the user. BizTalk [BizTalk 2005] and AlViz, see section 2.4.4, are examples of a mapping tool which uses the enter formulas interaction.

### 2.6.4 Data Grid

The data grid interaction has been used in the Webscripter mapping tool, see section 2.5.1. With this interaction users are shielded from the complexities of ontologies by only dealing with their instance information. The interaction requires the user to manage correspondences by constructing tables and placing the data within a table. Correspondences with equivalent relationships are then generated between instances which can then be used to generate correspondences between concepts. In the author's opinion, this kind of interaction will be very tough for ordinary users to grasp as the interaction asks a lot of the user from managing to maintain the data. It would also be a burden for the users as it requires the user to deal with the entire volume of data for mapping.

### 2.6.5 Structure Sentences

The structure sentence interaction has been proposed by the Ontogame tool, see section 2.5.9, but has yet to be implemented or evaluated. The interaction allows users to modify the relationship of the correspondences by selecting a sentence, in Ontogame case a word, which best describes the relationship. This is a very similar to the interaction type of natural language interfaces (NLIs), see section 2.4.8. Ordinary users should be able to use this type of interaction as long as sentences used and the information provided is understood by the users. This type of interaction has no flexibility with the users only allowed to select one of the provided sentences.

### 2.6.6 Tagging

The tagging interaction has not been used by any mapping tool. This type of interaction would allow users to enter their own words for the relationship of a candidate correspondence. Tagging makes the interaction easier for ordinary users as there is an unlimited number of ways to answer and there is no "wrong" choice. However there would be an issue on how to classify the relationship from the tags used. Several popular web sites use tagging. Flickr[12] is a site where users can upload their photographs and add tags to organise the photographs and share them with other users. Delicious[13] is a site where users upload bookmarks for web pages and add tags to organise the web pages and share them with other users.

## 2.7 User Evaluation of Mapping Tools

There has been a distinct lack of user studies carried out on mapping tools to date. The following sections present five published user evaluations of mapping tools. Each of the evaluations only considers mapping with one-to-one correspondences.

### 2.7.1 PROMPT User Study

A user study was conducted on PROMPT [Noy 2002]. The experiment involved the users downloading the tool, reading through the documentation and tutorial example given, completing a set of tasks and emailing the author the results of the tasks. The experiment concentrated on evaluating the correspondence suggestions provided by the tool and did not evaluate the usability, user's performance of generating correspondences and user's experience with the tool. The experiment also only involved four users and the author states "*The number*

[12] http://www.flickr.com/
[13] http://www.delicious.com/

*of users was still too small and the variability in user's expertise with Protégé too large to get meaningful estimates"*. This suggests the number of users is too small to gain any significant conclusions from the results.

### 2.7.2 PROMPT and Chimaera User Study

Lambrix and Edberg [Lambrix 2003] performed a user evaluation of PROMPT and Chimaera for the specific use case of merging ontologies in bioinformatics. The user experiment involved eight users, four with computer science backgrounds and four with biology backgrounds. Participants were given a set number of mapping tasks to perform. A user manual was given to the user and the software help system was available to support the user while they performed the mapping tasks. The experiment was an observational case study and participants were instructed to "think aloud" while an evaluator took notes. Afterwards the users were instructed to complete a questionnaire about their experience. The user interfaces were evaluated using the REAL (Relevance, Efficiency, Attitude and Learnability) approach [Lowgren 1993]. PROMPT outperformed Chimera, however participants found learning how to merge ontologies with each tool difficult. The participants also found it particularly difficult to perform any non-automated procedures in PROMPT, such as creating user-defined merges. The author mentioned they investigated the difference in the user interface evaluation between the two groups of users and found "*there was no significant difference between the results of these two groups"*. This suggests that the results of the user groups were to the same standard. However, in the evaluation provided there was no analysis of the user's performance with the mapping tools. Instead the evaluation just provided analysis on the quality of candidate correspondences made by each tool rather than the correspondences generated by the users with the mapping tools.

### 2.7.3 Visualisation Mapping User Study

Robertson et al. [Robertson 2005] evaluated their novel approach to visualisation of mappings, which used several standard visualisation techniques to improve the interface of mapping tools see section 2.5.2, with a user-study. The Biztalk mapping tool was used as the baseline for the evaluation, as the prototype was built on top of Biztalk. The prototype allowed several of the visualisation techniques to be turned off which allowed for four different versions of their prototype to be tested against BizTalk, to assess the impact of each feature. The user experiment involved eight users, all of which had experience using BizTalk. The experiment was an observational case study and the participants were required to complete a user satisfaction questionnaire and interview after the experiment. The results revealed the new feature set received significantly higher ratings. In this experiment the efficiency of the user performance was evaluated by analysing the mapping task time. However, the effectiveness of the user performance, the quality of the mappings generated, was not evaluated. Also the usability of the mapping tools was measured with a customised questionnaire.

### 2.7.4 Schema Mapper User Study

Schema Mapper [Raghavan 2005] was evaluated against MapForce[14], which is an XML mapping graph-based tool, in a user-study. The user experiment involved nine participants, five of whom had previous experience with XML editing. Each user was given instructions on how to use each tool. The users were also given additional reading material for both tools and asked to explore the tools until they felt comfortable. The experiment was an observational case study. The results revealed that users found it easier to locate nodes with Schema Mapper as there was not as much scrolling or searching with the hyperbolic tree compared to the graph-based interface of MapForce. Also the recommendations from Schema Mapper helped navigation of the mappings. The users found Schema Mapper easier to use as they found confusing the lines that MapForce used to denote mapping across the screen. Finally, users did not find it confusing to view just the local view of the term when mapping in Schema Mapper and actually preferred it and would use this view again for mapping. Once again only the efficiency of the user performance was evaluated by the time taken to perform a mapping task.

### 2.7.5 CoGZ and PROMPT User Study

CoGZ was evaluated against PROMPT in a user-study [Falconer 2009]. The user experiment involved 18 participants, all recruited from the University of Victoria computer science department[15]. The experiment was an observational case study and participants were instructed to "think aloud" while an evaluator took notes. For the experiment each user would attempt to complete 9 mapping tasks, each a different level of task, with both tools. The procedure was the following: first the user received training with the tool, then they practiced using the tool with a different set of ontologies, next they answered the set of mapping tasks for the tool, and finally they completed a SUS questionnaire to measure the user satisfaction. The tools were used in different orders and an interview was conducted with the user after they were finished using both tools. Both the efficiency and effectiveness of the user performance was evaluated for both mapping tools. The results revealed that the CoGZ tool improved the mapping performance of the users and had better user satisfaction response from the users.

---

[14] http://www.altova.com/mapforce.html
[15] http://www.csc.uvic.ca/

### 2.7.6 Summary

Table 2-6 displays a summary of evaluation methods used by the experiments detailed in the previous section.

**Table 2-6:** Summary of evaluation methods in mapping tool user studies

| | PROMPT | PROMPT & Chimaera | Visualisation Mapping | Schema Mapper | CoGZ & PROMPT |
|---|---|---|---|---|---|
| **Background of Participants** | Medical | Medical & Computer Science | Computer Science | Computer Science | Computer Science |
| **Participants' Mapping Experience** | Novice | Novice | Expert | Novice & Expert | Novice & Expert |
| **Number of Users** | 4 | 8 | 8 | 9 | 18 |
| **Experiment Type** | Field | Controlled | Controlled | Controlled | Controlled |
| **Experiment Protocol** | None | Think-aloud | Observational | Observational | Think-aloud |
| **Focus of Experiment** | Tool | Tool | User | Tool | User |
| **Group Compare[a]** | No | Yes | No | No | No |
| **Statistical Tests Used** | No | No | ANOVA | No | t-Test |
| **Questionnaire Used** | None | Yes | Yes | None | SUS |
| **Interviews** | No | No | Yes | Yes | Yes |

[a]The user study compares the mapping results of different groups of users

Only one experiment occurred in the user's work environment. However there was no evaluation on the user performance or feedback from the users using the tool in this experiment. The others occurred in a controlled lab environment where an evaluator observed the participants interacting with the mapping tools. Two of the experiments required the user to think-aloud while performing mapping tasks. Using the think-aloud protocol has several problems such as unease on the part of the participant on having to speak and if there is silence prompting the participant to speak may become overbearing which can affect the participant's performance[16]. In the author's opinion it may be better for controlled lab experiment to be observational and try to replicate the natural setting of the user's work environment. **The majority of the experiments in this section focused on measuring the performance of the mapping tool rather than the user's performance with the mapping tool.** Table 2-7 displays a summary of the areas of user evaluation and metrics, used by the experiments referred to in this section.

[16] http://www.tiresias.org/tools/think_aloud.htm

**Table 2-7:** Summary of Mapping Tool Evaluation Areas

| | PROMPT | PROMPT & Chimaera | Visualisation Mapping | Schema Mapper | CoGZ & PROMPT |
|---|---|---|---|---|---|
| **Efficiency** | No | Questionnaire | Task Time | Task Time | Task Time |
| **Effectiveness** | No | No | No | No | Gold Standard |
| **User Satisfaction** | No | Questionnaire | Questionnaire & Interview | Interview | Questionnaire & Interview |
| **Accessible** | No | No | No | No | No |
| **Convenience** | No | No | No | No | No |
| **Simple**[a] | No | Questionnaire | Interview | Interview | Interview |

[a]Evaluates how easy the goal of generating mapping is to accomplish for users

The majority of the experiments measure the efficiency of users generating mappings with the mapping tool and use the time taken to complete a mapping task as the metric. Only one of the experiments measured the effectiveness of users generating mappings with the mapping tool by evaluating the quality of the mappings generated by the participants against a gold standard[17]. Questionnaires and interviews were used in the experiments to measure the user's satisfaction. Specifically how simple a mapping task was to perform was judged by the responses given to questionnaires and comments given in the interviews. None of the experiments evaluated if the mapping tools were accessible or convenient to use.

User experiments for mapping tools are necessary for the improvement of mapping tools [Bernstein 2007] [Falconer 2009] [Shvaiko 2008]. However other than the few examples in this section, there have been little user evaluation published on ontology mapping tools. Also these user experiments evaluated just the usability of the mapping tools and did not evaluate the accessibility or convenience of using the tools in the users computing environment. Moreover, there is no standard defined benchmark on how to perform user evaluation of mapping tools. A proposed mapping evaluation benchmark STBenchmark [Bogdan 2008] presents a simple usability model to evaluate the usability of mapping tools through the effort the user exerts, through the number of keystrokes and buttons pressed, but this model still lacks necessary feedback from users. For example it is important to measure the cognitive load of the task which will not be captured with the number of keystrokes and buttons pushed. Finally, the majority of the user experiments that have been carried out have been done in a laboratory setting and only one published study has been a field test conducted in the user's computing environment. If mapping usage is to increase user experiments need to be evaluated in their own work environment to gain an understanding of the user behaviour with the tool in a non laboratory setting.

[17] A Gold Standard is the optimum mapping between two ontologies generally constructed by a knowledge engineer and is used to measure the quality of other mappings between the two ontologies

## 2.8 Requirements for Ordinary User

The current mapping process and tools are inadequate for ordinary users. The four actions framework, from Blue Ocean Strategy [Chan Kim 2005], is an analytical tool which can be used to help identify the requirements needed for ordinary users to engage in the mapping process. The framework asks four key questions: which factors that are taken for granted need to be *eliminated*, which factors need to be *reduced* well below the current standard, which factors need to be *raised* well above the current standard, and which factors need to be *created* that have never been offered. Figure 2-24 gives the author's opinion on the requirements needed to support an ordinary user participating in the mapping process in an eliminate-reduce-raise-create (ERRC) grid.

| **Eliminate** | **Reduce** |
|---|---|
| Any Matching API Configuration by User | Complexity of Mapping Process<br>Visual Emphasis on Global View |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | Easy to Understand Interface<br>Familiar Mapping Interaction<br>Evaluation Methods |

***Figure 2-24:*** *Requirements for Mapping Approach and Tool for Ordinary Users*

The matching configuration steps need to be removed as this brings unnecessary complexity to the mapping process for ordinary users. It can be sufficient for the mapping system to judge which algorithm to use based on the domain. The current mapping process is too difficult for ordinary users to engage in and its complexity needs to be reduced. For example the 'decide when to map' step can be removed as a user-specific mapping task and instead be left as a system-specific mapping task. In the majority of mapping tools the interface emphasis is on displaying the global structure of the ontologies and where the correspondences are between the ontologies. It is left to the user to disambiguate the information for each candidate correspondence. In the author's opinion, the interface needs to primarily display the information for each candidate correspondence when suggested and reduce the emphasis of the global structure of the ontologies. Furthermore the interface needs to represent the information in a way that will be easy to understand for ordinary users. Current mapping tools represent the information which requires users to have modelling experience which ordinary users may not have. Finally, evaluation methods need to be created to analyse whether the mapping tools are accessible and convenient to use within the computing environment of ordinary users. In addition, there are no standard methods for analysing the usability of mapping tools. Such standard methods are needed in the evaluation of mapping tools.

## 2.9 Summary

In this chapter, an overview of the state of the art in ontology mapping tools research was presented. Despite significant research contributions on automatic ontology mapping there is still a requirement for a "human in the loop". The standard mapping process is not performed at runtime and requires plenty of user involvement. Traditionally, the mapping process and tools are designed for knowledge engineers or domain experts, who have received substantial training. However the mapping process can be hard even for knowledge engineers.

State of the art mapping tools are still highly complex and predominantly assume the user to be a knowledge engineer or user who has received a great amount of training. There have been few user studies conducted on mapping tools and those which have been carried out have been undertaken with small user groups. In fact, there have been few field experiments attempted with mapping tools. The benefit of field experiments is the ability to observe the natural behaviour of users with the mapping tool. The chapter concluded with several requirements suggested by the author on how the mapping process and tools can be engineered to allow ordinary users to engage in the act of mapping.

In the next chapter, an evaluation is carried out on two alternate approaches to semantic mapping with different mapping process and interfaces. The first uses the traditional method of allowing the user to traverse the list of candidate correspondence themselves with a graph-based visualisation of the ontologies. The second approach restricts the navigation of candidate correspondences to only suggestions from the tool and displays the ontological information for a candidate correspondence with natural language. The mapping performance of three user groups of users with different technical experience is examined to measure the difference in performance between ordinary users and knowledge engineers. The user-study is used to evaluate where the problem areas are in mapping for ordinary users and the relative merits of each technique in solving these issues. The results will be used in the design of an incremental mapping process, where the act of mapping will be done at runtime, and the design of the tag-based mapping tool.

# 3 Mapping Tool User Experiment

This chapter presents the first of three experiments conducted to address the research objectives of this thesis. The goal of this experiment was to address research objective **RO1**: Determine whether ordinary users can actively participate in the construction of semantic mappings to a similar standard as knowledge engineers. The experiment is designed to evaluate whether ordinary users can generate mappings. This was investigated by analysing the mapping interaction results, i.e. the precision of the mapping made and the time taken, generated from two different mapping tools. Three user groups were chosen which each had different technology expertises. This allowed the results of ordinary users to be contrasted with both knowledge engineers and users who have modelling experience. The experiment was carried out in a controlled lab situation. The experiment also aided in addressing research objective **RO2:** Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts. The two mapping tools used in the study have different approaches for mapping which will help in identifying the difficult parts of the mapping processes respectively, for example which parts are the cognitively toughest for users and which are the most time consuming. Using two mapping tools will help to distinguish the relative effectiveness of each approach. The study presented in this chapter was first presented in [Conroy 2007].

The structure of this chapter is as follows: first the design of the study is presented, secondly the goals of the experiment are stated, thirdly the Natural Language (NL) prototype mapping tool which is built to reflect the approach of Natural Language Interface (NLI) authoring tools is described, fourthly the methodology for the experiment is detailed, fifthly the results of the experiment are presented, sixthly the analysis of the experiment is detailed, seventhly the finding of the experiment are described, eighthly the limitations of the experiment are given, ninthly the requirements needed for ordinary users to engage in the mapping process, see section 2.8, are updated based on the results of the evaluation, finally a summary of the chapter is given.

## 3.1 Experiment design

In this study, the mapping interactions and feedback of three user groups are observed, recorded and evaluated. The user groups were:

- **Ontologically aware** users who have experience of working with ontologies. These users are representative of the current users of mapping tools and the participants used in other user-based mapping studies, see section 2.7. Thus this group reflects the skill level of knowledge engineers and was chosen to be used as the control group for the study.
- **Technically aware** users who have database/UML modelling experience but no ontology experience. This group have modelling skills which can be translated into the skills needed for mapping tools. In the author's opinion, this group also represents a high percentage of web users who are the early adopters of new technologies. This group was chosen to examine how significant are the differences in the results between ordinary users and users that have modelling skills. This group was also chosen to be used as a control group for the study.
- **Non-technical** users who have basic computer experience with no database/UML modelling or ontology experience. These users have the lowest level of technical experience and have no modelling skills. This group is taken to be a proxy for the ordinary users.

The study design allowed for an analysis of differences in the mapping performance between the user groups. The purpose of this analysis was to evaluate whether users with the least technical knowledge, non-technical users, can construct mappings to the standard of users who have ontology work experience, the ontologically aware user group. In addition, the analysis also evaluated whether the users level of modelling skills affect their performance with the mapping tools.

Two different mapping tools were used in the experiment, COMA++ and the Natural Language (NL) prototype mapping tool. COMA++ was introduced in section 2.3.2 and was primarily selected as it presents a traditional approach to semantic mapping using a graph-based graphical display where navigation of candidate correspondences is left to the user. The NL prototype mapping tool, see section 3.4, was an initial prototype mapping tool developed by the author to use natural language to display mapping information, and supports a question and answer approach to navigation of candidate correspondences. The NL prototype mapping tool was designed primarily to reflect a current approach of a state of the art ontology authoring tools which aim at getting domain-expert users involved in the process by displaying ontological information using natural language, see section 2.4.8.

The two mapping tools were also selected for other reasons. First they both support user-interaction in different ways. COMA++ first computes the candidate correspondences between ontologies, displays the results visually in a standard graph tree-type display, and then the user interacts to remove invalid correspondences and append any correspondences missed by the algorithm. The NL prototype mapping tool on the other hand produces a list of candidate correspondences with the INRIA Alignment API [Euzenat 2006] that the user then validates by answering the candidate correspondences asked in an iterative loop. Finally the user-interfaces for both tools are distinctly different, allowing the investigation of which type of interface is easier to understand for each user group.

To evaluate if the mapping results differ between the user groups both the efficiency and effectiveness of each tool was measured. The user satisfaction with each tool was also measured to analyse the user perceived quality with each approach. The outline for evaluation is displayed in the table 3-1 below, from section 1.5.2.

**Table 3-1:** Experiment evaluation outline with current experiment (1) highlighted

| **Evaluation Category** | | **Experiment 1** | **Experiment 2** | **Experiment 3** |
|---|---|---|---|---|
| **Efficiency** | | Time Taken (section 3.5.1) | Yes | Yes |
| **Effectiveness** | | Precision & Recall (section 3.5.2) | Yes | Yes |
| **Accessibility & Convenience** | ***Unintrusive*** | Not evaluated in this experiment | Yes | Yes |
| | ***Engaging*** | Not evaluated in this experiment | Yes | Yes |
| **Simple** | | Interview (section 3.5.3) | Yes | Yes |
| **User Satisfaction** | | Interview (section 3.5.3) | Yes | Yes |

In summary, in this study the questions being investigated are as follows:

**Q1:** Can non-technical users complete a mapping to the same level of performance as ontologically aware and technically aware users?

**Q2:** Is a Question & Answer process preferred by non-technical users for mapping navigation than the process of searching tree graph-based display?

**Q3:** Does a natural language representation of the local structure of a concept (e.g. parents, children, properties) make the display of the concept more identifiable to non-technical users than a graph based representation, for mapping correspondences?

## 3.2 Goals

From the questions stated above, several goals were formulated to be evaluated. The study was designed to provide data on each of these goals:

> **G1**: There will be no significant difference in mapping performance between user groups with different technology expertise, derived from **Q1**;
>
> **G2**: Non-technical users will be able to map correspondences quicker using the approach of NL prototype mapping tool than with the approach of COMA++, derived from **Q2**;
>
> **G3**: Non-technical users will perceive the difficulty of mapping to be easier when using the mapping interface of the NL prototype mapping tool than with COMA++ mapping interface, derived from **Q2**;
>
> **G4**: Non-technical users will prefer the mapping interface of the NL prototype mapping tool and find it easier to use than the COMA++ mapping interface, derived from **Q3**;
>
> **G5**: Non-technical users will be able to understand a concept's structure more accurately with the mapping interface of the NL prototype mapping tool than with mapping interface of COMA++, derived from **Q3**.

## 3.3 The NL Prototype Mapping Tool

A natural language mapping tool was chosen in the experiment to reflect a current approach of ontology authoring, see section 2.4.8. The NL prototype mapping tool was developed as there is no publically available semantic mapping tool which displays the ontological information in natural language. The NL prototype mapping tool is built using a client-server web server approach which allowed the mapping interface to be displayed in a web browser. The NL prototype mapping tool was developed to be a questionnaire in the web browser and was deployed on a Tomcat web server [Chopra 2004]. The source code for the NL prototype mapping tool is provided in the accompanying DVD media under Experimental Datasets – Experiment 1.

The NL prototype mapping tool uses a linear questionnaire approach to the mapping process by asking the user a series of sequential candidate correspondences questions. A series of fixed question templates using the words ‘similar’ and ‘correspond’ were used to represent the candidate correspondence questions. Different templates were used to help the questions from becoming too repetitive. However the results of the experiment revealed that the participants were answering differently based on the questions used, see section 3.5.4. Figure 3-1 gives an example of how the candidate correspondence question is presented to the user.


***Figure 3-1:*** *The NL prototype mapping tool user mapping interface*

The concepts are represented in a type of natural language which displays the parents, properties and children of the concept. This information is displayed in bullet point format. The parents of the concept are chained all the way back to the top level concept `owl:Thing`, while the child concepts are grouped together, and the properties of the concept are displayed in a fixed template depending on its cardinality and if it is restricted to a concept. The template for the natural language statements are the following:

- *Parent*: ‘<concept name> is a <parent name>’
- *Children*: ‘<child name> is a type of <concept name>’
- *Property*: ‘<concept name> may have a <property name> which is a type of <restricted concept>’

A clear distinction is kept between the concepts of the candidate correspondence as each are displayed in a different colour and positioned differently. The source concept is coloured in blue and positioned to the left while the target is coloured in red and positioned to the right. They are positioned in the same way as current mapping tools that separate the source information to be on the left and the target information to be on the right. The different colours were used to help people identify if there was a difference between the concepts shown. However in the experiment, one participant who was colour-blind mentioned they did not see any difference in colour between the information but this did not hamper their decision making. For each candidate correspondence a list of alternative candidate correspondences for the two concepts is displayed to the user. This gives the user additional information in making their decision by giving the candidate correspondence some context. Figure 3-2 below shows the overall interface for the NL prototype mapping tool where the alternate candidate correspondences are displayed to the right of the candidate correspondence asked and a list of all the correspondences is displayed below.

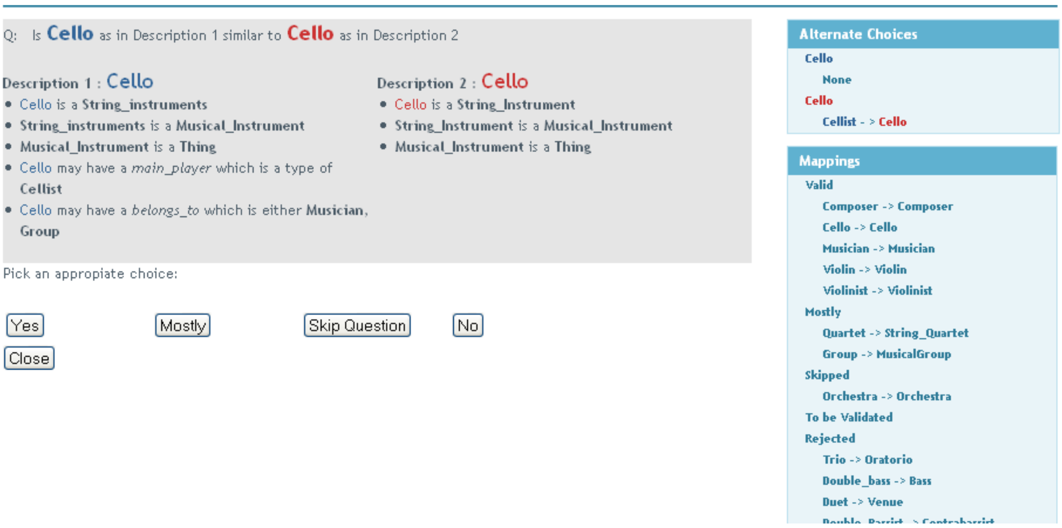



***Figure 3-2:*** *The NL prototype mapping tool overall interface*

The list of candidate correspondences is generated by the INRIA alignment API [Euzenat 2006] using the `NameAndPropertyAlignment` algorithm. This algorithm was chosen as during the trial of the NL prototype mapping tool this alignment generated the most candidate correspondences, in comparison to the other algorithms, for the ontologies used in the experiment. The ontologies used are provided in the accompanying DVD media under Experimental Datasets - Experiment 1. The candidate correspondences are categorised within one of the following groups: **valid/mostly/skipped/rejected** based on the user choice. Mostly is used to express the relationship of the correspondences as more general. All the buttons are disabled for four seconds after a candidate correspondence has been displayed, this is to give the user time to think and look in detail at the candidate correspondence rather than just clicking without consideration one button all the time. However whether the user takes this time to think and consider the candidate correspondence is entirely dependent on them. In addition due to the variety of the information being displayed the physical location of the buttons will change, again ensuring that the user cannot simply click in the same spot continuously. This is potentially very annoying for the user and could reduce the usability of the tool.

## 3.4 Method

This section gives the details of how the experiment was carried out. Firstly the details of the participants in the experiment are given. Next the materials used in the experiment are described. Afterwards the procedure of the experiment is detailed. Then a summary is given of the data collected in the experiment. Finally the analysis methods used in the evaluation are explained.

### 3.4.1 Participants

24 participants were recruited for the study. The participants were segregated into three different groups (8 per group) based on their technology expertise: *Ontologically aware* users who have ontology work experience. This group is labelled in diagrams as "onto aware". *Technically aware* users who have database/UML modelling experience but no ontology experience. This group is labelled in diagrams as "tech aware". *Non-Technical* users who have basic computer experience with no database/UML modelling and ontology experience. This group is labelled in the diagrams as "non-tech". Participants were recruited via emails to candidates, from both within the college and outside it. Each of the individuals were identified based on their technology experience, the ontologically aware group were made up of researchers, both PhD students and post-docs, within the KDEG research group[18]. The technically aware group was made up from undergraduate students and people outside of the college who have either completed or were currently doing a computer science degree. Each of these people had no ontological experience. The non-technical aware group were made up of staff of the college and people from outside of the college. All of these people had no computer science background and only had basic computing experience.

[18] Knowledge and Data Engineering Group in Trinity College Dublin, http://kdeg.cs.tcd.ie/

### 3.4.2 Materials

The participants in the study were required to use each mapping tool, COMA++ and the NL prototype mapping tool, and give their impressions on the tools. Since the mapping tools were used one after another, two different mapping domains were needed so the participants would not get familiar with the domain. This helped to mitigate the learning effect between tasks and mapping tools. The two domains chosen were **university** [Iswc 2007][Swrc 2007] and **music** [Kanzaki 2007][Maciej 2007] which are completely disjoint to each other and had no overlap of concepts. The ontologies are also provided in the accompanying DVD media under Experimental Datasets - Experiment 1. The university domain was chosen as the majority of the concepts would be familiar to the majority of the participants, and at least it would be understandable to all the participants. While the music domain was chosen as it would also be familiar to the participants but also had certain concepts which might not be familiar to the participants, for example concepts such as 'Contrabassist' and 'Double Bassist'.

A gold standard was produced by the author for each domain, with input from a domain expert for the music domain. The gold standard for each domain was verified by a knowledge engineer from the KDEG research group, who had three years ontology modelling experience. The gold standard is provided in the accompanying DVD media under Experimental Datasets – Experiment 1. The gold standard had 25 valid concept correspondences for the university domain and 24 valid concept correspondences for the music domain. For the NL prototype mapping tool a set of 33 candidate correspondences were generated (15 were valid) by the Align API for each domain. While for COMA++ the results varied with 19 candidate correspondences (12 were valid) for the university domain being generated, while 10 candidate correspondences (7 were valid) were generated for the music domain. Note that although it is possible to compare the mapping results between the mapping tools, if the candidate correspondences used are the same for each, since each tool generated different candidate correspondences in this experiment it made it impracticable to compare the results between the tools. Instead the focus was on contrasting the approaches of each tool.

### 3.4.3 Procedure

The study was conducted in a controlled lab experiment where participants were required to use both mapping tools one after another with different domains for each tool[19]. This allowed the experiment to be done in four different ways, see Table 3-2, and each order was done twice for each user group.

[19] The experiment was carried out before an ethical approval process had been introduced to the school

**Table 3-2:** Participant experimentation order, -- indicates there is no value in the cell.

| | COMA++ | | NL prototype mapping tool | |
|---|---|---|---|---|
| **Order** | *Music* | *University* | *Music* | *University* |
| 1 | First | -- | -- | Second |
| 2 | -- | First | Second | -- |
| 3 | Second | -- | -- | First |
| 4 | -- | Second | First | -- |

When using the NL prototype mapping tool the participants were required to validate the 33 candidate correspondences in a question and answer system. After validating all the candidate correspondences the participants were asked if they wished to make any changes to their answers. They were finished when they were happy with the answers they had provided. When using the COMA++ mapping tool the participants were required to validate the candidate correspondences provided and then to locate any missed correspondences. The participants were considered to be finished when they were happy with the answers they had provided.

The first part of the experiment began with participants receiving training with the first assigned mapping tool. They were instructed briefly about the ontologies, how to make the mapping, and the tool's user interface (using the alternate domain). Finally the ontologies and candidate correspondences were loaded and the participant was asked to carry out the task of mapping with the tool.

The experiment was observational, where the author observed the participant's actions while they were mapping correspondences. The participants were encouraged to mention any thoughts they had during the completion of the tasks. Once the tasks were completed the same procedure was followed for the second tool (including training). The study concluded with a short interview as described in section 3.4.4.

### 3.4.4 Data Collection

The experiment was an observational study which allowed the participant's behaviour with the mapping tools to be observed. During the course of the study participants were encouraged to mention any thoughts they had and notes of these thoughts were taken during the sessions. The mapping results from the participants were saved for both tools during each session and used for later inspection. The entire set of anonymised data records are provided in the accompanying DVD media under Experimental Datasets - Experiment 1. All interviews were transcribed for later analysis. The list of questions asked in the interview is given in Appendix B. Examples of the questions asked during the interviews were:

- Which display was easier to understand COMA++ or the NL prototype mapping tool?
- Which tool had the better interface for the mapping correspondences?
- Which tool was the most intuitive?
- How difficult was the task of making the mapping with each tool?

### 3.4.5 Analysis method

The ontology aware group are the control group for the experiment as they reflect the technical experience of current users of mapping tools. The technically aware group reflect users with modelling skills needed for mapping tools and are also used as a control group. The quantitative data analysed for each mapping tools were the time taken to map correspondences and mapping precision, used to measure the efficiency and effectiveness respectively. Note that there were two different sets of candidate correspondences for each tool which made comparing the results between tools impracticable. An unpaired two-tailed t-Test was used to compare both the time taken for mapping correspondences and mapping precision between user groups. Any significant result from these tests will be highlighted while the full results can be found in Appendix A. The qualitative data helped in revealing why one tool was preferred over the other.

## 3.5 Results

This section outlines the results from the experiment. The evaluation is split into the three areas, displayed in table 3-1, as follows: firstly the efficiency of the two mapping tools is measured, next the effectiveness of both mapping tools is evaluated, and finally the user satisfaction with each mapping tool is presented. Table 3-3 below displays a snapshot of the experiment results. Furthermore, the entire set of anonymised data records are provided in the accompanying DVD media under Experimental Datasets - Experiment 1.

**Table 3-3:** Overall mapping results for the user groups.

| | | COMA++ | | | NL prototype mapping tool | | | |
|---|---|---|---|---|---|---|---|---|
| | **Experiment Order** [a] | **Overall Time** [b] **(MM::SS)** | **Correspondences Made** [c] | **Correspondences Correct** [d] | **Overall Time** [b] **(MM:SS)** | **Correspondences Done** [e] | **Correspondences Made** [c] | **Correspondences Correct** [d] |
| **Non-tech 1** | 1 | 5:11 | 8 | 8 | 8:56 | 33 | 17 | 11 |
| **Non-tech 2** | 1 | 8:59 | 18 | 7 | 8:04 | 33 | 18 | 11 |
| **Non-tech 3** | 2 | 10:25 | 25 | 21 | 5:32 | 33 | 22 | 14 |
| **Non-tech 4** | 2 | 11:52 | 19 | 19 | 15:16 | 33 | 23 | 14 |
| **Non-tech 5** | 3 | 9:55 | 14 | 8 | 12:42 | 33 | 22 | 12 |
| **Non-tech 6** | 3 | 9:19 | 11 | 8 | 16:35 | 33 | 18 | 13 |
| **Non-tech 7** | 4 | 16:45 | 24 | 22 | 22:49 | 33 | 17 | 12 |
| **Non-tech 8** | 4 | 7:02 | 22 | 18 | 15:36 | 33 | 20 | 15 |
| **Non-tech group** | | **9:56** | **18** | **14** | **13:16** | **33** | **20** | **13** |
| | | | | | | | | |
| **Onto aware 1** | 1 | 7:03 | 9 | 9 | 14:19 | 33 | 20 | 14 |
| **Onto aware 2** | 1 | 9:10 | 8 | 8 | 19:13 | 33 | 15 | 15 |
| **Onto aware 3** | 2 | 7:30 | 23 | 21 | 8:41 | 33 | 18 | 14 |
| **Onto aware 4** | 2 | 11:11 | 16 | 15 | 15:00 | 33 | 26 | 12 |
| **Onto aware 5** | 3 | 8:23 | 11 | 11 | 13:42 | 33 | 16 | 15 |
| **Onto aware 6** | 3 | 6:00 | 3 | 3 | 9:57 | 33 | 19 | 15 |
| **Onto aware 7** | 4 | 8:04 | 17 | 16 | 15:10 | 33 | 20 | 14 |
| **Onto aware 8** | 4 | 7:56 | 22 | 21 | 7:06 | 33 | 18 | 14 |
| **Onto aware group** | | **8:10** | **14** | **13** | **12:54** | **33** | **19** | **14** |
| | | | | | | | | |
| **Tech aware 1** | 1 | 11:56 | 10 | 10 | 9:45 | 33 | 16 | 15 |
| **Tech aware 2** | 1 | 8:08 | 13 | 12 | 7:23 | 33 | 15 | 15 |
| **Tech aware 3** | 2 | 3:18 | 13 | 12 | 13:55 | 33 | 13 | 12 |
| **Tech aware 4** | 2 | 10:51 | 23 | 21 | 14:14 | 33 | 12 | 12 |
| **Tech aware 5** | 3 | 8:25 | 7 | 7 | 19:06 | 33 | 24 | 14 |
| **Tech aware 6** | 3 | 6:35 | 14 | 8 | 7:34 | 33 | 23 | 13 |
| **Tech aware 7** | 4 | 15:19 | 24 | 17 | 16:57 | 33 | 26 | 15 |
| **Tech aware 8** | 4 | 8:12 | 20 | 20 | 11:11 | 33 | 16 | 14 |
| **Tech aware group** | | **9:06** | **16** | **13** | **12:54** | **33** | **18** | **14** |

[a] The experiment order refers to the order the user used the mapping tools and with which domain
[b] The overall time is the total time taken by the user using each mapping tool
[c] Correspondences made are the number of correspondence each user said were valid
[d] Correspondences correct is the number of correspondences made which are valid
[e] Correspondences done is number of correspondences suggested by the tool that the user validated

### 3.5.1 Efficiency

The metric taken to measure the efficiency of each mapping tool is the time taken to complete mapping a correspondence, see Table 3-1. This is the commonly used metric to evaluate the efficiency of mapping tools, see section 2.7. Figure 3-3 below gives the results for the mean time taken for mapping correspondences (41 seconds with COMA++ and 23 seconds with the NL prototype mapping tool). It suggests that for each user group it takes longer to complete mapping correspondences with COMA++. Note that the navigation time, i.e. the time taken to scan the graph, is also included in the COMA++ time. Although the results seem to indicate a difference between the ontologically aware and non-technical user groups, a non-paired two-tailed t-Test revealed that there was no statistically significant difference in the time taken for each user group, see appendix A.

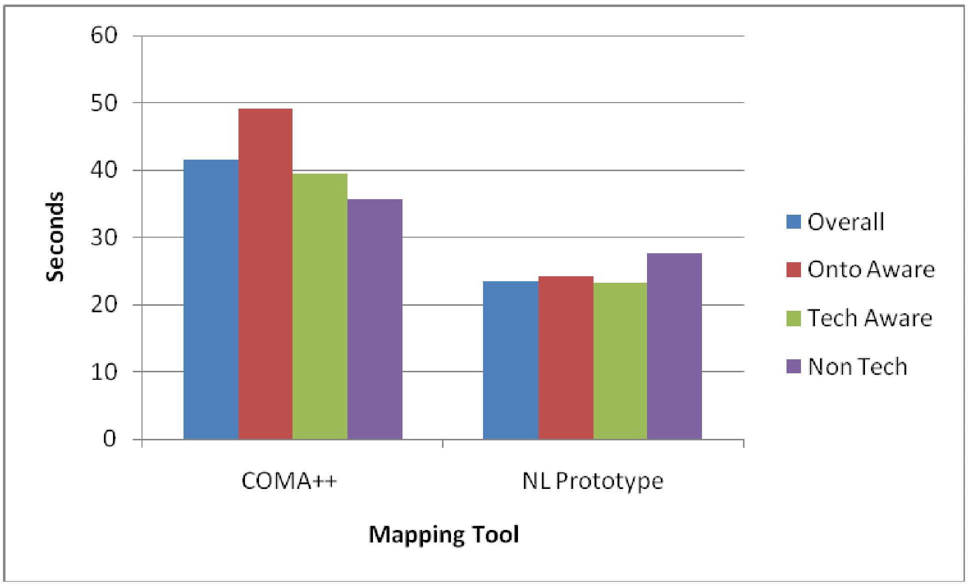



***Figure 3-3:*** *Mean Time for Mapping Tasks for User Groups*

The time taken in each domain was also investigated. The results for the NL prototype mapping tool showed very little change with the maximum difference being 5 seconds between domains for the non-technical user group, the music domain seeing the increased time. However with COMA++ the differences were large, as shown in the Figure 3-4. In particular the difference for the ontologically aware group is over 40 seconds.

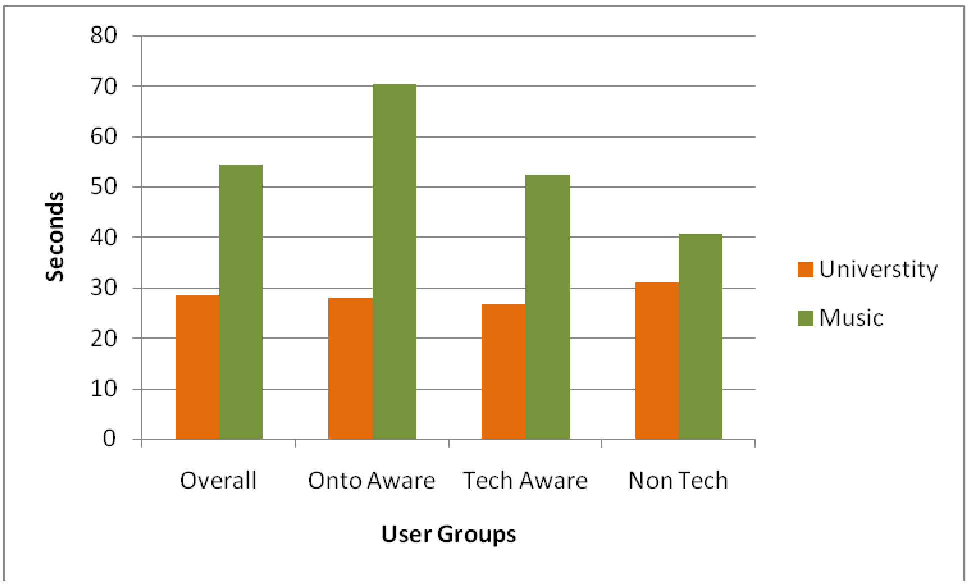



***Figure 3-4:*** *Time taken for mapping correspondences with Coma++*

### 3.5.2 Effectiveness

The metrics used to measure the effectiveness of each mapping tool is the mapping precision and recall. Precision measures how many of the correspondences completed are correct and is commonly used as a metric to evaluate the effectiveness of mapping tools, see section 2.7. Recall measures how many correspondences were discovered out of the total number of correspondence in the gold standard. It is worth noting that COMA++ allows users to append any missed correspondences while the NL prototype mapping tool does not. This should give COMA++ an unfair advantage as the NL prototype mapping tool results are dependent on the matching algorithm. However, the COMA++ system generated less candidate correspondences than the NL prototype mapping tool which may impact this advantage. The recall results for COMA++ should reveal how adapt each user group is at discovering missed correspondences.

Figure 3-5 below displays the mapping precision for each user group. It would not be appropriate to compare the results between mapping tools compared as there were different candidate correspondences generated by each tool. The results are high with each user group scoring above 70% with each mapping tool (88% with COMA++ and 79% with the NL prototype mapping tool). This is quite comparative with the only other published user studies which analysed the precision result [Falconer 2009]. Their evaluations revealed that CoGZ achieved a precision of 1.69 (84.5%) and PROMPT a precision of 1.58 (79%). The results of a unpaired two-tailed t-test indicated that there was a statistically significant difference between the ontologically aware user group and the non-technical user group for both the COMA++

mapping tool (t = 2.4211, p = 0.04 < 0.05) and the NL prototype mapping tool (t = 3.6641, p = 0.01 < 0.05), see Appendix A.

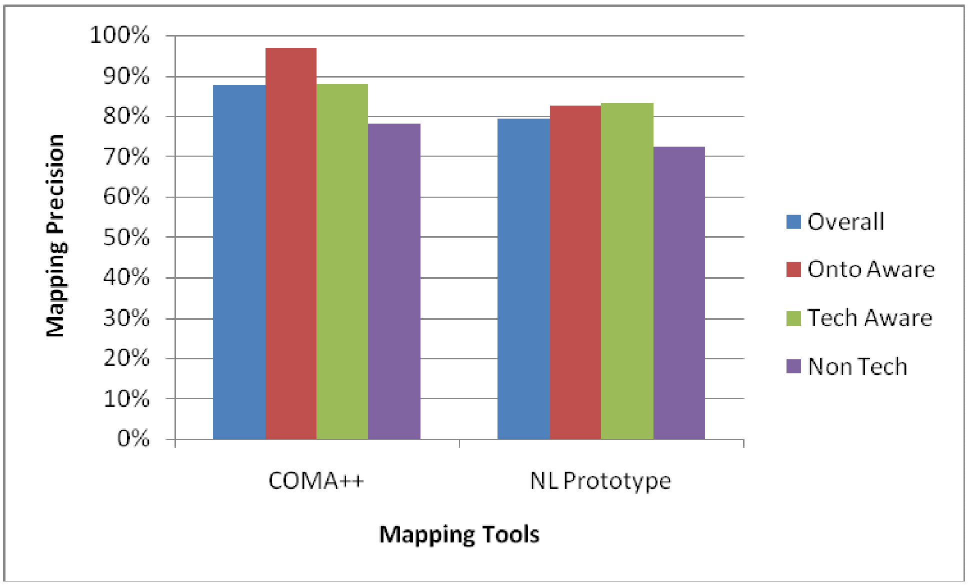



***Figure 3-5:*** *Mean mapping precision for each user group*

Figure 3-6 below displays the differences in mapping precision between user groups using the COMA++ mapping tool for the two mapping domains (University and Music). The mapping precision in the university domain is similar for each user group. However, the mapping precision drops significantly in the music domain for the non-technical user group.

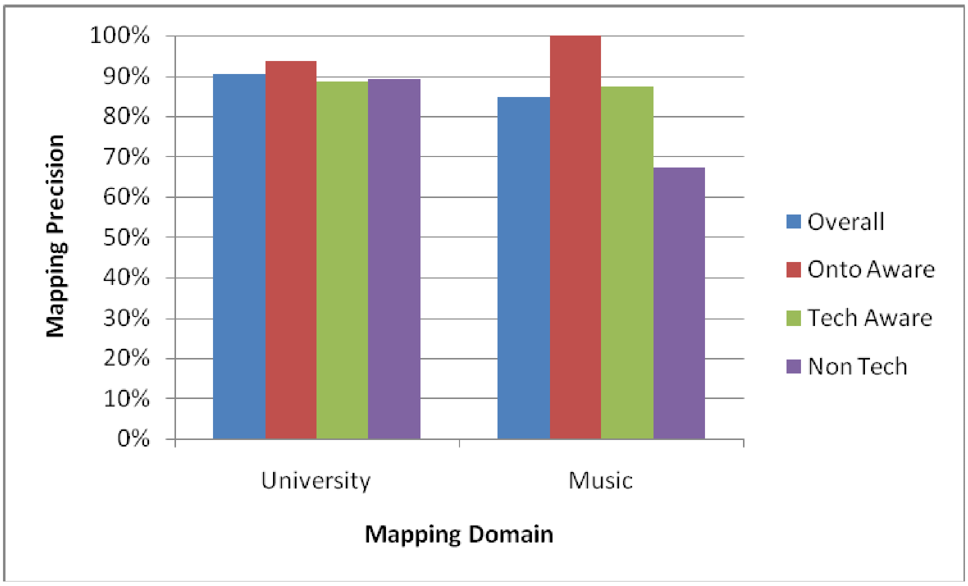



***Figure 3-6:*** *Mean mapping precision with COMA++ mapping tool*

Figure 3-7 below displays the mapping recall results for each user group. The results reveal that both mapping tools had an overall mean of approximately 51%. Furthermore the user groups have similar results for each tool. The recall result for the NL prototype mapping tool was 51%. This is only 6% less than the number of valid correspondence in the initial candidate correspondences list (57%), see section 3.4.2.

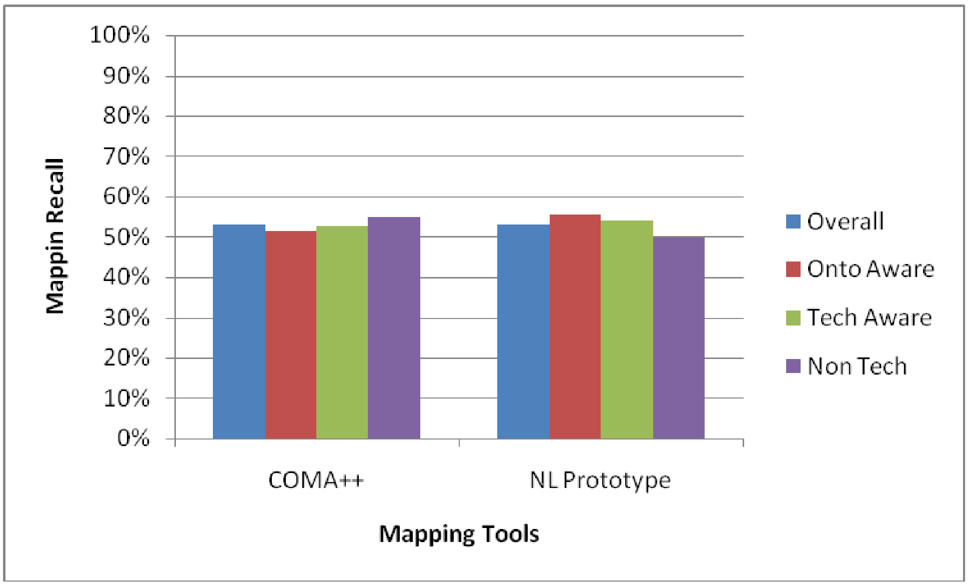



***Figure 3-7:*** *Mean Mapping recall for each user group*

The recall results for the COMA++ mapping tool revealed that the users found an extra 10% of the valid correspondences as only 39% valid correspondences were in the initial candidate correspondences list, see section 3.4.2. Figure 3-8 shows the difference mapping recall between each domain for the COMA++ mapping tool. These results reveal users had trouble finding additional correspondences in the music domain which is not surprising as that domain has concepts which were unfamiliar to the participants, see section 3.4.2. Indeed the number of correspondences found with the COMA++ mapping tool is very low and suggests users find it hard to find correspondences when they are not suggested especially if they are non-obvious.

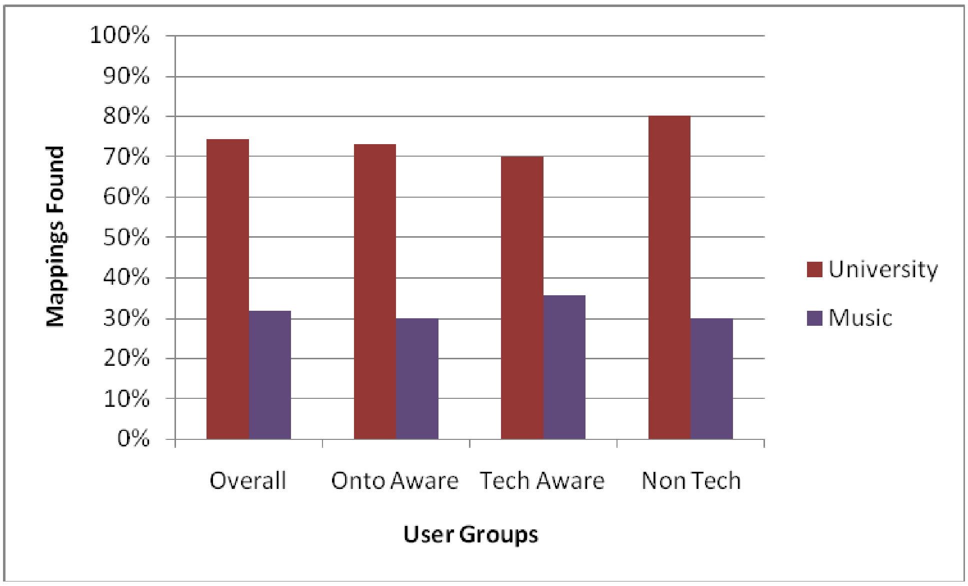



***Figure 3-8:*** *Mean mapping recall with COMA++ mapping tool*

### 3.5.3 User satisfaction

Interviews were carried out with the participants after they had used both mapping tools. The interviews focused on evaluating the participant's satisfaction with both mapping tools. In particular the questions asked were designed to address several of the experimental goals, see Appendix B for the questionnaire. The rationale for the design of the questionnaire is described in the following paragraphs.

To evaluate whether users perceived the difficultly of mapping correspondences to be easier using the approach of the NL prototype mapping tool rather than the approach of COMA++ the participants were asked: *How difficult was the task of making the mapping with each tool?* The majority of the responses from participants indicated that they found the NL prototype mapping tool straightforward to use but with COMA++ it required more thought and it was hard to scan for potential correspondences. Table 3-4 highlights some of the comments from the participants. In addition, 84% of the participants responded that the NL prototype mapping tool was easier to use than COMA++.

**Table 3-4:** Comments from participants on difficulty of mapping with tools

| **Users** | **Comments** |
|---|---|
| Non-tech | *The NL prototype mapping tool was easy as I just have to read and go through questions while COMA++ was difficult as there were too many things to scroll through* |
| Onto Aware | *COMA++ is difficult to browse while with the NL prototype mapping tool it is easier- I just have to say yes or no* |
| Tech Aware | *The NL prototype mapping tool was the easier one with simple relation on left to right while with COMA++ there are too many possible relations, nothing as defined* |

To measure whether users found it easier to use the mapping interface of the NL mapping prototype tool than the mapping interface of COMA++ the participants were asked: *Which tool had the better interface for representing the tasks?* 79% of the participants said they preferred the NL prototype mapping tool interface instead of the COMA++ interface. Table 3-5 highlights some of the responses from participants. Furthermore, when the participants were asked to choose which mapping tool they preferred 75% chose the NL prototype mapping tool while only 17% chose COMA++.

**Table 3-5:** Comments from participants on interfaces of mapping tools

| Users | Comments |
|---|---|
| Non-tech | *The NL prototype mapping tool is much cleaner, while COMA++ doesn't give the same level of detail and scrolling is hard* |
| Onto Aware | *with the NL prototype mapping tool [I am] not getting data overload, it is just one question after another* |
| Tech Aware | *all the information is right in front of you with the NL prototype mapping tool, there is not as much clutter as with COMA++* |

To evaluate if the users understood the concept structure better with the NL prototype mapping tool interface rather than the interface of COMA++ participants were asked in the interview: *Which display was easier to understand COMA++ or the NL prototype mapping tool?* 84% of the participants responded that they found the NL prototype mapping tool easier. Table 3-6 highlights some of the responses from participants.

**Table 3-6:** Comments from participants on difficulty of mapping with tools

| Users | Comments |
|---|---|
| Non-tech | *the task is just more straightforward, I have never used a system like COMA++* |
| Onto Aware | *the visual cues of the NL prototype mapping tool are nice and the task is defined better* |
| Tech Aware | *The NL prototype mapping tool gave better description and information for the task* |

## 3.5.4 Observations

This section gives the author’s observations of the participants’ mapping actions made during the experiment. During the study it was observed that some non-ontologically aware participants were just label matching with COMA++ rather than looking at the structure of the concepts. This observation came from noticing that some participants would map a concept on the left hand side of the graph to multiple concept/properties on the right hand side as long as they had the same label. For example the concept ‘Violin’ would have a user mapping to the property ‘plays’ in ‘Violist’ as it had ‘Violin’ beside the property. In the interviews, when questioned about this, these participants responded that they were just label matching.

Also it was observed that during the course of the experiment users were answering differently based on the template of the mapping candidate correspondence question being asked, i.e. there were different templates of questions using the words ‘corresponds’ or ‘similar’, see section 3.3. When asked in the interviews the users confirmed that they would answer differently based on the use of ‘similar’ or ‘corresponds’ in the question when in fact they were intended to be interpreted as synonyms. For each incorrectly classified correspondence generated by each user, they were asked in the interviews if they would have answered correctly if 'equivalent' had been used instead. The majority replied yes they would have answered correctly however there were still were a few who said they would have answered the same. Some of the comments were *“made different choices based on correspond and similar would have chosen right if equivalent”* by a technically aware user, *“I didn't make different choice based on similar/correspond”* by an ontologically aware user, and *“was using the similar as not exactly the same but similar, would have chose differently with equivalent”* by a non-technical user.

## 3.6 Analysis with respect to Goals

This section gives a summary of the analysis of the experiment. Table 3-7 displays the evaluation of the experiment goals.

**Table 3-7:** Summary of the Analysis of Goals

| **Goals** | | **Suggests support for** | **Evidence** |
|---|---|---|---|
| **G1** | There will be no significant difference in mapping performance between user groups with different technology expertise | No | Section 3.5.1 & 3.5.2 |
| **G2** | Users will be able to map correspondences quicker using the approach of NL prototype mapping tool than with the approach of COMA++ | Yes | Figure 3-3 |
| **G3** | Users will perceive the difficulty of mapping correspondences to be easier when using the mapping interface of the NL prototype mapping tool than with COMA++ mapping interface | Yes | Section 3.5.3 |
| **G4** | Users will prefer the mapping interface of the NL prototype mapping tool and find it easier to use than the COMA++ mapping interface | Yes | Section 3.5.3 |
| **G5** | Users will be able to understand a concept's structure more accurately with the mapping interface of the NL prototype mapping tool than with the mapping interface of COMA++ | Yes | Section 3.5.3 & 3.5.4 |

The first question to be investigated in the experiment was *whether non-technical users could complete a mapping to the same level of performance, i.e. mapping precision and time taken, as ontologically aware and technically aware users*. The first goal **G1** was designed to address this question. Analysis of the mapping results of each user group showed that there was a statistically significant difference between the ontologically aware group and the non-technical user group for the precision of mappings generated with both mapping tools. This finding rejected the goal. Examining the mapping results revealed that this finding can be attributed to the correspondences generated in the music domain. This is hardly surprising as the music domain had many terms that the non-technical users would not know and might be unable to understand, see section 3.4.2, and the number of candidate correspondences was very low, requiring users to search for correspondences, see section 3.5.2. In addition participants made different decisions based on how the question was phrased, see section 3.5.4. This may also explain why there was a difference in performance between the ontologically aware user group, who understood that there was no difference between these constructs, and the non-technical user group when using the NL prototype mapping tool.

The second question of this experiment aimed at evaluating *whether a Question & Answer process is preferred by non-technical users for mapping navigation than the process of searching tree graph-based display*. The second and third goals (**G2** and **G3**) addressed this question. In the evaluation it was revealed that mapping correspondences took longer to complete with COMA++, with an overall mean of 41 seconds, than with the NL prototype mapping tool, with an overall mean of 23 seconds, see section 3.5.1. The feedback from the participants also mentioned that scanning the graph for correspondences was hard. 84% of the participants found the NL prototype mapping tool to be easier than COMA++ for generating correspondences. A major reason given was that it was difficult to browse the correspondences with COMA++. In particular, general feedback from participants mentioned there were too many things to scroll through with COMA++ and no designated start and end point making it hard to understand.

The final question of this experiment examined *does a natural language representation of the local structure of the concept make the display of the concept more identifiable to non-technical users than a graph based representation*. The fourth and fifth goals (**G4** and **G5**) addressed this question. 79% of the participants preferred the NL prototype mapping tool interface over COMA++ interface. Some of the feedback given mentioned that COMA++ did not give the same level of detail for the task and all the information was "right in front of you" with the NL prototype mapping tool interface while with COMA++ the information required was cluttered.

84% of the participants found they could understand the task better with the NL prototype mapping tool. Some of the comments mentioned that there was a better description of the task, the task was defined better and the task was more straightforward with the NL prototype mapping tool. It was also observed during the course of the experiment that some participants were just label matching with COMA++, see section 3.5.4. Label matching is just mapping based on the name of the concept and not the structure. This can lead to several incorrectly generated correspondences. This indicates that ordinary users have a hard time understanding graph-type displays due to insufficient modelling knowledge

## 3.7 Overall Findings with respect to RO2

Although **RO2** is not specifically addressed in this experiment there were several conclusions and issues that emerged which impacted the design and implementation of the mapping tool presented in this thesis. The research objective **RO2** is: Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts. This analysis is based on the results derived from the study.

### 3.7.1 Natural Language superior way of representing ontological information to ordinary users

As shown in the interview results given and observations from the study some participants were found to be only label matching with the graph-type COMA++ mapping tool. While in the interviews, participants said with the NL prototype mapping tool they read the information provided before answering with comments like '*I preferred the NL prototype mapping tool because it gives more information which helps to understand what is being asked of you'* from a non-technical user and '*with the NL prototype mapping tool there is better description for task much more information'* from a technically aware user. Also the majority of participants at 75% said they preferred the NL prototype mapping tool.

### 3.7.2 A Question & Answer approach aids in mapping navigation

On average mapping correspondences took longer to complete with COMA++, with an overall mean of 41 seconds, than with the NL prototype mapping tool, with an overall mean of 23 seconds, see section 3.5.1. This time includes navigation between candidate correspondences and suggests that using the question and answer approach decrease the time spent navigating through candidate correspondences. In addition in the interviews participants mentioned some other improvement of the Question & Answer approach. First, it allows users to know when to start and when to finish unlike scanning. A non-technical user commented that '*using COMA++ is harder as you have to scan'.* Secondly, question answering is a familiar process for users unlike scanning. A different non-technical user commented that '*I have never used a system like COMA++'.* Finally, navigation is much easier for users using the question and answer approach. A comment from a technically aware user mentioned that '*scrolling up and down was a bit of a pain with COMA++*'.

### 3.7.3 Confusion with certain ontological terms and candidate correspondences asked

Throughout the study each participant was first bemused and then baffled when inappropriate candidate correspondences were displayed, for example 'Violin -> Double Bassist'. The concept descriptions also had certain ontological terms which confused participants, in particular the word `owl:Thing`, which lead them to answer incorrectly. For example, a non-technical user commented: "*If each concept had Thing in its description I would answer mostly the same even if the descriptions were different, for example Tempo - > Composer*" and another non-technical user said *"I was unsure of some of the terms being used"*. These issues could make users lose confidence with the mapping tool. To reduce the confusion, ontology specific terms will need to be eliminated from the descriptions. In addition, showing the benefits of the mappings to the users could keep them engaged in the semantic mapping process even when inappropriate candidate correspondences are displayed.

### 3.7.4 Different question templates led to different answers

During the course of the study it was observed that participants were answering differently based on the template of the candidate correspondence question being asked, i.e. if "corresponds to" or "similar to" were used, see section 3.6 for further details. It was also observed frequently during the study that when participants went to choose an answer for the task a lot of the time the participant would then go back and read the question again. This may indicate that the participants wanted to respond with an answer that was not available. Using a tagging approach, see section 2.6.6, would allow users to answer in their own words and may also remove the burden on choosing the 'correct answer'. This approach would also remove the issue of users answering differently based on the template of question being asked, as the question would just ask them what the relation is.

### 3.7.5 Users unaware of any added aid for mapping correspondences

Despite the NL prototype mapping tool interface having a list of the candidate correspondences that the user could navigate as well as an alternative list for each candidate correspondence asked, see section 3.3, the majority of participants in the study mention they did not look at either list and in fact forgot about it (only 3 said they looked at the list). This was despite the list being highlighted to each participant in the training section of the experiment. Some of the comments from the participants were "*I didn't pay attention to it*" from a non-technical user and" *No I did not notice it and I could not understand it*" from another non-technical user. It was hoped that it would aid the participants in making their decision by giving the candidate correspondence some context but instead the participants were more focussed in the candidate correspondences being asked rather than the added support at the side of the screen with a comment given "*I did not look at it much if it was more in centre and a bit bigger it might have helped*" by ontologically aware participant. This is both positive and negative, with the participants actually reading the information of the task but being unaware of any added support. This indicates that ordinary users were just focussed on the task at hand. In the author's opinion, this suggests that a mapping system which would just ask candidate correspondences to the user without any navigation may satisfy the needs of ordinary users.

## 3.8 Updated Requirements for Ordinary Users

In this section the mapping requirements needed for ordinary users, detailed in section 2.8, are updated based on the results of the evaluation. Figure 3-9 displays the updated list of requirements with any modifications made highlighted in italics.

| **Eliminate** | **Reduce** |
|---|---|
| Any Matching API Configuration by User<br>*Visual Emphasis on Global View* | Complexity of Mapping Process |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | *Iterative Mapping Process*<br>*Familiar & Rich Mapping Interaction*<br>Easy to Understand Interface<br>Evaluation Methods |

***Figure 3-9:*** *Updated Mapping Requirements for Ordinary Users*

When ordinary users were using COMA++ some of the users were revealed to be just label matching, see section 3.5.4. These users found it difficult to disambiguate the information for the candidate correspondences from the global graph. Furthermore ordinary users found it tough to find any missing correspondences with COMA++, see section 3.5.3. In the author's opinion, this indicates that finding missing correspondences is too tough a task for ordinary users to achieve. In fact, as ordinary users were using the NL prototype mapping tool they were fully concentrated on the task of validating candidate correspondences and were unaware of any of the added support of the tool, see section 3.7.5. This leads to the requirement that this task be removed from the process to reduce its complexity. Ordinary users will then just be needed to validate candidate correspondences suggested by the mapping system. Furthermore the mapping system itself can guide the user through the process of mapping candidate correspondences. The Question & Answer approach of NL prototype mapping tool was shown to be beneficial in navigating ordinary users through candidate correspondences, see section 3.7.2. This removes the need for the global structure of the ontologies to be shown to ordinary users. Instead only the local neighbourhood of each candidate correspondence needs to be shown.

Adding weight to this assertion is the issue of short-term management with users. It is a well-known fact of cognitive science that human short-term memory (SM), when compared to other attributes of our memory systems, is exceedingly limited [Miller 1956]. As we go about our daily lives, short-term memory makes it possible for you to engage with all manner of technology and the environment in general. SM is a temporary memory that allows us to remember a very limited number of discrete items, behaviours, or patterns for a short period of time. SM makes it possible for you to operate without constant referral to long-term memory, a much more complex and time-consuming process. This is critical because SM is fast and easily configured, which allows one to adapt instantly to situations. Human short-term memory is also highly volatile [Scheidermann 1984]. This means it can be erased instantly, or more importantly, it can be overwritten by other information coming into the human perceptual system. For example in the authors' opinion, with the COMA++ approach, the user is required to iterate through the candidate correspondences or locate any missing correspondences to construct all the possible correspondences. As the user iterates through the candidate correspondences they store a global view of the ontologies in their working and short-term memory. The approach will then require the user to make a judgement on the correspondence they select based on the local neighbourhood which will then enter their working memory and erase the short-term memory of the global views. After mapping the correspondences the user will be required to continue the mapping process by searching for more correspondences. However with the short-term memory of the global views being erased the user will feel disoriented and will need time to build up the global views again in working memory before continuing the process. This approach makes the task of mapping much tougher for ordinary user.

The overall process of mapping the ontologies was revealed to be quite time-consuming with on average it taking ordinary users 10 minutes to finish a mapping with COMA++ and 13 minutes with the NL prototype mapping tool, see section 3.5.1. This type of a 'one-shot' process would be too time-consuming for ordinary users. However, the time taken to map a correspondence was not as time-consuming with on average it taking ordinary users 24 seconds with the NL prototype mapping tool and 40 seconds with COMA++. In the author's opinion, the mapping process needs to be segmented into smaller sessions so ordinary users can interact in the process iteratively over time. However an iterative mapping process creates the problem that only a partial portion of a mapping can be deployed. This is discussed further in section 4.4.

There were issues with ordinary user constructing correspondences with both mapping tools. They found the interaction approach of COMA++ unfamiliar and hard, see section 3.5.3, while the button press validation interaction of the NL prototype mapping tool was too restricting in its answers, see section 3.7.4. The different formatting of question with the NL prototype mapping tool also lead to different answers, see section 3.7.4, with no consensuses from the users on the correct format to use. The mapping interaction needs to be both familiar and rich. The requirement for a rich interaction is due to the inability of ordinary users to append missing correspondences based on the requirement stated above. Using a 'tagging' paradigm will allow ordinary users to enter a rich relation for the correspondence, see section 2.6.6. By allow ordinary users to enter rich relationships for correspondences, these 'rich' relationships may be used to impact the generation/validation of new candidate correspondences. This is not investigated in this thesis but is discussed further in the conclusion section, see section 10.3.2.

Finally a previous requirement stated there was a need for an easy to understand interface for ordinary users. The natural language interface of the NL prototype mapping tool was shown to be generally understandable to ordinary users, see section 3.7.1. However several modifications will need to be made to the representation of the natural language, see section 3.7.3.

## 3.9 Limitations of Experiment

The number of participants in this experiment is quite large in comparison to other user studies undertaken related to mapping tools, see section 2.7. However the number is still too low to allow for any meaningful statistical test to be performed on the data.

There was bias in the selection of participants for the experiment, as the author selected the individuals to take part. This may have led to the participants being more favourable towards NL prototype mapping tool might be perceived as the "preferred tool". However to try and combat this issue, the order in which the participants used the tools was randomised. There was also potential bias in the selection of the music and university domains which may be easier to understand than other domains, e.g. biology domain. However using these domains allowed people with basic computing skills to participate in the experiment. The experimental design was restricted to one-to-one correspondences and did not incorporate complex candidate correspondence such as one-to-many. This was purposefully done as most modern matching tools are incapable of generating complex candidate correspondences. Additionally the author constructed the gold standard which may have caused bias. Furthermore, in the experimental design different candidate correspondences were generated for each mapping tool which made it impracticable to measure the differences in mapping results between the two tools. The experiment would have benefitted from allowing the results to be compare between the two mapping tools.

Finally during the course of the experimentation a participant who was colour blind mentioned they could not make a distinction between the colours of the concepts in the interface of the NL prototype mapping tool. An initial pilot study may have identified this issue as well as any other issues there may have been.

## 3.10 Summary

In this chapter the first experiment is presented which evaluated two different tools and approaches to semantic mapping on three user groups with different technology experience. The study compares two different tools and approaches to evaluate where the difficulty is in these approaches and which techniques work best at solving the difficulty respectively.

The experiment highlighted that using natural language to visualise the ontological information allows users to understand the information more clearly than using a graph-type display. Using a question and answer approach improves navigation by focussing the user on mapping the candidate correspondence at hand rather than navigation between different candidate correspondences. This approach is less flexible than allowing the user to search the candidate correspondences and may not suit some people but the results revealed the approach is suitable for ordinary users. In the study the results showed that non-technical users can generate mapping results to a high standard but there is still a significant difference in the precision of the mappings generated in comparison with ontologically aware users.

The study also brought some conclusions. First, although the natural language interface made the information more readable there were certain ontology specific terms which confused users, for example the term 'Thing'. The natural language needs to be filtered to remove any confusing ontology specific terms. Secondly, users were confused by some of the inappropriate candidate correspondences asked. This could lead to a loss of confidence of the system. In the author's opinion, there is a need to show the user the benefits of the mappings so that they will remain engaged even when inappropriate tasks are asked. It was decided not to investigate how to improve the candidate correspondences generated by matching tools as there are already multiple researchers tackling this issue. Finally, using a 'yes'/'no' approach is limiting to the user as it is non-flexible and forces the user to take one of the answers, even when none of the answers are correct, and can lead to different answers depending on the phrasing of the question. In the author's opinion, a 'tagging' approach will allow the user to answer in their own words and may simplify the interaction.

In the next chapter, a new incremental mapping process is presented. The process is focused on both reducing the cognitive load to the user allowing the mapping process to occur over time.

# 4 Incremental Mapping Process

This chapter addresses research objective **RO3**: Develop a mapping process that will aid ordinary users in the act of mapping over time. The new incremental mapping process: (1) reduces the difficulty of the mapping process by eliminating several user-specific mapping tasks with high cognitive load, (2) allows the mapping process to occur intermittently when required within the computing environment of users, and (3) shortens the workload of users by both decomposing the process to occur over multiple sessions and removing unnecessary user-specific mapping tasks. A specific implementation of the incremental mapping process aimed at assisting ordinary users during the act of mapping over time is presented in section 5.1 as part of the tag-based mapping tool that was developed.

This chapter is structured as follows: First a brief overview of the traditional user mapping process is given, highlighting where there are issues for ordinary users. Next the requirements needed for the new incremental mapping process are discussed. Then this new incremental mapping process is described. Afterwards a comparison is given between the incremental mapping process and the traditional user mapping process. Finally, a task walkthrough of the process is given. The incremental mapping process described here was first presented in [Conroy 2008a].

## 4.1 Traditional user mapping process

The user mapping processes used by current state of the art mapping tools were presented and described in detail in chapter 2, see section 2.3. A summary of the steps in the mapping process requiring user participation was provided in that section, see table 2-3. In this section these steps are discussed. The steps in the process requiring user interaction are the following:

- *The user decides when it is time to generate or modify a mapping (Step 1)*: This requires the user to understand when a mapping is needed and indeed when a mapping that has already been generated needs to be improved. This is non-trivial and requires deep amount of thought. This is a task with a **high cognitive load**;
- *The user opens the mapping tool, ontologies to be mapped and mapping if available (Steps 2-4)*: These tasks require the user to open and use the mapping tool in their computing environment;
- *The user selects and executes a matching algorithm to generate candidate correspondences (Steps 5-7)*: These steps are **optional** as there normally there is a

default algorithm in mapping tools. They require that the user needs to understand the matching algorithm provided by the mapping tool and the circumstances of when the algorithm should be used, as well as its benefits and downsides. These are tasks with a **high cognitive load**;

- *The user decides which mapping approach to take (Step 8)*: This is non-trivial and will require deep amount of thought from the user on which approach will suit their mapping objectives. Furthermore, a study carried out on the mapping community showed that there is no consensus on which mapping approach to take [Falconer 2009]. Some of the approaches used by the respondents included *"Tackle the easy matches first"*, *"Focus on a familiar area of the ontology"*, and *"Rely on the tool's ordering of mappings"*. This is a task with a **high cognitive load**;
- *The user constructs correspondences by verifying candidate correspondences and appending any missing correspondences (Steps 9-10)*: The first step involves reading the information of the candidate correspondences to make an informed decision on the correspondence relationship. This necessitates that the user can **understand the information being displayed** and has **knowledge and understanding of the domain** of the ontologies. The second step requires the user to search through both the ontologies to locate and add any missing correspondences. This requires the user have **modelling knowledge** to be able to explore and comprehend the ontologies. Both of these tasks have a **high cognitive load**;
- *The user decides when the correspondences generated satisfy the requirements for mapping (Step 11)*: This task requires the user to comprehend when the correspondences generated satisfy the need for mapping. The user must understand when the level of work done is sufficient. This is a task with a **high cognitive load**;
- *The user saves and deploys the mapping constructed (Steps 12-13)*: This step requires the user to save the mapping constructed and to know how to deploy the mapping.

The issues for ordinary users with the above sequence of steps are that they are very cognitively demanding. These steps also reveal that the mapping process is both labour intensive and require that the user has strong knowledge and understanding of the domain. Furthermore the user must have modelling knowledge to be able to understand the information being displayed due to the interfaces used by the current state of the art mapping tools, see section 2.4. These points indicate that the user in this process can only be a person with a knowledge engineering background or a person who has been trained by a knowledge engineer and not an ordinary user. In its current form the mapping process is even considered complex by current users. A survey carried out on the ontology mapping community revealed that the majority found the creation of ontology mappings difficult [Falconer 2009]. Furthermore the traditional mapping process does not typically take place at runtime, when the ontologies and mappings are being used by users. This is despite a need for the combined use of applications and ontologies at runtime [Klein 2001] which can require several mapping changes to be done over time at runtime.

## 4.2 Requirements

In this section the requirements for the incremental mapping process are detailed. In section 2.8, several mapping requirements were specified to allow ordinary users engage in the mapping process. Figure 4-1 displays this list of requirements with modifications made due to the evaluation results from the previous experiment, see section 3.8. The specific requirements for the incremental mapping process are highlighted in bold.

| | |
|---|---|
| **Eliminate**<br>**Any Matching API Configuration by User**<br>Visual Emphasis on Global View | **Reduce**<br>**Complexity of Mapping Process** |
| **Raise**<br>Visual Emphasis on Local Neighbourhood | **Create**<br>**Iterative Mapping Process**<br>Familiar & Rich Mapping Interaction<br>Easy to Understand Interface<br>Evaluation Methods |

***Figure 4-1:*** *Specific Requirements for the Incremental Mapping Process*

The matching configuration steps (Step 5-7) need to be removed as user tasks as they bring unnecessary complexity to the mapping process for ordinary users. It can be sufficient for the mapping system to judge which algorithm to use based on the domain. The current mapping process is too difficult for ordinary users to engage in and its complexity needs to be reduced. Three user-specific mapping tasks which have significant cognitive load associated with them, see section 4.1, and need to be changed to system-specific mapping tasks to remove the burden from the user are:

- *The user decides when it is time to generate or modify a mapping (Step 1)*: The mapping system, or application, itself can identify when a mapping is required to be generated. By monitoring the use of the mapping the application can also notice when the mapping need to be refined.

- *The user decides which mapping approach to take (Step 8)*: There has been significant work done in the state of the art on improving matching algorithms to present better candidate correspondences to the user, see section 2.2. In the author's opinion, as this work continues to improve the need for the user to find missing correspondence diminishes particularly when the mapping input from users can help identify missing correspondences [Shi 2009]. The mapping approach can then just be left to the mapping system to decide which candidate correspondences (and in which order) need to be mapped in each mapping session.
- *The user decides when the correspondences generated satisfy the requirements for mapping (Step 11)*: The mapping system, or application, is best placed to monitor the correspondences generated. This can allow the system to check if the mapping has satisfied the requirements it needed to.

By removing the previous three steps as necessary user-specific mapping tasks, the complexity in the process will be reduced. Changing these to system specific mapping tasks can also bring structure to the mapping process.

Also in the author's opinion, the workload in mapping sessions needs to be reduced. Traditionally, mapping work occurs in a couple of long sessions. Ordinary users will have neither the time nor the effort to interact in these long sessions. Instead if mapping occurs in smaller sessions over time it will allow the workload and time per session to be reduced. Although the number of sessions may increase, the sessions will take place intermittently which ordinary users should be able to manage. Thus the mapping process needs to become iterative and occur over time. A positive consequence of an iterative process is it will allow users to see the impact of their mapping decisions between sessions. A downside is that an iterative process will only develop partial mappings that will continue to be refined over time.

## 4.3 Incremental Mapping Process

The incremental mapping process, see figure 4-2, is designed to support the requirements identified above.

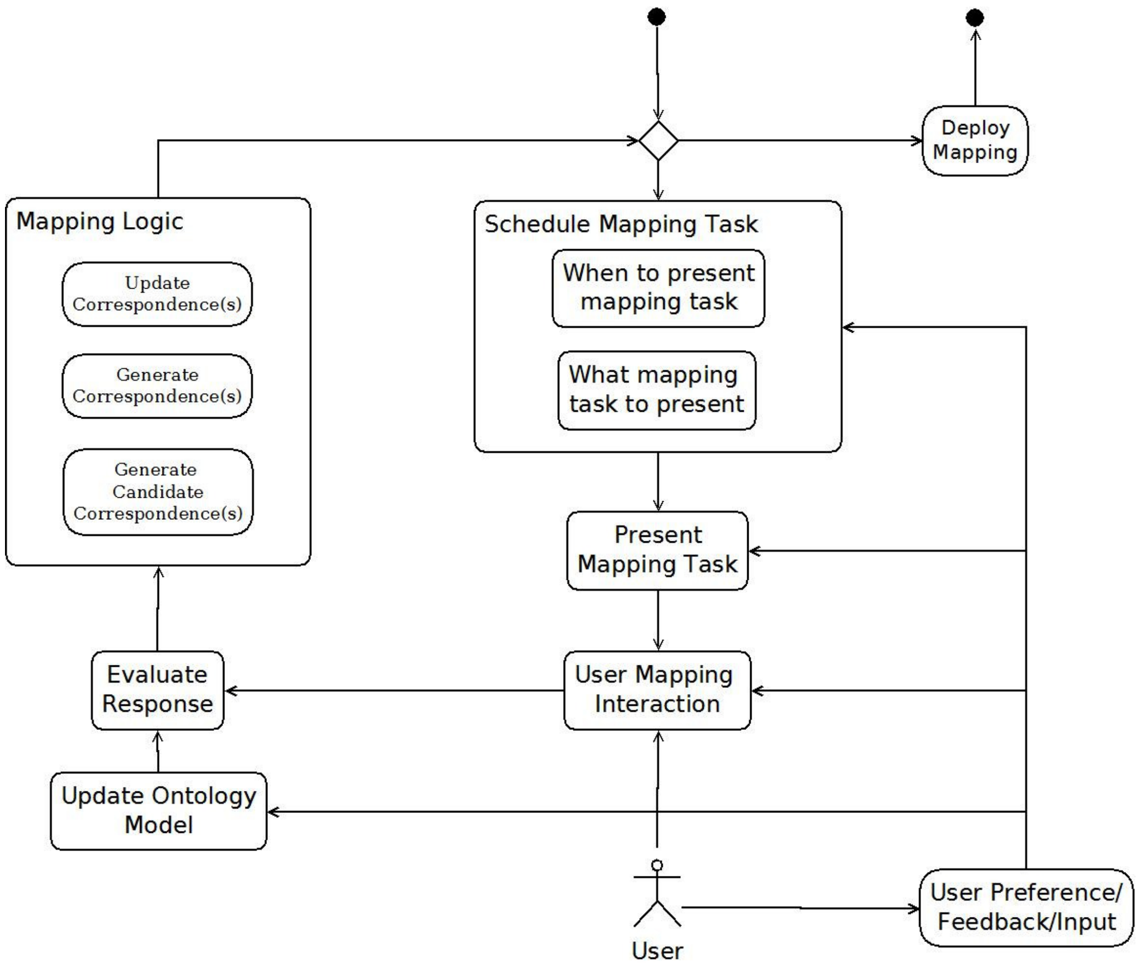



***Figure 4-2:*** *The Incremental Mapping Process*

The incremental mapping process decomposes the process of mapping into a series of user-specific mapping tasks. A user-specific mapping task can be for example verifying a candidate correspondence or selecting the matching algorithm to use to develop candidate correspondences. Table 4-1 lists the full set of user-specific mapping tasks supported by the process. In addition the incremental mapping process identifies and separates the steps of a mapping task which can allow the mapping process to be personalised through feedback from the users at each step. For example how the mapping task is presented can be tailored to the person skill level or preference, see section 4.3.2. Furthermore the design of the incremental mapping process allows the mapping process to become iterative and occur at runtime. In particular the incremental mapping process is,

- ***Simple & Usable***: This incremental mapping process reduces the workload for the user to just completing certain user-specific mapping tasks, see table 4-1. The **only required** user-specific mapping task that users must complete is to map candidate correspondences, see section 4.4. The other mapping tasks can be dealt with by the mapping system, see table 4-1 for the list of mapping tasks.

- ***Convenient & Accessible***: The incremental mapping process reduces the cognitive load on the user by removing three difficult user specific tasks and making them mapping system tasks: *When it is Time to Map*, *Which Mapping Approach to Take, When the requirement for the Mapping is Satisfied*. The 'Schedule Mapping Task' step takes care of these tasks by calculating if, when and what to map. The number of mapping tasks done by users can be tailored to allow mapping sessions to be as short as the user wants and carried out when the user wants.

In the following sections the steps of the incremental mapping process are discussed. **For the remainder of the thesis, the term "mapping tasks" is used to refer to user-specific mapping tasks.**

### 4.3.1 Schedule Mapping Task

The 'Schedule Mapping Task' step calculates if a mapping task is to be presented to the user, when to present the mapping task and what mapping task to present. Mapping sessions are constructed with multiple mapping tasks and this step manages when and what mapping sessions are to be done. It has two sub-steps that are discussed in the following sections.

#### 4.3.1.1 When to Present Mapping Task

This step involves calculating at what time to present a mapping task to the user. This allows the mapping process to become incremental, iterative and progressive. Another major benefit of this step is the ability for 'Just in time' tasks to be prompted. For example if a user is submitting a query but needs to map their ontology to another one for the query to work. This is a real benefit over the traditional approach as it allows tasks to be prompted when needed at runtime. Furthermore, it permits mapping tasks to be asked in a context that is agreeable to the user, for example not when the user is busy. When this step calculates it is appropriate to present a mapping task, the user should be given an indication asking if they wish to perform a mapping task at this time. Some examples of when a mapping task could be presented are the following:

- 'Just in time'
- When the user is perceived as being idle.
- After a specific time period, e.g. display a mapping task every hour when available.
- Having a threshold for the number of mapping tasks asked over a specific time period, e.g. no more than three mapping tasks asked per day.

#### 4.3.1.2 What Mapping Task to Present

This sub step involves deciding which mapping task is presented to the user. The mapping task can range from validating candidate correspondences to selecting a matching algorithm, see table 4-1. The tasks can be prioritised depending on their needs. For example candidate correspondences need to be generated before they are validated. Another example is that certain candidate correspondences may have higher priority based on their impact, i.e. some may allow more instances to be shared or have a big impact on other correspondences between the ontologies [Shi 2009]. The mapping task itself can force when to present the mapping tasks. An example of this is the “Just in time” example above. In contrast, the context of when the mapping task is presented can also force which mapping task is presented. For example if the user is doing something related to one of their interests then the mapping task asked could be relative to this interest. Some other examples of strategies as to what mapping task to present are the following:

- Priority based on expected need (based on user interaction patterns).
- A specific matching strategy, e.g. only lexically made candidate correspondences.
- A specific candidate correspondence confidence threshold, e.g. highest candidate correspondence above 70% confidence.

### 4.3.2 Present Mapping Task

This step involves deciding the appropiate way to display the mapping task to the user given their technical experience and current computing environment. There are multiple ways of viewing the information of a mapping task. The presentation can depend on the mapping task being asked. For example if the mapping task needs the user to prioritise which concepts are the most important in mapping (this could be part of the ‘decide mapping approach’ task) the visualisation of the ontology will need to be represented with a graph based structure whether it is tree-type or thesaurus like. It can also be dependent on the skills of the user or even the context of the user, for example if the user does not understand graph-based displays or if the device the user is using does not support a good representation of a graph based structure, e.g. PDA. Some other examples of how a mapping task can be presented are the following:

- If the user prefers to see the ontological information in natural language rather than displayed graphically.
- If the user prefers to see the natural language in bullet point form rather than in a paragraph description, i.e. different forms of representation.

- Users might want to filter away some of the ontological information dependent on whether it is necessary or unnecessary.

### 4.3.3 User Mapping Task Interaction

This step involves the user interacting to complete the mapping task. How the user interacts can be dependent on the way the task is represented or the level of technical skill of the user. The method of interaction should be familiar and simple to the user. Some examples of user mapping task interactions are the following:

- Requiring the user to answer a question with a simple 'yes' or 'no' answer.
- Requiring the user to graphically connect or drag and drop items to generate correspondences.
- Requiring the user to 'tag' the relationship for a candidate correspondence.

### 4.3.4 Evaluate Response

This step involves evaluation of the response given by the user, or system, to take appropriate mapping actions. The response from the user for a task may be that they want to generate a correspondence with a particular relationship. The responsibility of this step is to disambiguate the response from the user in order to take the appropriate action. In addition, this action may also generate several other mapping actions. For example it may cause other correspondences to become invalidated as the relations are in conflict with one another. The response from the system may be that the ontology has been changed or a mapping needs to be constructed. The responsibility of this step will be then to disambiguate the response from the system and update the correspondences that the ontology change affects or in the second case generate candidate correspondences that the user will need to validate.

### 4.3.5 Mapping Logic

This step involves executing the appropriate actions once the response has been evaluated. It encapsulates the logic of the mapping generation step and matching generation step, which involves generating candidate correspondences. For example this step may involve both (re-)categorising the mapping category of several candidate correspondences and generation of several candidate correspondences based on one user response for a mapping task.

### 4.3.6 Update Ontology Model

This step involves updating the ontology model. It does not necessarily have to be the personal model of the user and can be another ontology the user has responsibility for. When the ontology is changed it impacts the mapping and candidate correspondences already generated. So the evaluate response step is called to make the sufficient changes.

### 4.3.7 User Preference/Feedback/Input

This step involves the user giving some manner of feedback to personalise each step of the incremental mapping process. This feedback can be explicit, from user inputs into the system, or implicit, from a preference file for the user. For example the user may specify that the information be displayed with a graph-based representation or the user model of the user may suggest their technology skills are low leading to the information being displayed in a natural language based representation. This can allow the process to become personalised for each user.

## 4.4 Process Comparison

The major difference between the incremental mapping process and the traditional mapping process is that the traditional process is reliant on user input for multiple mapping tasks. For example the tasks ‘decide mapping approach’ and ‘decide when mapping finished’ both need to be completed by the user. In contrast, the incremental mapping process allows these tasks to become system mapping tasks to reduce the cognitive load. Furthermore the incremental mapping process can be implemented in multiple different ways. A specific implementation of the incremental mapping process aimed at supporting ordinary users in the mapping process is presented in the next chapter, see section 5.2. Table 4-1 displays the full list of user mapping tasks and identifies which mapping tasks need to be implemented for each process.

**Table 4-1:** Comparison of User Mapping Tasks for each Process

| Step | Mapping Task | Traditional Mapping Process | Incremental Mapping Process | Implementation of Incremental Mapping Process in Tag-based Mapping Tool |
|---|---|---|---|---|
| 1 | Decide when to map | Yes | Indeterminable– The process allows the user to decide through the user feedback step | No (System-Specific Task) |
| 2 | Open mapping tool | Yes | No (System-Specific Task) | No (System-Specific Task) |
| 3 | Open ontologies | Yes | No (System-Specific Task) | No (System-Specific Task) |
| 4 | Open mapping | Indeterminable – if one is already constructed | No (System-Specific Task) | No (System-Specific Task) |
| 5 | Select matching algorithm | Indeterminable – if the tool support is available | Indeterminable– if the tool support is available | No (System-Specific Task) |
| 6 | Configure matching algorithm | Indeterminable – if the tool support is available | Indeterminable– if the tool support is available | No (System-Specific Task) |
| 7 | Execute matching algorithm | Indeterminable – if the tool support is available | Indeterminable– if the tool support is available | No (System-Specific Task) |
| 8 | Decide mapping approach | Yes | Indeterminable– The process allows the user to decide through the user feedback step | No (System-Specific Task) |
| 9 | Verify candidate correspondence | Yes | Yes | Yes |
| 10 | Append missing correspondence | Indeterminable – if the tool support is available | Indeterminable – if the tool support is available | No (System-Specific Task) |
| 11 | Decide when mapping finished | Yes | Indeterminable – The process allows the user to decide through the user feedback step | No (System-Specific Task) |
| 12 | Save mapping generated | Yes | No (System-Specific Task) | No (System-Specific Task) |
| 13 | Deploy mapping | Yes | No (System-Specific Task) | No (System-Specific Task) |

A major drawback of the incremental mapping process is that any mapping tool using the process will only produce mappings that are partial and are undertaken over time. Any application using the mapping tool will then have to deal with an incomplete mapping. Table 4-2 displays the main benefits of the incremental mapping process in comparison to the traditional mapping approach.

**Table 4-2:** Comparison between mapping process

| **Ability** | **Traditional Mapping Process** | **Incremental Mapping Process** |
|---|---|---|
| *Embed mapping process within the users work environment* | No | Yes |
| *Mapping process to occur at runtime and over time* | No | Yes |
| *Users reduced to just answering mapping tasks* | No | Yes |
| *Mapping tasks asked in suitable contexts* | No | Indeterminable |
| *Runtime feedback of mapping actions between mapping tasks* | No | Indeterminable |
| *Mapping task selection based on the context of the user* | No | Indeterminable |
| *Personalisation of what mapping tasks are asked* | No | Indeterminable |
| *Personalisation of mapping process* | Yes | Indeterminable |

## 4.5 Task Walkthrough

This section sequentially steps through the incremental mapping process with an example task. The scenario chosen is that a person has just arrived at an art gallery when the gallery information source is discovered by their phone but requires a mapping to be developed to their interest ontology for information to be shared. The correspondences generated will be between the user's interests in art and the information for the art pieces within the gallery. For example if the user was interested in the renaissance then their interests in painting types would reflect this and candidate correspondences would be generated to the different types of paintings on display within the gallery. The user would then be responsible for mapping these correspondences. The valid correspondences would then be used to both guide the user around the gallery and provide information on the paintings.

The following situation details what happens within the incremental mapping process once the user's phone has located the gallery ontology. The first step is *schedule mapping task*. The mapping task is to be displayed to the user now as it is a 'Just in time' situation. The task selected is 'select matching algorithm' as candidate correspondences need to be generated. The second step is to *present mapping task* to the user. The phone vibrates to gain the user's attention. Once the user unlocks the phone, the task is displayed to the user on the screen. A list of different algorithms is suggested along with a description of the ontologies. For example the depth and width of the ontologies is given along with a list of the concept relationships used in the ontologies. The third step is *user mapping task interaction* where the user selects which algorithm to use based on the information given. The fourth step is *evaluate response* where the system evaluates the user's input and decides the next course of action. The 'configure matching algorithm' and 'execute matching algorithm' are system-specific mapping tasks in this scenario so the next course of action would be to generate the candidate correspondences with the algorithm selected by the user. The final step is *mapping logic* where the candidate correspondences are generated. After this sequence the process returns to *schedule mapping task* where the next user-specific mapping task is scheduled. In this scenario the next user-specific mapping task would be the 'decide mapping approach' task, where the user decides the best method of mapping the ontologies given the candidate correspondences generated, which would be prompted immediately to the user as the mapping is driven by a 'Just in time' situation. If the 'configure matching algorithm' was a user-specific mapping task then no action would take place in the mapping logic step and instead the next user-specific mapping task would be 'configure matching algorithm' rather than the 'decide mapping approach' task.

## 4.6 Summary

In this chapter an incremental mapping process was introduced. The incremental mapping process reduces the workload of the user to undertaking a series of mapping tasks. This process was designed to allow ordinary user to interact within the mapping process over time within their work environment. A specific implementation of the incremental mapping process for ordinary users will only require users to map candidate correspondences over time and is presented in the next chapter. Also in the next chapter a tag-based mapping tool which is built using this process is described.

# 5 Tag-Based Mapping Tool

In this chapter, a tag-based mapping tool is presented, in section 5.1, which is aimed at allowing ordinary users to engage in the mapping process over time within their computing environment, addressing research objective **RO4**. In the previous experiment, the button press interaction of the NL prototype mapping tool was found to be too restrictive for users, see section 3.7.4. In the author's opinion, using a 'tagging' approach allows the user to answer in their own words which would remove this issue and may also remove the burden on choosing the correct answer. Indeed the tagging interaction should be familiar to most users as it is extremely popular among ‘web 2.0’ applications and web sites, see section 2.6.6. In addition, in section 5.2, a personal information delivery tool called SportsFlows is also presented which uses the tag-based mapping tool to personalise sporting information for users.

## 5.1 Design & Implementation

In this section the details of the tag-based mapping tool are provided. First several requirements needed for the tag-based mapping tool are specified. The tag-based mapping tool architecture is then presented. The tag-based mapping tool has two major components: a browser plug-in and a mapping server. Next an overview of the architecture of the tag-based mapping tool is given. Subsequently the browser plug-in is described with emphasis upon how each step in the incremental mapping process is implemented. Afterwards the mapping server is detailed, again emphasising how each step in the incremental mapping process is implemented. Consequently the cognitive support of the tag-based mapping tool is analysed with respect to a cognitive support framework that was introduced in the literature review, see section 2.5.4. Finally, a walkthrough is given showing how the act of mapping is achieved with the tag-based mapping tool.

### 5.1.1 Requirements for Tag-Based Tool

Several mapping requirements have been specified for modifying the mapping process and tools to allow ordinary users to engage in the act of mapping, see section 2.8. These requirements were then updated with the evaluation results of the first experiment, see section 3.8. Figure 5-1 displays this updated list of requirements. The specific requirements for the tag-based mapping tool are highlighted in bold.

| Eliminate | Reduce |
|---|---|
| **Any Matching API Configuration by User** | **Complexity of Mapping Process** |
| **Visual Emphasis on Global View** | |
| **Raise** | **Create** |
| **Visual Emphasis on Local Neighbourhood** | **Iterative Mapping Process** |
| | **Familiar & Rich Mapping Interaction** |
| | **Easy to Understand Interface** |
| | Evaluation Methods |

***Figure 5-1:*** *Specific Requirements for the Tag-Based Mapping Tool*

The tag-based mapping tool is designed using the incremental mapping process, see section 4.3, which both reduces the complexity of the mapping process and allows the process to occur iteratively over time. The incremental mapping approach also reduces the user-specific mapping task to just validating candidate correspondences. However this may lead to an incomplete mapping being generated due to the inability for users to append missing correspondences, this is discussed further in section 5.1.5.1. This also removes any user-specific matching configuration step. Instead the tag-based mapping tool will use a default algorithm. The tag-based mapping tool uses a natural language interface to display the mapping task to the user. This allows the local neighbourhood of each concept in the candidate correspondence to be emphasised and highlighted to the user. Using natural language was already shown to be a suitable way of representing the information for ordinary users, see section 3.7. The mapping tasks are prompted to the user in a question and answer system over time. This allows the system to guide the user through the candidate correspondences and removes the need for the global structures of the ontologies to be displayed. The interface uses a tagging interaction to allow the user to express the relationship of the correspondence in their own words. This **changes** the **user-specific mapping task** from **validating candidate correspondences** to **mapping the relationship of candidate correspondences**. The tagging paradigm should also be familiar to ordinary users as it is used by many websites, see section 2.6.6.

### 5.1.2 Tag-Based Mapping Tool Design

Figure 5-2 gives a UML architectural diagram of the tag-based mapping tool and also identifies where the steps of the incremental mapping process are integrated. The tag-based mapping tool uses a client-server architecture and is split into two distinct parts: a browser plug-in and a

Mapping Server. The client-server architecture was chosen as it allows the users' mapping interactions and behaviour with the mapping tool to be recorded in a central location, which eases the collection of experimental results. It was also important to have a light client application on the user's machine, as the application is running in the background over time and it is crucial the application does not interfere with and make the browser run slowly which could annoy and frustrate the users.

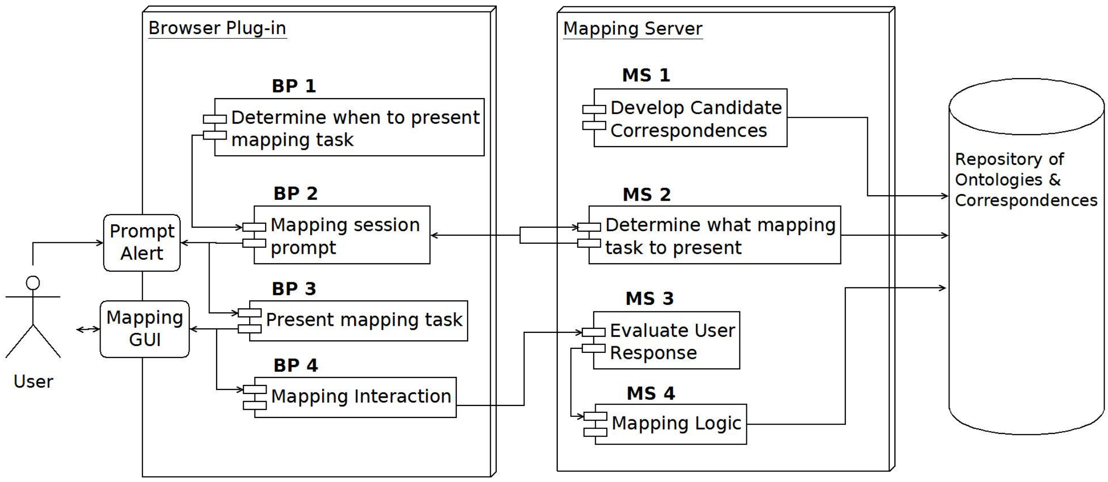



***Figure 5-2:*** *UML architectural diagram of the Tag-Based Mapping Tool*

The browser plug-in, which is a Firefox browser extension, implements the user interface of the tool. It is also designed, using the incremental mapping process, to decompose mapping into multiple mapping sessions which will allow the user to see the feedback of their actions between sessions. The mapping server, in contrast, is used to determine the mapping tasks for each mapping session and execute the mapping logic of the system. It also stores both the correspondences and ontologies. The mapping server also logs all of the users' actions within the mapping process, for experimental evaluation purposes.

The steps were separated between the client and the server based on the requirement for a light client implementation. The "determine when to present" step was implemented on the client as the step was based on whether the user has their browser open and for how long. Furthermore, this step could potentially have been implemented on the mapping server but it would have required a persistent request/response channel to be opened with the mapping server. However, this persistent channel could impact the performance of the user's browser. The "determine what to present" step was dependent on the information gained from the mapping task so it needed to be on the server to calculate the information gained by each candidate correspondence. The "session prompt", "present mapping task" and "user interaction steps" were required to be on the client as that contained the user interface. Finally the "evaluate response" step was implemented to be on the mapping server to keep the client light. The next section gives an overview of the implementation.

### 5.1.3 Implementation Overview

The tag-based mapping tool uses a client-server architecture. A Firefox browser extension[20] was the implementation of the client. The browser extension, which is compatible with all Firefox browsers up to version 2.5, uses XUL [Feldt 2007] to overlay and display elements in the browser, i.e. XUL is used to define the user interface. The primary language of Mozilla browsers is JavaScript [Musciano 2000] and thus JavaScript is used to implement the logic of the browser extension. Ajax http requests [Musciano 2000] are used to send information back and forth to the mapping server. The information is sent to the mapping server in name-value pairs. The information is responded in JSON [Crockford 2009]. However any ontology information is responded in XML[21]. The browser extension uses XSL[22] parsing to display ontology information onto the user interface.

The Mapping Server is a web application which runs on Tomcat [Chopra 2004] web server version 6. The mapping server uses SPRING framework [Walls 2005] to allow for a plug-in architecture which allows for the quick replacement of any part of the implementation. A MySQL [Widenius 2002] database is used to store information on the server with Hibernate [Bauer 2005] used as a database persistent layer in the SPRING framework. Hibernate allows for information to be cached on the mapping server. The ontologies and mapping files are stored in local folders on the machine running the Tomcat server. VTD-XML[23] was used to read and write information to both the ontologies and mapping files as it is the world's fastest XML parser[24] which allows the information to be read and written quickly to the files. The candidate correspondences are generated using the INRIA Alignment API [Euzenat 2006].

### 5.1.4 Firefox Browser Extension Implementation

A Firefox browser extension is used to implement the user interface of the tag-based mapping tool, see Figure 5-3. The user mapping interface is displayed transparently over the web page the user is viewing when the user enters a mapping session, see Figure 5-3(1). There is an icon, which looks like a question mark in a circle, that represents the tag-based mapping tool and is

[20] https://addons.mozilla.org/en-US/firefox/extensions/
[21] http://www.w3.org/TR/REC-xml/
[22] http://xml.coverpages.org/xsl.html
[23] http://vtd-xml.sourceforge.net/
[24] http://vtd-xml.sourceforge.net/benchmark1.html

displayed in the bottom right-hand corner of the browser, see Figure 5-3(2). The icon changes colour to indicate if something has occurred. Its default colour is blue. A summary of how users interact with the mapping tool is as follows: First mapping prompts are shown to the user when there is a mapping session available, see section 5.1.4.2. The user can choose to click on the prompt or ignore the prompt. If the user clicks on the prompt the mapping session will begin. Upon entering the mapping session, the mapping user interface will be transparently shown over the web page the user is currently on. The user is then asked to answer all the mapping tasks in the session. Each mapping task is iteratively shown to user. After the user has exited or finished the session the mapping interface is removed from the web page.

During the implementation phase, two volunteers were used to test and evaluate the Firefox browser extension. These volunteers were also used to compute the timing of when prompts should be displayed to ordinary users, the numbers of mapping prompts displayed per day and the number of mapping tasks in a mapping session via trial and error and the feedback of the volunteers. These results were also evaluated by a larger group in the evaluation of the tag-based mapping tool in the next chapter. In the following sections, the incremental mapping process steps included in the Firefox browser extension are detailed.

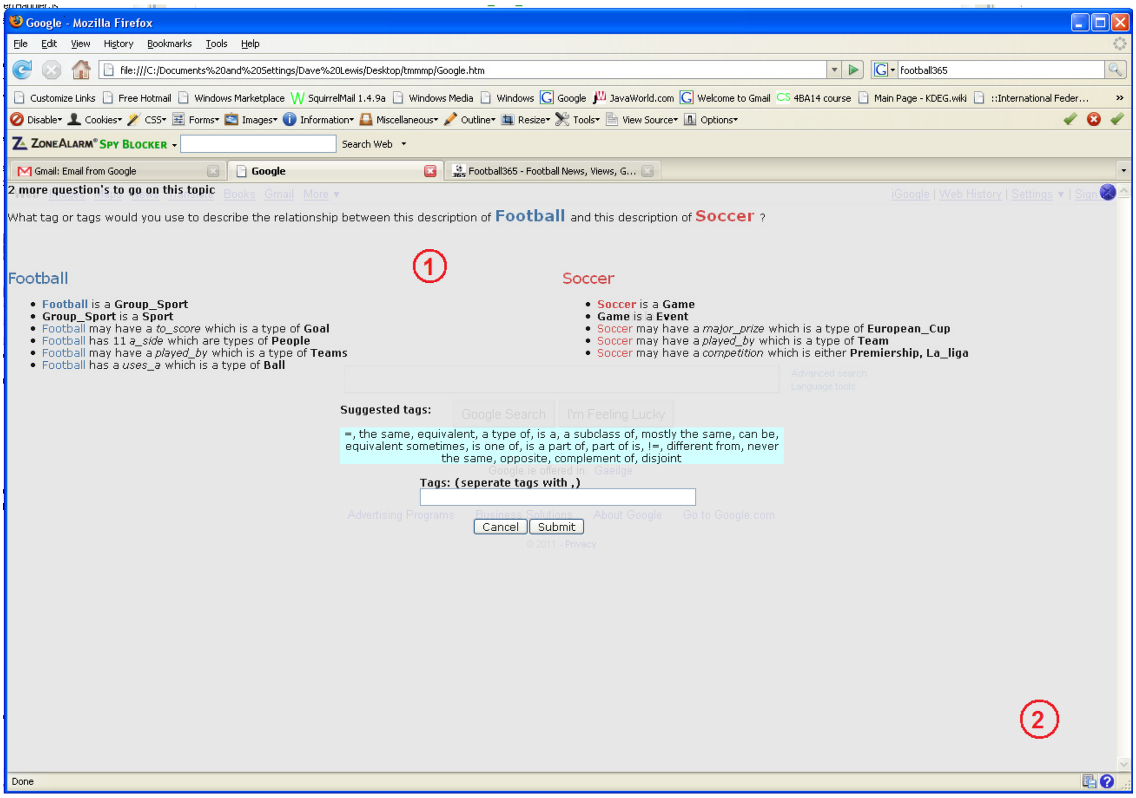



*Figure 5-3: Screenshot of the Firefox Browser Extension*

#### 5.1.4.1 Determine when to present mapping session (BP1 in figure 5-2)

Mapping sessions are prompted based on a **time-interval**. The time-interval approach will replicate the mapping process occurring over time. After testing the tag-based mapping tool with the volunteers the following was decided upon: The mapping session is to be made up of 3 mapping tasks. The maximum number of mapping prompts which can be displayed to the user each day is 3. This number was kept low due to the low number of candidate correspondences generated for the evaluation of the tag-based mapping tool, see section 6.3.3. The first mapping session prompt is displayed 10 minutes after the user opens their browser. This should give users enough time to do any urgent daily browsing work, for example checking for and reading any important emails, before being displayed with a prompt. This time was found through feedback from the volunteers and was assessed in the evaluation of the tag-based mapping tool. The next two session prompts are displayed hourly after the first prompt was displayed to give enough time between the sessions, based on the volunteer's feedback. If the user closes their browser then opens it again the number of prompts displayed will revert back to 0 for the day, this was a technical error with the implementation and was not meant to happen. This led to a number of participants in the evaluation of the tag-based mapping tool finishing all of the mapping tasks before the second week of the experiment, see section 6.3.5. The `MappingAlertHandler` JavaScript class implements this step.

#### 5.1.4.2 Mapping Session Prompt (BP2 in figure 5-2)

A mapping session is made up of three mapping tasks (or less if there are fewer than three mapping tasks for the user to complete). This low number of tasks reduces the workload of the mapping process and was evaluated from volunteer feedback. Each mapping session is prompted to the user with a message alert which is displayed in the bottom right hand corner of the browser. The prompting approach replicates the way email clients display latest emails and thus should be familiar to users. The title of the prompt is called ‘Feed Questions’ with the text ‘Answer questions to refine your feeds’, see Figure 5-4. The user enters the mapping session by clicking the prompt which will then present each mapping task in order, see section 5.1.4.2. However the user can also choose to just ignore the prompt. The tag-based mapping tool icon will then change colour to green to let the user know there is a mapping session available, see Figure 5-3(2). This allows the user to enter the session at a later time, by double-clicking on the icon. After twenty minutes if the user has not entered the mapping session the session becomes unavailable, due to the assumption that the user will not enter the session and the icon colour will revert back to blue. Again this timing was decided based on the behaviour and feedback from the volunteers. The `MappingAlertHandler` JavaScript class implements this step.

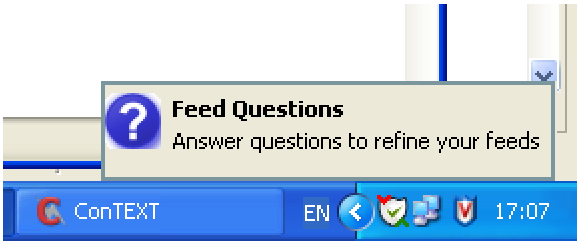



***Figure 5-4:*** *Mapping Session Message Alert.*

#### 5.1.4.3 Present Mapping Task (BP3 in figure 5-2)

The user is iteratively asked 3 mapping tasks in a mapping session. The user can exit the mapping session at any time by clicking the cancel button or the close icon in the top right-hand corner of the screen. After the three mapping tasks have been answered the mapping session closes and the mapping user interface is removed from the screen. Figure 5-5 displays the user mapping interface. The mapping tasks require the user to map the relation between concepts from the source and target ontologies in a candidate correspondence, see section 5.1.5.1. The ontological information is displayed to the user in natural language, see Figure 5-5(1). The concepts are displayed with a bullet point list with each bullet point representing an aspect of its structure, i.e. child, parent, properties. The full details for how the concepts are represented in natural language can be found in section 3.3. In contrast to the previous NL prototype mapping tool, no ontologically specific term is displayed, for example `owl:Thing` or `Agent`. Also the phrasing of the question is changed to ask the user what they think the relationship is rather than asking if the concepts are similar or correspond. These changes were made based on feedback given in the previous experiment. The interface uses the same colour and alignment as NL prototype mapping tool, see section 3.3, to make a distinction between the concepts. The mapping interface contains a suggested relations list where users can select a suggested relation instead of entering one of their own, see Figure 5-5(2). The user enters the relationship for the correspondence in the tag box, see Figure 5-5(3). The next section gives more details regarding how the user interacts with the interface. The mapping interface is displayed transparently over the web page the user is viewing, which gives the user the feedback that they are still on the web page they are viewing. This again was decided upon based on the feedback from the volunteers where they liked the transparency as they felt they were still on the webpage and was assessed in the evaluation of the tag-based mapping tool. The `MappingSessionHandler` JavaScript class implements this step.

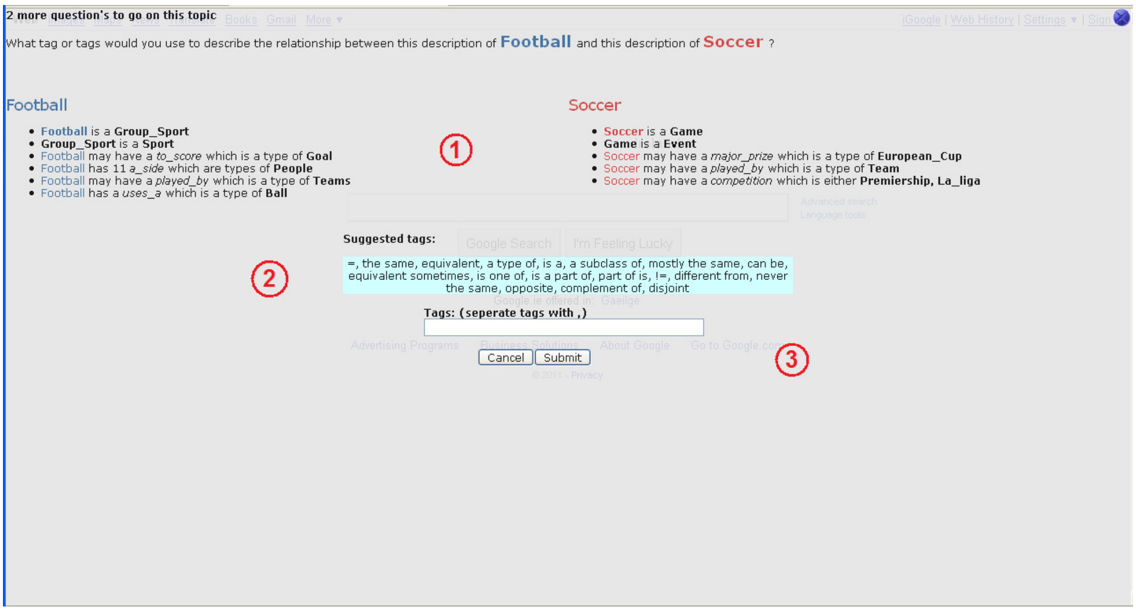



***Figure 5-5:*** *The mapping interface for the tag-based mapping tool*

### 5.1.4.4 Mapping Interaction (BP4 in figure 5-2)

The mapping task presented requires the user to 'tag' the relation for each candidate correspondence, which will then be used to categorise the correspondence, see section 5.1.5.3. The interaction uses both a tagging paradigm and a structured sentences paradigm, see section 2.6. Users can express in their own words (tag(s)) the relationship by entering it in the tag box on the mapping interface, see Figure 5-5(3). Users can also select tag(s) from a suggested relations list by clicking on them which adds them to the tag box. Users then click on the submit button to submit the tag(s) entered as the relationship of the candidate correspondence. If the user clicks on the cancel button the mapping session will end and the mapping interface will be removed with no tag(s) sent to the server for the current candidate correspondence presented. Any previously answered candidate correspondence in the mapping session will have their tag(s) sent to the server. The tag(s) entered by the user which are not already in the suggested relation list are called **user-defined relations**. The suggested relations list is constructed from several default relations suggested by the tool and any user-defined relations entered by the user, see Figure 5-6 for the full list of default relations suggested. In addition multiple tags can be used to define the relationship, each tag has to be separated with ','. The type of interaction used should be familiar to most users as tagging is extremely popular among 'web 2.0' applications and web sites, see section 2.6. The `MappingInteractionHandler` JavaScript class implements this step.

### 5.1.5 Mapping Server

The mapping server is used to store information and to execute several system-specific mapping tasks. The information stored includes ontologies, candidate correspondences and any complete or partially complete mapping developed. The candidate correspondences are generated when the mapping server is compiled. The updated version of the mapping tool allows candidate correspondences to be generated at runtime, see section 7.2.4.1. The mapping server decides what candidate correspondences are to be presented in each mapping session. The mapping server is also required to evaluate the user's actions and to take the appropriate mapping actions subsequently. In the following sections, the incremental mapping process steps in the mapping server are detailed.

#### 5.1.5.1 Develop candidate correspondences (MS1 in figure 5-2)

The candidate correspondences are generated using the INRIA Alignment API. The `NameAndPropertyAlignment` alignment method was selected to be used. This alignment was chosen as during the trial of the tag-based mapping tool this alignment method generated the most useful candidate correspondences, in comparison to the other alignment types, for the ontologies used in the experiment. Most useful means that the candidate correspondences generated involved more of the concepts that mattered in the use-case used in the evaluation of the tag-based mapping tool, see section 5.2. The `CorrespondenceService` Java class implements this step.

An issue with the tag-based mapping tool is the inability of the user to append any missed correspondences. However it can be assumed that as the precision and recall continue to increase in matching tools [OAEI 2011], the matching tools themselves will provide improved candidate correspondences. Furthermore, in the updated tag-based mapping tool, a new component is developed which looks periodically for potential candidate correspondences, see section 7.2.4.1, to offset the need for users having to append missed correspondences.

#### 5.1.5.2 Determine what mapping tasks to present (MS2 in figure 5-2)

The only user-specific mapping task asked by the mapping server requires the user to map the relationship between concepts in a candidate correspondence. The other mapping tasks are all system-specific mapping tasks and are implemented in the steps of the incremental mapping process, for example this step. The previous step generates a list of candidate correspondences, for each user. The decision as to which candidate correspondence to present to the user is based on the number of instances available through the correspondence, i.e. the number of instances the target concept has. A ranked list is devised based on the number of instances and the candidate correspondences are presented iteratively from the highest ranked to the lowest ranked. This will allow the candidate correspondences which are the most beneficial at pruning the information, to be asked first to the user, see section 5.2. If two or more candidate correspondences have the same number of instances available, then the candidate correspondence with the higher confidence value will be presented first, see section 2.1. The `CorrespondenceService` Java class implements this step.

#### 5.1.5.3 Evaluate User Response (MS3 in figure 5-2)

The user enters tag(s) to define the relationship of candidate correspondences, see section 5.1.4.4. The mapping server is then responsible for evaluating what relationship is meant by the tag(s) the user entered. Once the relationship has been identified the correspondence can be verified by the mapping server, see section 5.1.5.4. The taxonomy of categories was devised by the author, see Figure 5-6, to allow the relationship of the tag(s) entered to be defined. The taxonomy groups the relationships concepts can have into mapping categories which are used to determine which correspondences are valid for the interchange of instance information. The valid correspondences are then used by SportsFlows to personalise the sporting information shown to users by mapping the interests of the users to sporting stories, see section 5.2. In the taxonomy the concept relationships contain the default relations suggested by the mapping tool. The concept relation and mapping category of a tag is found by identifying where the tag lies in the taxonomy. For example if the tag was 'Is a type of' then the concept relation would be subclass and the mapping category would be *equivalent sometimes*. Any tags which are user-defined relations will have no concept relation in the taxonomy and have *unknown* mapping category. The taxonomy was validated during the implementation phase when two volunteers were using the tool and it worked as intended, however there has been no validation of the design of the taxonomy. The `MappingService` Java class implements this step.

- **Equivalent**
  - *The same*
    - The same
    - Equivalent
    - =
  - *Super class*
    - Mostly the same
    - Can be
  - *Intersection*
    - Part of is
- **Equivalent Sometimes**
  - *General*
    - Equivalent sometimes
  - *Subclass*
    - Is a type of
    - Is a subclass of
  - *One of*
    - Is one of
  - *Union*
    - Is a part of
- **Different**
  - *Different from*
    - Different from
    - Never the same
    - Different
    - Not the same
    - !=
  - *Complement of*
    - Opposite
    - Complement of
  - *Disjoint*
    - Disjoint
- **Unknown**

***Figure 5-6:*** *The taxonomy of categories. Mapping categories are the top level and bolded, the corresponding list of concept relation categories are the sub level and are in italic. The third level represents the default relations used in the mapping tool.*

The set of possible mapping categories are: *Equivalent* (all instances of target concept are contained in the source concept), *Equivalent Sometimes* (some instances of target concept are contained in the source concept), *Different* (no instances of target concept are contained in the source concept) and *Unknown*. The *Equivalent* and *Equivalent Sometimes* categorises are used to categorise correspondences which serve as valid correspondences for interchange of information in the personal information delivery use-case, see section 5.2. These mapping categories could serve a different purpose in other use-case implementations. For example a stricter use-case implementation may only consider equivalent categories as valid for information transfer. Equivalent sometimes correspondences are specified where the target concept is more general than the source concept or only part of the source and target intersect.

The *Different* category is used to categorise correspondences which are not considered valid. The *Unknown* category is used to categorise user-defined relations. The user-defined relations are not redefined into other categories in this implementation of the tag-based mapping tool. However the utility of the user-defined relations is measured in the evaluation of the tag-based mapping tool, see section 6.5.2.2. The rationale behind the organisation of the taxonomy is given as follows:

- *Superclass* belongs under equivalent as this relationship implies that all the instances of the target concept will belong to the source concept. For example, 'Spanish Football Team' is a superclass of 'Barcelona F.C.';
- Conversely, *Subclass* belongs under equivalent sometimes as the relationship implies that the source is a subclass of the target and that only some of the instances of the target will belong to the source. For example, 'Lionel Messi' is a subclass of 'Argentine Football Player';
- Only one-to-one mappings are considered in this thesis so the *Union* and *Intersection* concept relations are considered as follows: for three concepts A, B and C if A ᴗ B = C (union) then A = C\B (intersection). Then:
  - *Union* is defined as the relation between A and C. Therefore the union relation belongs under equivalent sometimes as the relationship implies that only some of the instances of the target concept C will belong to the source concept A, i.e. A is part of C. For example 'Men' ᴗ 'Female' = 'Human Beings' then only some of the instances of 'Human Beings' are also the instance of 'Men';
  - *Intersection* is defined as the relation between C and A. Therefore the intersection relation belongs under equivalent as the relationship implies that all the instances of the target concept A belong to the source concept C, i.e. part of C is A. For example 'Men' ᴗ 'Female' = 'Human Beings' and all the instances of 'Men' are also instance of 'Human Beings.

#### 5.1.5.4 Mapping Logic (MS4 in figure 5-2)

This step has responsibility to evaluate the mapping category for correspondences based on the mapping category for each tag entered by the user, which was calculated in the previous step. The default mapping category for each correspondence generated is the *Equivalent* category.

After the mapping category of each tag has been calculated the categories for the tag(s) are then put into a category list for the correspondence. The mapping category of the correspondence is then calculated based on the category list via the following steps:

1. Remove all *Unknown* categories from the list,

2. If the list is empty then the mapping category of the correspondence remains as the default mapping category *Equivalent*,

3. If there is a majority mapping category in the remaining list then this category is the new mapping category of the correspondence,

4. There is a majority tie between two or more mapping categories. Then the first mapping category in the list (of the categories which have tied), i.e. the first tag the user entered, is the new mapping category of the correspondence.

For example if the user enters a tag from the *one of*, *super class* and *the same* relation categories then the correspondence would be categorised as *equivalent* as there are two tags with that mapping category and only one with the *equivalent sometimes* mapping category. Figure 5-7 shows the mapping category cycle of correspondences. The `MappingService` Java class implements this step.

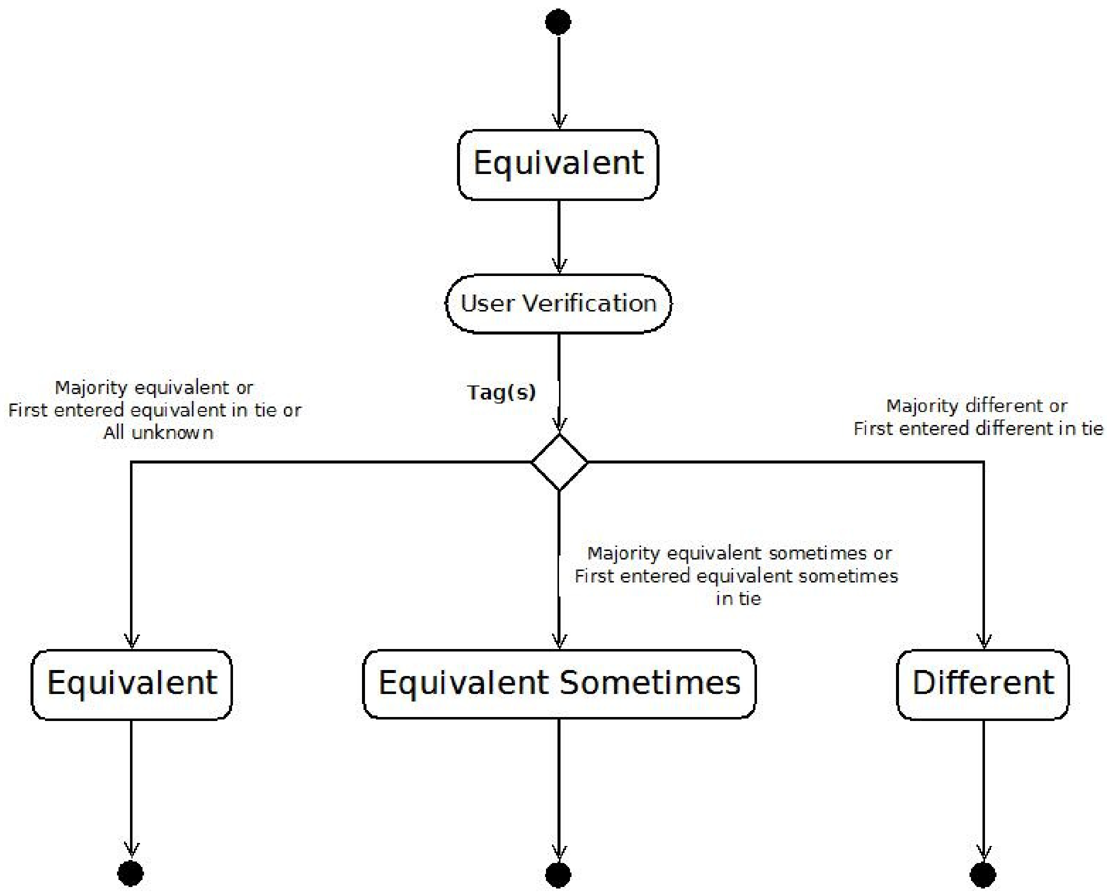



***Figure 5-7:*** *The mapping category lifecycle of a correspondence*

### 5.1.6 Cognitive Support

In section 2.5.4 a theoretical framework for cognitive support in mapping tools was presented. Table 5-1 details the requirements of this framework and specifies which requirements are supported and are not supported in the tag-based mapping tool. The rationale for why certain requirements are not supported is given.

**Table 5-1:** Theoretical Framework for Cognitive Support

| Number | Principle | Requirement | Supported |
|---|---|---|---|
| *Analysis and Decision Making* | | | |
| #1.1 | Discover mappings | Support ontology exploration and manual creation of mapping correspondences. Provide tooling for the creation of temporary correspondences that the user can address at a later time. | **No** – It was discovered that ordinary users found it difficult to find missing correspondences, see section 3.5.2. |
| #1.2 | Make mapping decisions | Provide a method for the user to accept/reject a suggested correspondence | **Yes** |
| #1.3 | Inspect definition of term | Provide access to full definitions of ontology terms | **Yes** |
| #1.4 | Inspect context of term | Show the context of a term when a user is inspecting a suggested correspondence | **Yes -** the local neighbourhood is represented<br>**No** – The full global ontology is not displayed |
| *Interaction Dimension* | | | |
| #2.1 | Explore ontologies | Provide interactive access to source and target ontologies | **No** – The only user-require mapping task is to map candidate correspondences and this task does not require the user to explore the ontologies |
| #2.2 | Explore/verify potential mappings | Support interactive navigation and allow the user to accept/reject candidate correspondences | **Yes** – The system decides which candidate correspondences to present to the user<br>**No** – the user cannot explore the candidate correspondences themselves |
| #2.3 | Explore/remove verified mappings | Support interactive navigation and removal of verified correspondences | **No** – the system only requires users to map correspondences to reduce the complexity and load for ordinary user |
| #2.4 | Perform search and filter | Provide support for searching and filtering the ontologies and mapping | **No** – the system only requires users to map correspondences to reduce the complexity and load for ordinary user |
| #2.5 | Direct creation & manipulation of the mappings | Support for adding details on verified correspondences and manually create correspondences | **No** – the system only requires users to map correspondences to reduce the complexity and load for ordinary user |
| *Analysis and Generation* | | | |
| #3.1 | Generate mappings | Support the automatic discovery of candidate correspondences | **Yes** |
| #3.2 | Execute mappings | Allow the user to test a mapping by automatically transforming instances from the source to the target ontology | **No** – the system only requires users to map correspondences to reduce the complexity and load for ordinary user |
| #3.3 | Save verification state | The verification process must support potential interruptions by automatically saving and returning users to a given state | **Yes** |
| #3.4 | Conflict resolution and | Support identification and guidance for resolving conflicts | **No** – the system only requires users to map correspondences to |

| | | | |
|---|---|---|---|
| | inconsistency detection | | reduce the complexity and load for ordinary user |
| *Representation Dimension* | | | |
| #4.1 | Source and target ontologies | Provide a visual representation of the source and target ontology | **Yes** – only local neighbourhood of the candidate correspondences<br>**No** – the global view of the ontologies is not displayed |
| #4.2 | Potential mappings | Provide a representation of the potential correspondence describing why it was suggested, where terms are in ontology, and their context | **Yes** – the correspondence is represented with identification of where the terms are in the ontology, i.e. the parents of the concept is given<br>**No** – description on why suggested is not given as not supported by matching API used, also the task asks the user to enter the relation by giving a reason for suggestion may sway the users answer |
| #4.3 | Verified mappings | Provide a representation of the verified correspondences that describe why the correspondence was accepted, where the terms are in the ontologies and their context | **No** – only user-specific mapping task is to map a candidate correspondences, management of correspondences is left to the system |
| #4.4 | Identify "candidate-heavy" regions | Identify visually candidate-heavy regions based on the automated mapping procedure | **No** – the system decides which candidate correspondences the user is to map |
| #4.5 | Identify possible starting points | Indicate possible start points for the user, e.g. flag terms that have exact name matches, as these are generally the most straight-forward correspondences to perform | **No** – the system decides which candidate correspondences the user is to map |
| #4.6 | Progress feedback | Provide progress feedback on the overall mapping process | **Yes** –progress feedback for mapping session<br>**No** – no progress feedback on overall mapping process |
| #4.7 | Reason for suggesting a mapping | Provide feedback explaining how the tool determined a potential correspondence | **No** – description on why suggested is not given as not supported by matching API used, also the task asks the user to enter the relation by giving a reason for suggestion may sway the users answer |

For example the "source and target ontologies" principle above is partially supported by displaying the local neighbourhood of the concepts during the mapping of the relationship of a correspondence task. The principles relating to this task are supported by the tag-based mapping tool. However the other principles are not supported due to those being assigned to mapping tasks which are not user-specific in the tag-based mapping tool. This is due to several mapping tasks being identified as too complex for ordinary users to achieve, see section 4.2.

### 5.1.7 Implementation Walkthrough

This section sequentially steps through creation of a mapping in the tag-based mapping tool and describes the role the user has in the mapping cycle. An implementation walkthrough is given of how candidate correspondences are mapped by the user and what implementations are involved. The mapping cycle does not include the generation of candidate correspondences as this step is done prior to runtime, see section 5.1.5.1. Figure 5-8 gives the UML diagram for the sequence of mapping steps within the tag-based mapping tool.

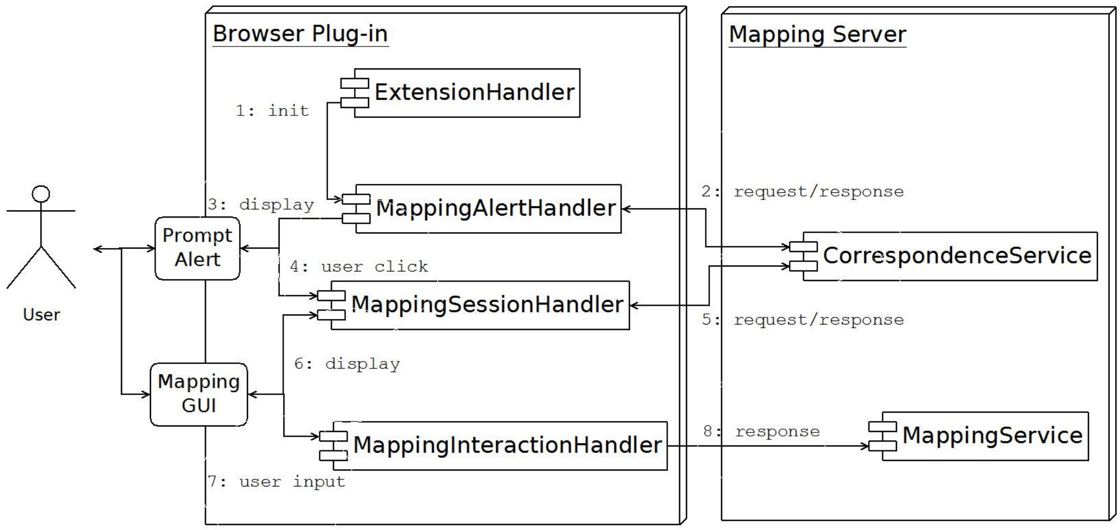



***Figure 5-8:*** *UML sequence diagram of Mapping Cycle*

The first step in the sequence involves determining when to present a mapping session to the user. The `MappingAlertHandler` class is responsible for deciding when to present a mapping session to the user. The availability of a mapping session is presented to the user with a message prompt, see section 5.1.4.2. The prompts are shown to the user at certain time intervals, see section 5.1.4.1. The prompt is only shown if there are candidate correspondences that need to be mapped. This is evaluated by sending a request to the `CorrespondenceService` on the mapping server to determine what mapping tasks the mapping session will consist of. The response includes the ids and priority order of three candidate correspondence. Candidate correspondences that will allow the most instance information to be shared have the highest priority and will be asked to the user first, see section 5.1.5.2. The user can enter into a mapping session by clicking the prompt or clicking the tool icon in the bottom right corner of the Firefox browser, see Figure 5.3(2). When the user clicks on the prompt the mapping session begins and the cycle moves onto the next step.

The next step involves presenting the mapping session to the user. The `MappingSessionHandler` class is responsible for presenting each candidate correspondence iteratively to the user. The ontological information for each candidate correspondence is obtained from the `CorrespondenceService` on the mapping server. The list of candidate correspondences is presented to the user iteratively in the mapping session with the highest priority asked first. The mapping interface is displayed to the user as a graphical overlay on the user's current web page, see Figure 5.2. The ontological information is represented in natural language, see section 5.1.4.3, and the user is tasked with entering the tag(s) which best define the relationship between the concepts from their point of view. The mapping interface also lets the user know how many candidate correspondences are left in the mapping session, e.g. '2 more questions to go'. The user can either choose from the suggested relation list or enter user-defined relation(s) which are not in the suggested list, see section 5.1.4.4. In addition several tags can be used to define the relationship. The user clicks on the submit button to verify that the tag(s) entered define the relationship. At any time, even if the user has not mapped the relationship for any candidate correspondences of the session, the user can exit the mapping session by clicking the cancel button or close icon. Once all the candidate correspondences have been mapped by the user or the user has exited the session, the `MappingInteractionHandler` class forwards the tag(s) for each completed task to the `MappingService`. If the user exits the mapping session before verifying the tag(s) for the relation then these tag(s) will not be forwarded.

The final step involves evaluating the user's response. The `MappingService` is responsible for mapping candidate correspondences. This involves calculating which mapping category the candidate correspondence belongs to. The tag(s) entered by the user to define the relationship are used to calculate the mapping category. First the mapping category of each tag is evaluated using the taxonomy of categories, see section 5.1.5.3. Then using the mapping category lifecycle the categories of the tag(s) are used to calculate the mapping category of the correspondence, see section 5.1.5.4. Then the correspondences mapping category is modified. After this step the mapping cycle moves back to the first step. No notification is given to the user of any mapping change. The user will only see the change in the modification made by tools using the correspondences mapped by the user.

## 5.2 SportsFlows

In this section a personal sporting information delivery tool called SportsFlows is presented. SportsFlows uses the tag-based mapping tool to personalise the sporting information shown to the users. It is important that the tag-based mapping tool supports a service which performs a job for ordinary users in a functional, social or emotional dimension. This is to allow the tag-based mapping tool to be adopted and used by ordinary users, and thus be evaluated in their computing environment. As Harvard marketing professor Theodore Levitt famously said *People don't want to buy a quarter-inch drill. They want a quarter-inch hole!* In particular, the idea was to develop SportsFlows to enable the tag-based mapping tool to be evaluated over two weeks within ordinary users' computing environments, see section 6.3. In the author's opinion, sports is very popular and will allow for a wider selection of people with different skills to be tested than with other domains like music or financial. SportsFlows was needed in order to be able to have a long experiment with ordinary users using the tag-based mapping tool within their own computing environment over time. Figure 5-9 gives an overview of the architecture of SportsFlows. The tool is built on top of the tag-based mapping tool and thus it uses the same client-server architecture.

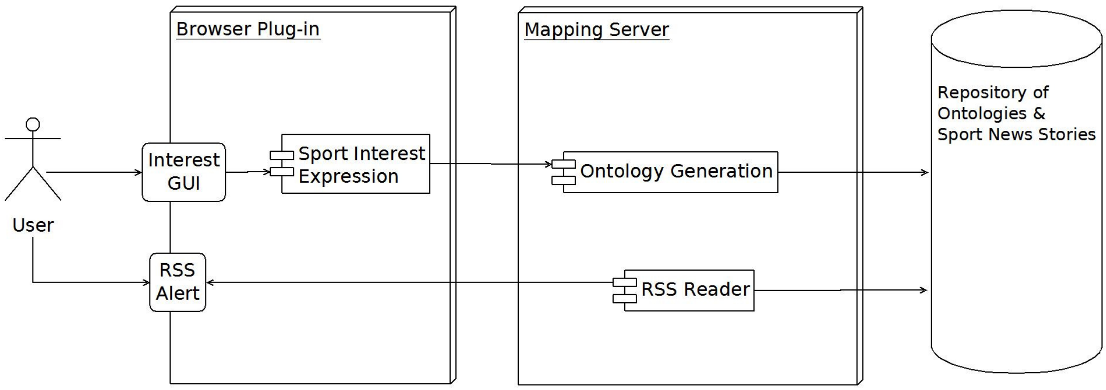



***Figure 5-9:*** *Architecture of SportsFlows*

The scenario used is first presented. Afterwards the information service ontology is described. Then an explanation of how personal ontologies are constructed for each user is detailed. Finally a description is provided of how the user will see the feedback on their mapping actions.

### 5.2.1 Scenario

The scenario chosen was the personal information delivery of sporting information. In this scenario a personal ontology will represent the user's interests and the information service ontology will represent the available information from streams on the web, i.e. RSS feeds. RSS feeds are a web feed format used to publish frequently updated works in a standardised format [Libby 1999]. The mapping constructed by the users will be used to personalise the information delivery from the RSS streams to the user's browser. The benefits the user will gain from mapping is that the information displayed to the user will get better over time and be refined to their interests as they continue to partake in the mapping process.

The sporting scenario was chosen as RSS feeds are generally updated regularly over the course of a day, particularly sporting news, which will allow the feedback of mapping actions undertaken by users to be observed more quickly than in many other information flows. Also there are numerous sporting RSS feeds available on the web allowing for multiple content stories available to be mapped to the users' interests each day. Finally the sporting area is familiar to most people which allows for a large set of users that can evaluate the tag-based mapping tool.

### 5.2.2 Information Service Ontology

The `RSSReader` class on the mapping server is used to read and store sporting news stories from sporting RSS feeds, see Figure 5-9. The RSS feeds used in SportsFlows are BBC[25], RTE (the Irish national broadcaster)[26], and Breaking news[27]. These RSS feeds were chosen as they provided the best source of sporting information for the users who were selected for evaluating the tag-based mapping tool, see section 6.3.1. The RSS feeds are searched for new stories every 20 minutes. This frequency was decided upon after observing that these RSS feeds are in general updated with new information every 20 minutes. SportsFlows has the ability to append other RSS feeds, by appending the RSS feed URL location to a table in the repository.

The mapping server has an information service ontology which is used to assign the sports news stories, from the RSS feeds, to a concept and store them as instances for that concept. The information service ontology is provided in the accompanying DVD media under Experimental Datasets - Experiment 2. The information service ontology was constructed by the author after a search for sport ontologies on the web found no suitable ontology to represent sport news feeds. The ontologies found during the search did not encapsulate the entire range of sports needed. Several of the concepts in the information service ontology were enriched with *added labels* to allow multiple different stories to be saved as instances, see Appendix C for list of concepts and their added labels. None of the concepts had any overlapping *added labels*. For example some of the added labels for the concept 'Soccer' were the following 'Liverpool', 'Arsenal', and 'Chelsea'. These added labels would be used to categorise sport news items from the RSS feeds by searching the title of the sport new item for these labels.

[25] http://newsrss.bbc.co.uk/rss/sportonline_world_edition/front_page/rss.xml
[26] http://www.rte.ie/rss/sport.xml
[27] http://feeds.breakingnews.ie/bnsport

#### 5.2.2.1 Populating the Ontology with Instances

The sports news stories, from the RSS feeds, would be added as instances to the concept which had the most *added labels* in the title. If no *added labels* were found in the title for a sport news story, for any of the concepts in the information service ontology, the story is not saved and discarded. If two concepts had the same number of *added labels* in the title, then the story is added as an instance to the concept which had the first *added label* in the title. The decision for this approach was made after observing sports news stories from sports feeds for 4 months. It was noticed that in general the first few terms mentioned in a sports story are more related to the subject of the story. In addition, the approach taken also allows for a high amount of flexibility as the *added labels* can be modified. In fact, during the implementation the *added labels* were continually modified until the results were to a high enough standard. Several other approaches were also evaluated such as taking specific RSS streams for each sport but none gave as high results with as much flexibility. Although it is not implemented in the current version of SportsFlows, using this approach allows for the stories to be filtered to interests rather than sports, see section 7.3. The sport news stories were saved to a table in the repository on the mapping server with a reference to the corresponding concept in the information service ontology.

### 5.2.3 Personal Ontology

The same sports ontology is used to represent each user's own personal sporting ontology. The sport ontology was constructed by the author after a search for sport ontologies on the web found no suitable ontology to represent the sporting interests of users. The sport ontology is provided in the accompanying DVD media under Experimental Datasets - Experiment 2. The sports ontology is personalised by each user selecting their 'concept interests'. The 'concept interests' are selected from a list of concepts from the ontology which represent various different sports. For example the user could select Football, Rugby and Tennis as their 'concept interests'. Figure 5-10 gives a display of how the user selects their 'concepts interests'.

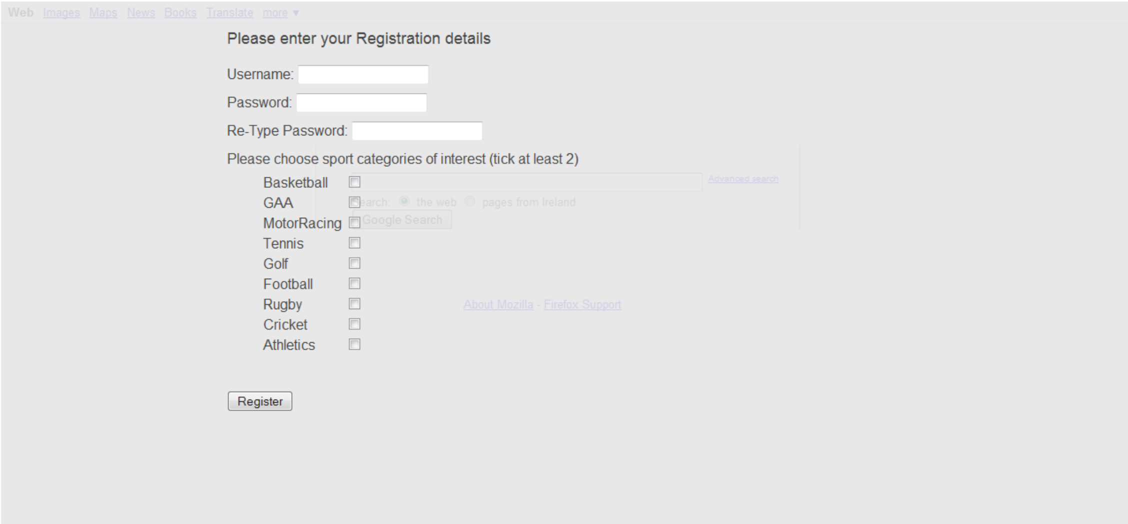



***Figure 5-10:*** *The 'Concept interest' Interface*

These 'concept interests' are used to personalise the sporting news for the user by only showing sport news stories received from the mapping server with correspondences generated from these 'concept interests'. Even though each user could have different 'concept interests' they are still required to answer all of the mapping tasks, i.e. map all of the candidate correspondences. This can potentially be annoying for users to be asked tasks not relating to their interests and may negatively affect the evaluation, see section 6.4. However it allows the mapping results of users to be compared between each other as the same ontologies and candidate correspondences are used in the evaluation. This is discussed further in section 6.3.

### 5.2.4 Mapping Incentives

SportsFlows displays sporting news to the user received from the mapping generated between the user's own personal sporting ontology and the information service ontology. When the candidate correspondences are first generated they are categorised as *equivalent*. This is before the user has mapped the candidate correspondence. Hence this provides all the sport news stories for the correspondences with 'concept interests', see section 5.2.3. Therefore the act of mapping is a pruning approach which will remove any sport news story which is not related to the user's own interests. This approach was chosen as the act of mapping should be able to show a quick change to the information being presented to the user using this method. Also if a branching out method[28] was chosen the user would not see any sport news stories for their interests until they mapped a candidate correspondence related to their interests.

[28]In contrast to the pruning approach the branching out method only shows information that is related to the users interests and thus the amount of information only increases as correspondences are mapped by the user, see section 7.2.4.1

Any 'latest' sport news story, stories within the last hour which have not already been displayed to the user, which are related to the users interests will be displayed in a message prompt in the right hand bottom corner of the screen, see Figure 5-11. The title of the prompt is the headline of the sport news story and the text is the first few words of the description of the sport news story.

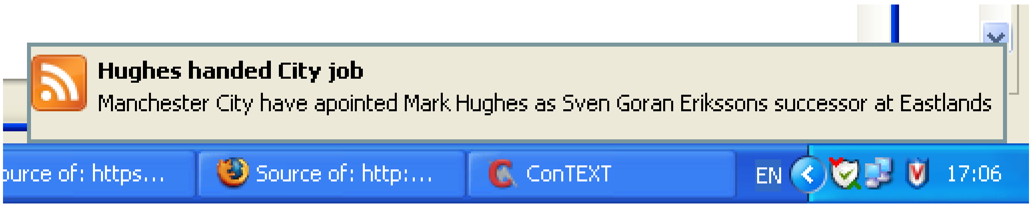



***Figure 5-11:*** *Latest sport news story prompt is displayed in the bottom right hand corner of the browser.*

If the user clicks on the prompt the sport news story will be opened in a new tab on the Firefox browser. Also a sidebar displaying all the sport news stories will be opened, see Figure 5-12. The prompt is displayed for 8 seconds before closing, giving the user enough time to see the story without it becoming too intrusive. This time was calculated from trial and error with the volunteers during implementation and was measured in the evaluation of the tag-based mapping tool. If the user chooses instead not to click on the prompt, the tag-based mapping tool icon in the bottom right hand corner of the Firefox browser will change colour to orange to indicate to the user that new sports news stories are available. If the user double clicks on the icon the sport news story sidebar will open displaying the new sport news stories and the icon reverts back to its original colour. After twenty minutes if the user has not opened the sport news story sidebar the icon will revert back to its original colour to signify the stories are no longer the latest, as the `RSSReader` iterates through RSS feeds every twenty minutes.

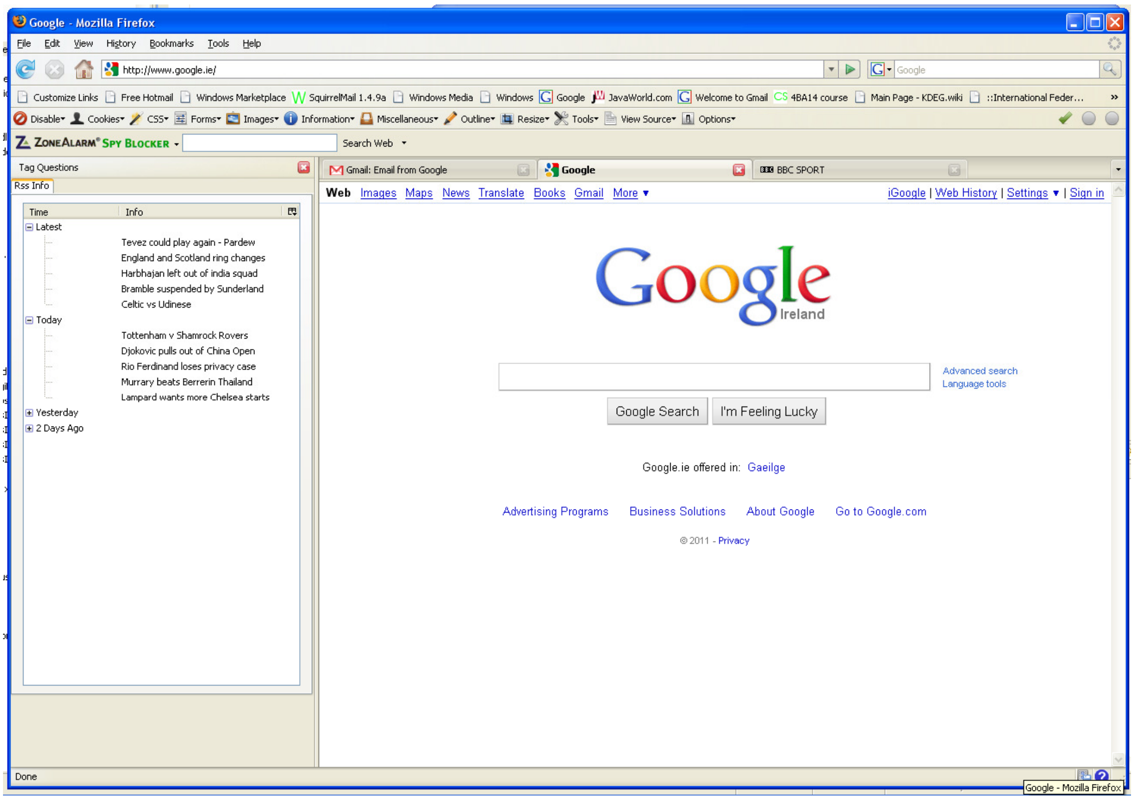



***Figure 5-12:*** *Sport news story sidebar*

The sidebar separates the sporting news into four different categories: *latest* the sport news received within the last hour, *today* the sports news for today excluding the last hour, *yesterday* the sport news for yesterday, and *2 days ago* the sports news two days ago. By double clicking on any sport news story the story will open on a new tab in the Firefox browser. Each category can be collapsed or opened. When the sidebar is opened only the latest and today category are opened. Thus, SportsFlows enables the personal information delivery of sports news by filtering the stories shown to the user using the mapping constructed by the users with the tag-based mapping tool.

## 5.3 Summary

This chapter described a tag-based mapping tool built using the incremental mapping process detailed in the previous chapter and techniques identified in both the state of the art literature review as well as results from the previous user evaluation. The tool is seamlessly embedded in the ordinary user's browsing environment using a Firefox browser extension which allows the mapping process to occur over time. The act of mapping is reduced to just one user-specific mapping task, i.e. mapping candidate correspondences. The tag-based mapping tool aims to simplify the task by using a tagging interaction and natural language to represent the ontological information. A personalised information tool, SportsFlows, built on top of the tag-based mapping tool is also presented. SportsFlows intends to provide incentives for mapping by displaying the feedback of mapping actions through the personalisation of sporting news.

In the next chapter, the tag-based mapping tool is evaluated by measuring the efficiency, effectiveness, convenience, accessibility and simplicity of the act of mapping with ordinary users using the tool over time within their own computing environment.

# 6 Tag-Based Mapping Tool Experiment

In this chapter the second of the three experiments is presented. The goal of this experiment was to address research objectives **RO1**: Determine whether ordinary users can actively participate in the construction of semantic mappings to a similar standard of knowledge engineers, and **RO5**: Evaluate the mapping tool with ordinary users over a long term period. The study evaluates whether it is possible for ordinary users to actively participate in the semantic mapping process within their own computing environment over time by analysing the mapping results generated by three user groups selected to have different levels of technical expertise. This field study was carried out by participants within their own browsing environments over a period of two weeks. The experiment also aided in addressing research objective **RO2**: Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts. The study investigated the behaviour of ordinary users with the tag-based mapping tool and the personal information delivery tool SportsFlows over time which allowed the collection of evidence for validation of the tool's approach to mapping and identification of weaknesses in the tool. This study was first presented in [Conroy 2009].

The structure of this chapter is as follows: first requirements for the evaluation of the tag-based mapping tool are specified, secondly the design of the study is discussed, thirdly the hypotheses of the experiment are stated, fourthly the methods for the experiment are explained, fifthly the results of the experiment are presented, sixthly the analysis of the experiment is detailed, seventhly the findings of the experiment are described, finally the limitations of the experiment are given.

## 6.1 Requirements for Evaluation

Several mapping requirements have been specified for modifying the mapping process and tools to allow ordinary users to engage in act of mapping, see section 2.8. These requirements have been updated throughout the thesis. Figure 6-1 displays this list of requirements. The specific requirement for the evaluation of the tag-based mapping tool is highlighted in bold.

| **Eliminate** | **Reduce** |
| --- | --- |
| Any Matching API Configuration by User<br>Visual Emphasis on Global View | Complexity of Mapping Process |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | Iterative Mapping Process<br>Familiar & Rich Mapping Interaction<br>Easy to Understand Interface<br>**Evaluation Methods** |

***Figure 6-1:*** *Specific Requirements for the Evaluation of Tag-Based Mapping Tool*

Evaluation methods need to be devised to evaluate whether the tag-based mapping tool is convenient, accessible, simple, efficient, effective and satisfactory for ordinary users. The methods and metrics for each of these evaluation areas are described and explained in section 6.5.

# 6.2 Study design

In this study, the mapping interactions of three different user groups are observed, recorded and evaluated. The user groups were:

- **Ontologically aware** users are those who have experience using or designing ontologies. These users are representative of the current users of mapping tools (knowledge engineers) and the participants typically used in other user mapping studies, see section 2.7. Thus this group reflects the skill level of knowledge engineers and was chosen to be used as a control group for the study.
- **Technical aware** users are those who have database/UML modelling experience but no ontology experience. This group have modelling skills which can be translated into the skills needed for mapping tools. In the author’s opinion, this group also represents a high percentage of web users who are the early adopters of new technologies. This group was chosen to examine how significant the differences are the interactions between ordinary users and users with modelling skills. This group was also chosen to be used as a control group for the study.
- **Non-technical** users are those who have basic computer or web experience with no database/UML modelling or ontology experience. These users have the lowest level of

technical experience and have no relevant modelling skills. However some of these users may be capable of/trained in abstract thinking which gives them necessary experience needed for modelling skills. This group is taken to be a proxy for ordinary users.

For convenience the outline for the evaluation approach is displayed again in table 6-1, from section 1.5.2.

**Table 6-1:** Experiment evaluation outline with current experiment (2) highlighted

| **Evaluation Category** | | **Experiment 1** | **Experiment 2** | **Experiment 3** |
|---|---|---|---|---|
| **Efficiency** | | Yes | Time Taken (section 6.5.1) | Yes |
| **Effectiveness** | | Yes | Precision & Utility (section 6.5.2) | Yes |
| **Accessibility & Convenience** | ***Unintrusive*** | No | Specific Metrics (section 6.5.3.1) | Yes |
| | ***Engaging*** | No | Specific Metrics (section 6.5.4.1) | Yes |
| **Simple** | | No | Specific Metrics (section 6.5.5.1) | Yes |
| **User Satisfaction** | | Yes | Questionnaire & Interview | Yes |

As mentioned, in section 2.7, there has been a distinct lack of user studies conducted on mapping tools. In particular, hardly any research has been carried out on mapping undertaken over time with mapping tools within the user's own work environment. This study is designed to investigate ordinary users using the mapping tool within their own browsing environment over time. The study is a *field experiment* where the experiment takes place in a naturally-occurring environment for participants[29]. This allowed the experiment to occur within the **user's own computing environment without supervision**.

Specifically, in this study the following questions are investigated:

**Q1:** Can non-technical users complete a set of mapping tasks to the standard of ontologically aware and technically aware users over time within their own work environment given appropriate tool support and an awareness of the domain?

**Q2:** Can the semantic mapping process be made non-invasive for non-technical users while occurring over time within their own work environment?

[29] http://en.wikipedia.org/wiki/Field_experiment

**Q3:** Does using a tagging paradigm simplify the mapping task interaction?

**Q4:** Does using a tagging paradigm maintain the quality of mappings made?

**Q5:** Does showing the benefits from mappings help to engage non-technical users to participate in the semantic mapping process?

## 6.3 Hypotheses

From the questions stated above, several hypotheses were formulated to be evaluated.

**H1**: There will be no significant difference in mapping performance between user groups with different technology expertise, derived from **Q1**;

**H2**: The mapping process will not disrupt non-technical users from their daily work, derived from **Q2**;

**H3**: Non-technical users will enter into mapping sessions and there will be no significant degradation of the number of sessions entered in the second half of the experiment, derived from **Q2**;

**H4**: Mapping sessions will not be time-consuming for the non-technical users, derived from **Q2**;

**H5**: Mapping tasks will be viewed as simple to accomplish for non-technical users, derived from **Q3**;

**H6**: Non-technical users will be able to achieve a high mapping precision derived from **Q4**;

**H7**: Non-technical users will finish mapping sessions they enter into and there will be no significant degradation of the number finished in the second half of the experiment, derived from **Q5**;

**H8**: Non-technical users will see the benefits generated from the correspondences they have developed, derived from **Q5**;

**H9**: Seeing the benefits, from **H8**, will then give non-technical users an incentive to continue participation in the semantic mapping process, derived from **Q5**.

Unlike the other hypotheses, which are evaluated from the user interactions and feedback with the tag-based mapping tool, hypothesis 8 is evaluated from the user interactions with the personal information delivery tool SportsFlows and feedback from the participants.

## 6.4 Method

This section gives the details of how the experiment was carried out. Firstly the details of the participants of the experiment are given. Next the materials used in the experiment are described. Afterwards the procedure of the experiment is detailed. Then a summary is given of the data collected in the experiment. Finally the analysis methods used in the evaluation are explained.

### 6.4.1 Participants

Three different groups of users were used in the experiment: *Ontologically aware* users who have ontology work experience. These participants were PhD students from the KDEG research group. This group is labelled in diagrams as "onto aware". *Technically aware* users who have database/UML modelling experience but no ontology experience. These participants were undergraduate students and postgraduate students who have either completed or were currently doing a computer science degree. This group is labelled in diagrams as "tech aware". *Non-Technical* users who have basic computer experience with no database/UML modelling and ontology experience. These participants were volunteers from the university sports centre who had no computer science background and only had basic computing experience. This group is labelled in the diagrams as "non-tech".

12 participants were recruited for the study with 4 participants for each user group. The low numbers used meant that we could gain more detailed feedback and analysis from each participant through post experiment interviewing techniques.

### 6.4.2 Materials

The experiment involved the participants using the SportsFlows tool within their browser on their own personal computer. The participant first filled in a pre-questionnaire which was used to assess their technical skills and thus which user group the participant would belong to. Then the participant constructed their personal ontology by selecting from a list of sport interests, see section 5.2.3. The tool then personalised the RSS sports shown to the participant based on the mapping that the participant constructed. The tag-based mapping tool displays prompts for mapping sessions three times during each day, see section 5.1.4.1, with each mapping session containing three mapping tasks, see section 5.1.4.2. These numbers were derived from trial and error with test users.

28 candidate correspondences were generated using the INRIA Alignment API [Euzenat 2006], although 6 had to be removed as they would confuse ordinary users, see section 3.7.3, for example ‘Individual’ to ‘Basketball’. 10 missing correspondences were added by the author. The total number of candidate correspondences required to be validated by the participants during the course of the experiment was 32. The reason for the low number of correspondences is that it is anticipated that the user’s sporting interests will be relatively shallow and that any mapping tasks will be done over time in small sessions as their interests vary and change. A gold standard for the candidate correspondences[30] was produced by the author and was verified by a knowledge engineer from the KDEG research group with domain knowledge of the sporting domain. The gold standard is provided in the accompanying DVD media under Experimental Datasets - Experiment 2. The gold standard was used to validate the mappings generated by the participants. Questionnaires were used to evaluate the user’s satisfaction with the mapping tool. In particular the questionnaires focussed on measuring whether mapping was unintrusive, engaging and simple enough for users. There were two questionnaires used. The first was asked halfway through the study while the second questionnaire was asked post study, see Appendix D for the questionnaires.

### 6.4.3 Procedure

Each participant was required to use the SportsFlows tool over a two week period on their own personal computer[31]. It included several steps. The participants were first asked to fill in a *pre-study questionnaire* which included several questions about their background to assess their technical skills. An *introduction to the tag-based mapping tool* was given to each participant describing how to use the tool and what they were required to do during the course of the experiment. The participants were shown the tagging interface and explained how to interact with the tool. Each participant was also required to select their personal interests from a list of sport concepts to form their ‘concept interests’ (i.e. the subset of the sporting ontology concepts that were of interest to them). After the introduction was finished the participants were asked to *use the tool* over a two week period to construct mappings and view any sports story that interested them. While the participants were using the tool all their actions with the tool were logged through Ajax calls to the mapping server, see table 6-2. During the two weeks each

[30] http://www.cs.tcd.ie/sports/goldstandard
[31] The experiment was carried out before an ethical approval process had been introduced to the school

participant was asked to answer the mapping session prompts when convenient and to look at any RSS feed items they found interesting. The participants were required to answer a *mid experiment questionnaire* to give their initial thoughts on the tool, which were focused on the usability of the tool, e.g. 'Are the questions and supplied information confusing?'. A four point scale was taken to allow users to express different levels of satisfaction. After the experiment, participants filled out a *post-study questionnaire* to evaluate their final thoughts. Again the questionnaire focused on the usability of the tool with a four point scale. The same questions were asked from the mid experiment questionnaire and other questions were added, e.g. 'How abrupt was the timing of the question prompts?'. Finally an *interview* was conducted with each participant to allow them to give their general impressions and reactions to the tool. The survey answers were taken into account during each interview. Some of the questions asked were: 'how simple was the mapping task', 'how engaging was the mapping task' and 'how interruptive were the mapping tasks', i.e. how intrusive were the mapping task to the participant.

### 6.4.4 Data Collection

The following data was collected: (a) *Pre-experiment Questionnaire* – used to evaluate which user groups the participants belong to, (b) *Mid experiment Questionnaire* – used to examine the initial usability and usefulness of the tool, (c) *Post experiment Questionnaire* – used to examine the long-term usability and usefulness of the tool (see Appendix D for questionnaires), (d) *Interviews* – used to record the in-depth account of the participants experience with the tool, (e) *Log of users actions* – used to track the participant's behaviour with the tool in particular how active they are in the mapping process, (f) *Resultant Mappings* – the quality of the mappings were analysed to give the precision for each group. The number of participants (12) was relatively small but the number of events was much larger (2,538 total giving 21 events per day for every user) giving a reasonable amount of data to analyse and this allows us to draw several indications and conclusions.

### 6.4.5 Analysis Method

The ontology aware group is a control group for the experiment as they reflect the technical experience of current users of mapping tools. The technically aware group reflect users with modelling skills needed for mapping tools and is also used as a control group. The quantitative data analysed were the mapping task time, mapping precision, numbers of mapping sessions entered and mapping session completed. To evaluate the differences in mapping results between each user group an unpaired two-tailed t-test was used. Any significant result from these tests will be highlighted in the evaluation while the full results can be found in Appendix A. The qualitative data gathered also helped in revealing the users' opinions about the tag-based mapping tool.

In the evaluation section, if there was little or no difference in the response given to the same questions asked in both the mid-experiment questionnaire and post questionnaire it is not highlighted or mentioned. The entire set of anonymised data records are provided in the accompanying DVD media under Experimental Datasets - Experiment 2. The mapping interaction results were measured against the overall browsing time taken by the participant's during the course of the experiment. Since the experiment was undertaken over a two week period, the mapping interaction results could also be measured weekly. However since there were only 32 candidate correspondences, many of the participant's finished all of the mapping tasks early in, or before, the second week and this would skew the results if the results were analysed on a weekly basis. As an alternative, since the total browsing time while creating the mapping was recorded, the browsing time was separated in half and the mapping interaction results was measured between each half to examine if there was any improvement over time. Note that when the participant has finished all of the mappings tasks for the experiment, the browsing time being recorded stops. However, for the mapping precision results the total number of correspondences mapped by the participants was split in half and compared. This was carried out to observe if there were any improvements in the mapping results after half of the correspondences were mapped by the participant.

## 6.5 Evaluation

This section outlines the results from the experiment. The evaluation is split into the five areas, displayed in table 6-1, as following: Firstly, the efficiency of the tag-based mapping tool is calculated. Next the effectiveness of the tag-based mapping tool is measured. Following this the non-invasiveness of the tag-based mapping tool is evaluated. Afterwards the engagement of the tag-based mapping tool is measured. Finally the simplicity of the tag-based mapping tool is analysed.

Table 6-2 below gives a snapshot of the experimental results. Furthermore, the entire set of anonymised data records are provided in the accompanying DVD media under Experimental Datasets - Experiment 2.

**Table 6-2:** Overall mapping results for each user group. A double dash, --, indicates there is no value in the cell.

| User | Browser Time (HH:MM:SS) | RSS Viewed | Correspondences | | | | Mean Mapping Task Time (Seconds) | Sessions Prompted[a] | Sessions Entered[b] | Sessions Finished[d] | Sessions Abandoned[e] | Sessions Not Started[g] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Undertaken | Correctly Categorised | Incorrectly Categorised | Inconclusively | | | | | | |
| **Non-tech 1** | 10:40:00 | 13 | 32 | 24 | 8 | 0 | 27 | 20 | 13 | 11 | 0 | 2 |
| **Non-tech 2** | 12:04:06 | 56 | 5 | 3 | 2 | 0 | 61 | 51 | 16 | 1 | 2 | 13 |
| **Non-tech 3** | 03:48:03 | 16 | 6 | 1 | 0 | 5 | 41 | 8 | 2 | 2 | 0 | 0 |
| **Non-tech 4** | 03:28:07 | 85 | 0 | -- | -- | -- | -- | 12 | 0 | 0 | 0 | 3 |
| | | | | | | | | | | | | |
| **Onto Aware 1** | 07:27:20 | 34 | 32 | 22 | 1 | 9 | 37 | 38 | 15 | 11 | 1 | 3 |
| **Onto Aware 2** | 04:43:47 | 8 | 32 | 19 | 13 | 0 | 53 | 18 | 14 | 10 | 1 | 3 |
| **Onto Aware 3** | 29:28:16 | 3 | 18 | 14 | 4 | 0 | 22 | 138 | 14 | 5 | 3 | 6 |
| **Onto Aware 4** | 08:16:26 | 29 | 27 | 12 | 15 | 0 | 61 | 26 | 12 | 9 | 0 | 3 |
| | | | | | | | | | | | | |
| **Tech Aware 1** | 04:23:23 | 10 | 32 | 9 | 0 | 23 | 37 | 41 | 14 | 11 | 0 | 3 |
| **Tech Aware 2** | 11:40:43 | 34 | 32 | 21 | 9 | 2 | 40 | 16 | 14 | 8 | 5 | 1 |
| **Tech Aware 3** | 06:28:06 | 64 | 32 | 19 | 2 | 11 | 20 | 22 | 14 | 10 | 1 | 3 |
| **Tech Aware 4** | 21:06:33 | 21 | 28 | 11 | 6 | 11 | 17 | 47 | 10 | 9 | 1 | 0 |

[a] The number of mapping session alerts which were prompted to the user
[b] The number of mapping sessions entered by the user
[d] The number of mapping sessions which were finished by the user
[e] The number of mapping sessions exited by the user after they completed a least one mapping task
[g] The number of mapping sessions exited without answering a mapping task

### 6.5.1 Efficiency

The metric taken to measure the efficiency of the tag-based mapping tool is the time taken to complete a mapping task. This is the commonly used metric to evaluate the efficiency of mapping tools, see section 2.7. Figure 6-2 below gives the results for the mean time taken to map a correspondence for each user group.

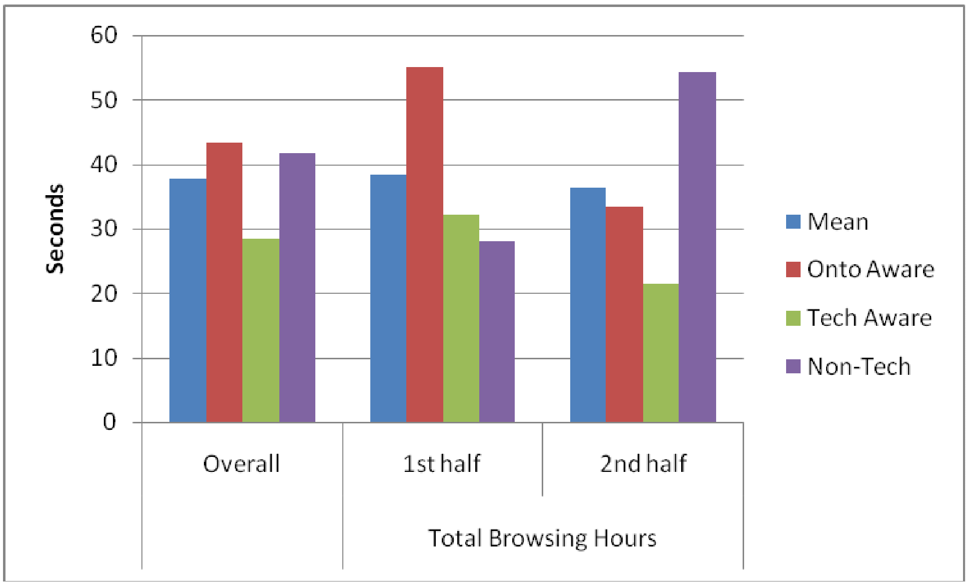



***Figure 6-2:*** *Mean mapping task time for User Groups*

The ontologically aware group had a mean of approximately 43 seconds, while the technically aware group had a mean of approximately 28 seconds and the non technical group a mean of approximately 42 seconds. These results suggest that non technical users do not take longer to complete a mapping task than ontological aware users or technically aware users, see Appendix A. The results also indicate that the mapping sessions are not time-consuming with a mapping task taking on average 38 seconds and a mapping session taking approximately 2 minutes. The mean task time reduced slightly (by 2 seconds) over the second half of the experiment, however with the non-technical aware group the time doubles. This is an alarming increase and is discussed in section 6.5.3.2.

### 6.5.2 Effectiveness (Quality of Mapping)

Since the mappings generated by users can have both correspondences mapped with suggested relations and with new user-defined relations, the effectiveness of the tag-based mapping tool is evaluated with two metrics: (1) the precision of correspondences that were generated with suggested relations or a combination of user-defined relations and suggested relations, (2) the utility of tags for correspondences mapped with **only** user-defined relations. In total there were 276 correspondences generated by the participants. 78% of these correspondences were generated with suggested relations while 22% were generated with new user-defined relations.

Mapping recall is not used as a metric in the evaluation of the effectiveness of the tag-based mapping tool. Mapping recall evaluates the overall percentage of valid correspondences in the set of correspondence mapped by the user and is primarily used to evaluate the match algorithms. The tag-based mapping tool only prompts candidate correspondences to the user and does not allow the user to append any missed correspondences. Thus the recall results would only evaluate the algorithm, which not the purpose of experiment and thesis. The recall results can also be used to measure the number of valid candidate correspondences mapped but this is just a subset of the overall precision results. Furthermore none of the user mapping studies conducted by the state of the art have evaluated the mapping recall, see section 2.7. The inability to allow the user to append any missed correspondences with the tag-based mapping tool was discussed in section 5.1.5.1.

#### 6.5.2.1 Precision of Suggested Relations

Figure 6-3 shows the mapping precision results, for the correspondences mapped with suggested relations, for each group. The precision was calculated by comparing the mappings constructed by users to the gold standard. If the mapping category of a correspondence is not the same as in the gold standard, the correspondence would be regarded as incorrect. The results showed that overall 76% of the correspondences mapped were deemed to be accurate.

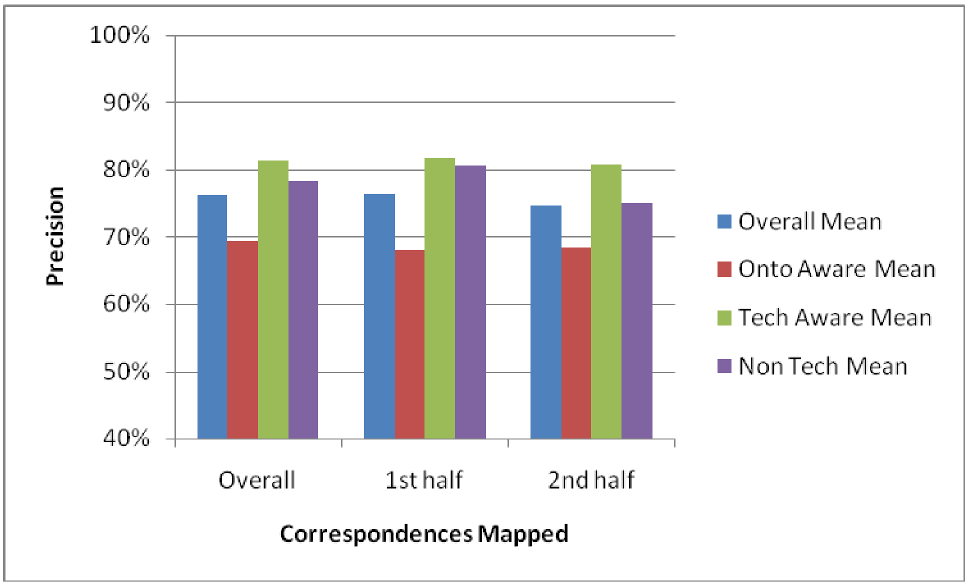



***Figure 6-3:*** *Mapping precision of correspondences mapped with suggested relations for User Groups*

Looking through each group shows the technically aware group achieved the highest accuracy with approximately 81% while the non technical group achieved 78% and the ontologically aware group had an approximate accuracy of 69%. These results suggest that non technical users mapped correspondences to the same standard as ontological aware users and technically aware users, see Appendix A. The results showed surprisingly that the ontologically aware group obtained the least precision. However examining the incorrectly categorised correspondences for the ontologically aware group revealed that in 25% of these correspondences correct tags were entered but additional incorrect tags were also entered which caused the correspondences to be incorrectly categorised. In addition, there is also a significant error rate of at least 19% for each group. Examining all of the incorrect correspondences revealed that 73% were correspondences which should have been categorised as different but had an association or relationships between the concepts, e.g. Golf and Golf Course, and were categorised by the user as Equivalent or Equivalent Sometimes. The majority of the others were correspondences in which the two concepts had a common parent concept and should have been categorised as different, e.g. Golf and Soccer, but were mostly categorised as Equivalent. Examining all of the suggested relations used in the incorrectly mapped correspondences showed:

- 43% were relations which, in the interviews, the participants mentioned they thought the relations had different meanings than the intended meaning, e.g. 'part of' was meant to be used to represent the union set theory relation but instead it was understood to be a way to express a concept to property correspondence. For example with the correspondence 'Golf Course' to 'Golf'.
- 27% were an 'equivalent sometimes' relation for correspondences which should have been mapped as 'equivalent'. For example 'equivalent sometimes' was used for the correspondence 'Football' to 'Soccer'.
- 15% were an 'equivalent sometimes' relation for correspondences which should have been mapped as 'different' but had an association, i.e. 'a subclass of' used for the match 'Basket' and 'Basketball'.
- 8% were relations which suggest the user misunderstood the descriptions of the concepts being matched. For example 'the same' used for 'Football', describing soccer, and 'American Football'.
- 7% were an 'equivalence' relation for correspondences which in fact had a association, e.g. 'the same' was used on the correspondence 'Wimbledon' and 'Wimbledon Trophy' where a 'different' relation or a user-defined relation like 'has trophy' would have been more appropriate.

The majority of these points indicate users made an incorrect choice of relation which could be due to certain relations being unclear, e.g. 'part of', while in other cases it may be due to users being unsure of what relation to use. It should also be noted that some users may not be inclined to add their own user defined relations and just preferred to click on a suggested relation that seemed appropriate to them. In the experiment 25% of the participants did not add any user-defined relation, with 50% of non technical users only choosing suggested relations. In contrast to the other points, the third point listed above suggests that some users were just tagging the relation based on the concept names and not the natural language description provided. This could be due to the descriptions being unclear or users not being inclined to read the description. These are important issues for mapping interface design and suggest that only standard relations that should have a clear meaning should be displayed initially to users, and any user-defined relations they use will then be appended, each of the user-defined relations will have a personal meaning to them. It would also be beneficial to track the use of the mappings to make sure they are categorised correctly to see if users are using suggested relations in a personalised way, i.e. evaluate if the correspondences are showing benefits to the user by monitoring the use of the correspondences.

#### 6.5.2.2 Utility of User-Defined Relations

Investigating how many user-defined relations were added showed that: 38% of the correspondences mapped by the technical aware group were correspondences with new user-defined relations, while the non-technical group had 12% and the ontological group was at 8%. This indicated that the technical aware group was more inclined to add user-defined relations than the other groups. The author reviewed the set of user-defined relations added by the users. The results revealed that 97% of the user-defined relations added made sense in defining a relationship and 92% accurately defined the relation between the concepts. The user-defined relations were categorised into the following categories:

- *Miss-spelt tags (6%)*: tags which are spelt incorrectly, for example 'drifferent',
- *Meaning of a suggested tag (39%)*: tags which have the same meaning as a suggested relation, for example 'more general than',
- *Tag giving semantic meaning (27%)*: tags which correspond to a property, for example 'has a league',
- *Alert to bad correspondence (10%)*: tags users have used to indicate that the suggested candidate correspondence is wrong, for example 'unrelated',
- *Other (18%)*: tags which indicate a general correspondence, for example 'similar'.

### 6.5.3 Unintrusive

In this section, first the metrics used to quantify an unintrusive mapping tool are given. Secondly, the analysis from the results is presented. Finally, some conclusions from this evaluation are given.

#### 6.5.3.1 Metric Definition

The term *unintrusive* is defined here to mean that the mapping tool will not disrupt the user from their daily work and that the process will blend into the background while they work, i.e. when doing mapping tasks regularly over the time period it does not hamper their work and that in fact they see it as a commonplace task. The quantitative metrics and qualitative feedback for unintrusive tasks are defined by the author with rationale for each as follows:

- **Quantitative Metrics**
  - *Percentage of mapping sessions entered*: an unintrusive task will have a high rate of response to task prompts.
  - *Mean mapping session time*: an unintrusive task will be quick to complete.
- **Qualitative Feedback**
  - *Responses to the questions about "How disruptive were the mapping tasks to your work?" measured through questionnaires*: an unintrusive task will not be perceived as disruptive to the user.
  - *Responses to the questions about "How interfering were the mapping tasks to your work?" measured through questionnaires*: an unintrusive task will not be perceived as interfering by the user.
  - *Responses to the questions about "How interruptive was the mapping task?" measured through interviews*: an unintrusive task will not be perceived as interruptive by users.

#### 6.5.3.2 Analysis

To assess whether participants found the tag-based mapping tool unintrusive, the number of mapping session prompts clicked on and ignored were tracked. The results, in Figure 6-4, give an indication of the participation rate of each user group and to a lesser extent the perceived utility of the task for participants within each group.

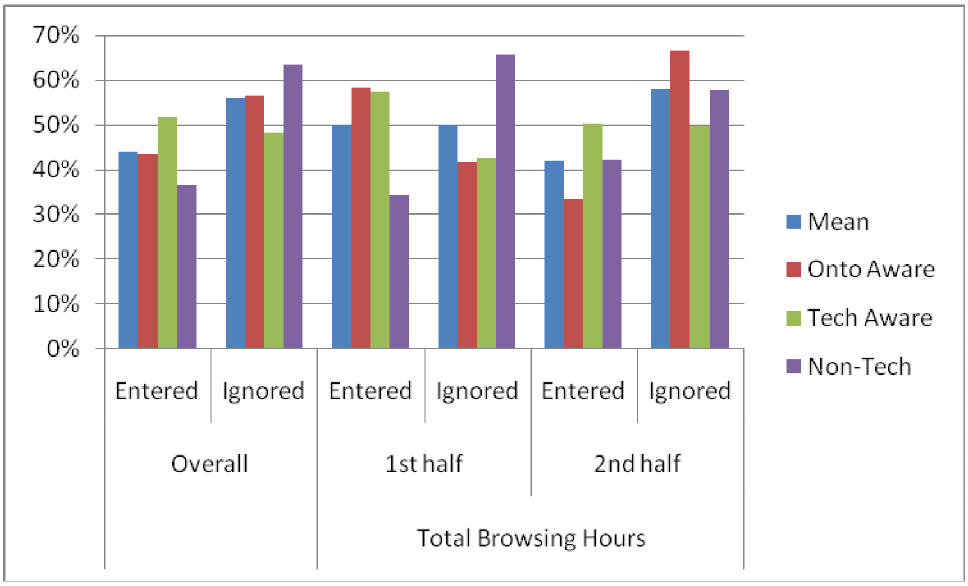



***Figure 6-4:*** *Mapping Session Prompts Results for User Groups*

The results show on average that almost half of the prompts issued were entered (approximately 45%). There is a slight increase (<10%) in the number of prompts ignored in the second half of the experiment. However, the entered rate is still quite high in the second half (approx 42%) which indirectly indicates that the participants did not generally find the mapping sessions to be overly inconvenient in their daily work, otherwise one might expect there to have been a greater drop in the number of sessions entered in the second half of the experiment. Furthermore the results suggested that the non technical users entered mapping sessions to the same standard as ontologically aware users and technically aware users, see Appendix A.

From section 6.5.1, the mean time taken for each user group to map correspondences indicates that the mapping sessions are not time-consuming. However, while the mean task time reduced slightly (by 2 seconds) over the second half of the experiment, for the non-technical group the mean task time doubles. This is an alarming increase. However investigating the number of correspondences mapped by non-technical users, see Figure 6-5, might suggest that this increase in part could be attributed to the low number of correspondences mapped by the users. The non-technical group only completed 45% compared to 85% for ontologically aware group and 97% technically aware group. In the final questionnaire, a majority (73%) felt that the application blended in to the background while they were undertaking their daily work. All the users felt the tag-based mapping tool was not disruptive to their daily work, rising from 92% of users in the mid-point survey. Finally, when asked about how distracting the mapping session prompts were, 93% of respondents said they did not find the mapping session prompts to be disruptive.

When looking at participation of the groups individually in entering mapping sessions (Figure 6-3), it shows that the ontology aware group's 'ignore rate' increases sharply in the second half of the experiment (by approximately 25%) which is in contrast to the non-technical users who have a reduction in the number of prompts ignored in the second half of the experiment (by approximately 8%). When asked in the interviews why they were ignoring the mapping session prompts, users answered that a lot of the prompts were displayed when they were busy with a daily task which would lead to the user ignoring them. Some of the comments given by participants were: '*The mapping tasks were not overly interruptive just when I was working it was, it didn't bother me at any other time*' and also '*When I am doing something and I did not want to be interrupted it was annoying but when I am just browsing the web it is fine. Just when I am concentrating it is not*'. The feedback given suggests that the participants would have preferred mapping tasks to be asked in a more suitable context, e.g. when they are not busy.

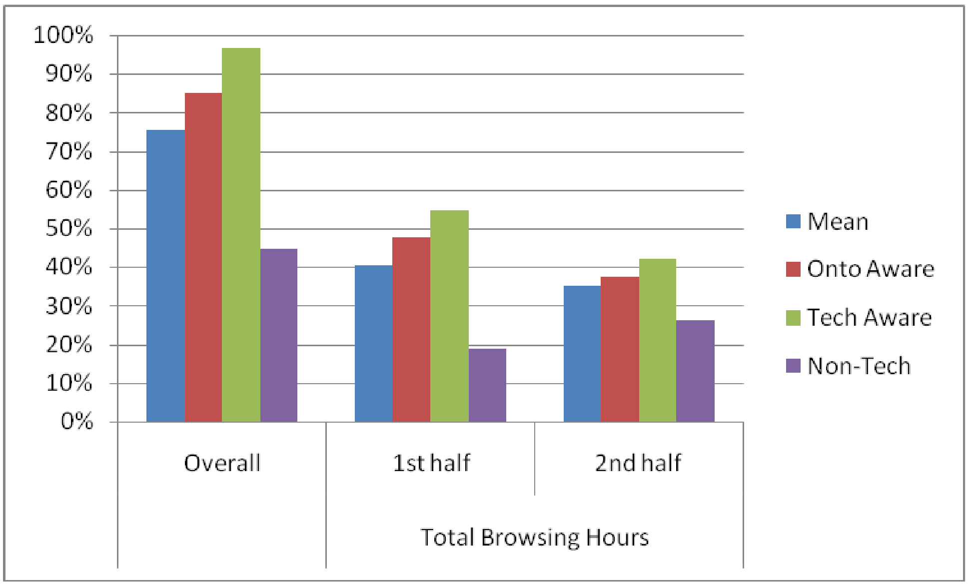



***Figure 6-5:*** *Percentage of the correspondences mapped for the User Groups*

#### 6.5.3.3 Conclusion

There was a general feeling from the participants that mapping sessions would be better if they were requested in a browsing context suited for each individual user. Participants also suggested that mapping tasks should be able to be requested by the user at any time when there are tasks available. Overall given the feedback from the participants and the results obtained, it seems to suggest that the mapping tasks were not hampering the participant's daily work and the tasks were blending into the participants every day work. Although when the prompts were displayed at times when the participants were busy they found it to be annoying and disruptive.

### 6.5.4 Engaging

In this section the metrics used to quantify an engaging mapping tool are first given, then the evaluation of the results is presented, finally some conclusions for this evaluation are given.

#### 6.5.4.1 Metric Definition

The term *engaging* is defined here to mean users will enter and finish mapping sessions regularly over time seeing the benefits of the mappings accomplished, i.e. personalised RSS sport news stories. The specific quantitative metrics and qualitative feedback for engaging tasks were defined by the author with a rationale for each as follows:

- **Quantitative metrics**
  - *Percentage of mappings completed*: an engaging task will have a high rate of mappings done.
  - *Percentage of mapping sessions entered*: an engaging task will have a high rate of response to task prompts.
  - *Percentage of mapping sessions finished*: an engaging task will have a high rate of mapping sessions finished.
  - *Number of RSS items looked at*: an engaging task gives benefits to the user in a functional, social or emotional dimension. With the SportsFlows application the user will view a high number of RSS sport stories.
- **Qualitative feedback**
  - *Responses to the questions about "How appropriate were the sports information and did this help motivate you to answer more tasks?" measured through questionnaires*: an engaging task will lead to appropriate benefits for the user and help motivate the user to engage in more tasks.
  - *Responses to the questions about "How engaging were the mapping tasks?" measured through interviews*: an engaging task will lead to users answering that they felt engaged to complete the task.

#### 6.5.4.2 Analysis

In the author's opinion, the number of RSS stories looked at by the participants, at an average of 3 stories viewed each day see table 6-2, suggests that the participants were engaging with the tool and might indicate that they were seeing the benefits of mapping.

From Figure 6-5 the overall number of correspondences mapped is quite high at 75% done. However looking at the results in depth there is a minor drop in the correspondences mapped in the second week but even more worrying is that the non-technical group only achieved 45% of the correspondences in total. The low percentage of correspondences undertaken by non-technical users can be attributed to the low number of hours browsing by the users over the course of the experiment, average of 7 and a half hours. The low number of browsing hours led to less mapping prompts being displayed to these users. Contrasting this with the technically aware users with average of 17 hours and ontologically aware user with average of 13 and a half hours shows the non-technical group had the browser open for less than 50% of the time of technically aware users and for less than 60% of the time for ontologically aware users. By considering the average number of correspondences undertaken per browsing hour it shows that 6% are done by non-technical users, 5.71% are done by technically aware users and 6.30% are done by ontologically aware users. In addition, a mean of 22.75 mapping session prompts were displayed to non-technical users while a mean of 55 mapping session prompts were displayed to ontologically aware users. This suggests that the non-technical users were quite competitive when compared to the two other groups when taking browsing statistics into account.

Every time a user responded to a mapping session prompt their actions were recorded which allowed for the calculation of the number of mapping sessions finished (Fin), abandoned (Abd) and not started (Non). A "not started" mapping session is one in which no more user events are recorded, this could be due to system unreliability as it is known that in some cases that the connection to the mapping server was lost, two of the participants mentioned this happened twice each (4 total) and it only happened in the first week of the experiment, or if a user simply ignored the session after it starts. These results give an indication whether the users are engaged in the mapping process and also reveal if the mapping task is perceived as too complicated or too time-consuming. In the results, see Figure 6-6, 8% of the sessions were abandoned which is quite low and might be explained by the user suddenly getting busy, e.g. they just got an email or they got a message on internet messenger. More alarming and interesting is that 30% of the mapping session entered were "not started", i.e. the user exited immediately without finishing a mapping task or logging failed as described above.

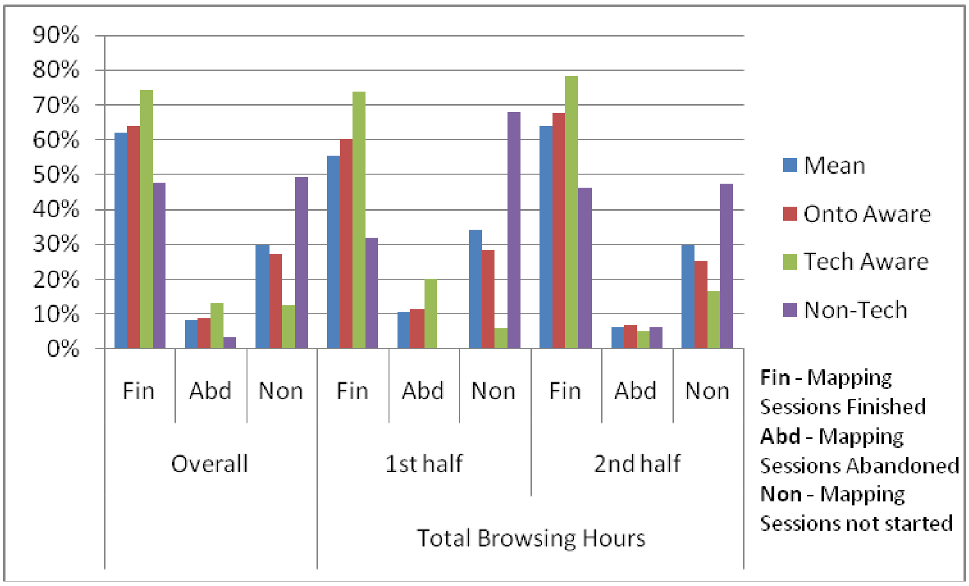



***Figure 6-6:** Breakdown of mapping sessions entered for User Groups*

During interviews, four of the participants answered that they exited the mapping task when it was asking them something they were not interested in, i.e. if a user had not included football as an interest then they were perplexed as to why it was being asked as a mapping task and chose not to answer. It should be noted that each participant was asked to map all candidate correspondences even the ones not related to their interests. This can potentially have a negative effect on engaging results and even the usability of the tool. A comment given by a non-technical user was '*I was engaged with the tasks when it was something I was interested in but when it was something I was not interested in I did not want to do it*'. Also three of the users mentioned that the submit and close mapping session buttons were placed too close together on the screen and that they accidentally pressed the close button when submitting tags for a mapping. Furthermore the results suggested that the non technical users entered and finished mapping sessions to the same standard as ontologically aware users and technically aware users, see Appendix A

In the survey only 25% of the users found that answering the mapping tasks tailored the sports news being provided to them. Instead a significant 42% found it did nothing at all and 33% said it only slightly changed. With 75% of the users not seeing the sport news getting more tailored over time, it would indicate that most users are not seeing the benefits that the mappings should be bringing. A proposed reason as to why some users did not see benefits may be due to the important correspondences, the correspondences which affect the users interests the most, being mis-categorised through the wrong choice of tags or undetermined tags being used. Through incorrectly mapped correspondences several stories not relating to the users interests may be shown which could lead to a user's frustration with the tool. Recall that 25% of the mappings were incorrectly classified with respect to the gold standard and this gives some supporting evidence for this proposed reason, see Figure 6-2. In the interviews the majority of the participants supported this reason by mentioning that after mapping there was still incorrect information shown for their interests. This indicates a need to improve the RSS reader of the SportsFlows application to only show information relevant to the user's interests.

#### 6.5.4.3 Conclusion

An issue that that was noticed in the experiment was keeping users engaged in the mapping process. It was a clear problem in the experiment that by prompting some users with mapping tasks that were not related to their interests that they became less engaged. The same mapping tasks were asked to each participant as their personal sporting interest were modelled using the same ontology and the number of candidate correspondence generated was too low to only allow users to answer candidate correspondences that effected them, see section 6.4.3. Thus each participant was asked to map each candidate correspondence. In future, only mapping task related to the users interests will be asked. Also there was a problem when correspondences were incorrectly mapped, this may lead to no benefits being observed by the user and thus lead them to become less engaged. In the author's opinion, decomposing the mapping process into small tasks helped to engage the user in the mapping process, as each task takes little time to do, and it also let the user see the feedback of their actions between sessions. Another potential reason for the participants' lack of motivation to use the tool may have been due to their lack of interest in sports. Additionally the participants may not have noticed the prompts. Although it was shown that a user can become engaged in the mapping process at the start it is important to show the benefits and just prompt with correspondences that relate to their individual interests in order to keep them engaged in the process for a longer time period. It should also be observed that the tag-based mapping tool only produces a partial mapping, which is continually updated by the user over time, and this could have also potentially affected the user's participation.

### 6.5.5 Simple

In this section the metrics used to quantify a simple mapping tool are first given. Secondly the analysis of the results is presented. Finally some conclusions from this evaluation are given.

#### 6.5.5.1 Metric Definition

The term *simple* is defined here to mean users will be able to achieve rich mappings in a non time-consuming way that is not confusing but instead is straightforward to the user. The specific quantitative metrics and qualitative feedback for simple tasks are defined by the author with rationale for each as follows:

- **Quantitative metrics**
  - *Mean mapping session time*: a simple task will be quick to complete.
  - *Percentage of mappings done*: a simple task will have a high rate of correspondences mapped.
  - *Mappings accuracy with respect to the gold standard*: a simple task will have a high creation rate of correctly mapped correspondence.
- **Qualitative feedback**
  - *Responses to the questions about "Are the questions and supplied information confusing?" measured through questionnaires*: a simple task will not be confusing to the user.
  - *Responses to the questions about "How simple was the mapping task?" measured through interviews*: a simple task will lead to users answering that they found the task straightforward to achieve.

#### 6.5.5.2 Analysis

The efficiency evaluation, see section 6.4.1, indicate that the mapping sessions are not time-consuming with a mapping task taking on average 38 seconds and a mapping session taking approximately 2 minutes. The unintrusive evaluation, see section 6.5.3.2, revealed that the mean percentage of candidate correspondences achieved by participant was high at 73%. The non-technical user group percentage (43%) was lower compared to the ontologically aware user group (85%) and technically aware user group (97%) but this was due to the low browsing hours by these users leading to less mapping session prompts being displayed to these users, see section 6.5.4.2. The effectiveness evaluation, section 6.5.2, showed the mean mapping precision was approximately 75%. This result is still comparative with the previous experiment (88% with COMA++ and 79% with NL prototype), see section 3.6.2, and the only other published user study on ontology mapping tools which analysed the precision result (CoGZ achieved precision of 1.69 (84.5%) and PROMPT a precision of 1.58 (79%)) [Falconer 2009]. The user-defined relations were shown to be useful for defining relationships with 92% of these relations accurately defining the relationship of the correspondence, see section 6.5.2.2

In the final questionnaire, a majority (83%) mention question and supplied information was not confusing. In the interviews users were asked how simple they found the mapping task. The general response from each group was that they found the mapping tasks very straightforward and not difficult at all. Some of the comments given were '*it was very simple, I have many options and the choice was obvious*' and '*it was pretty simple just click on tags. I understood the information it was pretty straightforward*'. The participants also mentioned they were confused by some of the candidate correspondences generated, e.g. Basket and Motor Racing, which were very poor candidates for mappings. The users did not like being asked to map these types of correspondences. The author's opinion from the interviews conducted with the users was that the poor candidate correspondences being asked were putting off some of the users from contributing and answering more tasks. Finally everybody mentioned in their interviews that they liked the tagging approach as it gave more choices than just 'yes' or 'no' and also allow them to be expressive.

#### 6.5.5.3 Conclusion

The results indicated that the users mapped a large number of the candidate correspondences with a high level of precision and utility. The logs showed that the time spent by the users doing the mapping task was quite acceptable and it was not time consuming although the increase in time spent by non-technical users in the second half is strange and maybe due to the reduced number of mappings undertaken, see section 6.5.4.2. From the interviews and questionnaires it appears that most users find the mapping task straightforward and easy enough to do although there is a need to show high quality candidate correspondences that have something in common rather than low quality candidate correspondences, e.g. 'Basket' to 'Motor Racing'. Also there is a need to reduce the ontological information shown to a more readable natural language as some of the non-technical participants found the volume of information for the description confusing to read.

## 6.6 Analysis with respect to Hypotheses

This section gives a summary of the analysis of the experiment. Table 6-3 displays the overview evaluation of the experiment hypotheses, which is then discussed in the following subsections.

**Table 6-3:** Summary of the Analysis of Hypotheses

| **Hypothesis** | | **Suggests support for** | **Evidence** |
|---|---|---|---|
| **H1** | There will be no significant difference in mapping performance between user groups with different technology expertise | Yes | Section 6.5.1 -- 6.5.4 |
| **H2** | The mapping process will not disrupt non-technical users from their daily work | Yes | Section 6.5.3 |
| **H3** | Non-technical users will enter into mapping sessions and there will be no significant degradation of the number of sessions entered in the second half of the experiment | Yes | Section 6.5.3 |
| **H4** | Mapping sessions will not be time-consuming for the non-technical users | Yes | Section 6.5.1 |
| **H5** | Mapping tasks will be viewed as simple to accomplish for non-technical users | Yes | Section 6.5.5 |
| **H6** | Non-technical users will be able to achieve a high mapping precision | Yes | Section 6.5.2 |
| **H7** | Non-technical users will finish mapping sessions they enter into and there will be no significant degradation of the number finished in the second half of the experiment | Yes | Section 6.5.4 |
| **H8** | Non-technical users will see the benefits from correspondences they have developed | No | Section 6.5.4 |
| **H9** | Seeing the benefits, from **H8**, will then give non-technical users an incentive to continue participation in the semantic mapping process | No | Section 6.5.4 |

### 6.6.1 Mapping Performance of Non-Technical Users

The first question of the experiment to be investigated was *whether non-technical users can complete mapping tasks to the same standard as ontologically aware users and technically aware users*. The first hypothesis **H1** was designed to address this question. The participants were separated into three user groups (ontologically aware, technically aware, and non-technical) based on their technical expertise, see section 6.4. The results of the experiment revealed that there was no significant difference between the user groups for the time taken to complete a mapping task, see section 6.5.1, and the precision of the correspondences generated with suggested relations, see section 6.5.2.1. There was a difference in the number of candidate correspondences finished by non-technical users (45%) and both technically aware users (97%) and ontologically aware users (85%) see section 6.5.3.2. However the low number of correspondences completed by the non-technical user group was due to the reduced number of browsing hours by these users which impacted the number of mapping session prompts shown to the users, a mean of 22.75 in comparison to the mean of 55 for the ontologically aware user group, see section 6.5.4.2. Analysis of the mapping results of each user group showed that there was no statistical significant difference between the mapping results of the groups.

### 6.6.2 Mapping Process Intrusiveness

The second question of this experiment aimed at evaluating *whether the mapping process can be non-invasive to non-technical users when occurring over time within their own work environment*. The second, third and fourth hypotheses (**H2**, **H3** and **H4**) were designed to address this question. The second hypothesis stated that the mapping process will not disrupt non-technical users from their daily work. The results of the post-experiment questionnaire, see section 6.5.3.3, revealed that all of the participants did not find the mapping tasks presented to be disruptive and a majority (93%) did not find the tasks interfered with their daily work. This indicates that the mapping process was not hampering the user's daily work and indicated that the tasks were blending into the background. The comments given in the interviews supported this claim. The third hypothesis stated that non-technical users will enter into mapping sessions regularly over time. Analysis revealed that the number of mapping sessions entered was quite high with approximately 45% of mapping task prompts shown entered, see section 6.5.3.3. Indeed the percentage of sessions entered is still high at just over 40% in the second week demonstrating that the participants engaged in the mapping process regularly over time. The fourth hypothesis stated that mapping sessions will not be time-consuming for non-technical users. The efficiency evaluation, see section 6.5.1, showed that mapping sessions were not time-consuming to the non-technical users with a mapping task taking on average 42 seconds, and the mapping session taking on average just over 2 minutes.

### 6.6.3 Mapping Process Simplicity

The third question of this experiment investigated *whether using a tagging paradigm would simplify the mapping task*. The fifth hypothesis **H5** was tasked with addressing this question. The fifth hypothesis stated that mapping tasks will be straightforward for non-technical users to accomplish. The simple evaluation, see section 6.5.5, revealed that the users found mapping tasks straightforward and easy enough to do. In the post-experiment questionnaire, the majority (83%) mention question and supplied information was not confusing, see section 6.5.5.2. The comments given in the interviews also supported the claim that non-technical users found the tasks simple to accomplish.

The fourth question of this experiment investigated *whether using a tagging paradigm would maintain the quality of mappings generated*. The sixth hypothesis **H6** addressed this question. The sixth hypothesis stated that non-technical users will be able to construct a mapping with high precision and utility. The effectiveness evaluation, see section 6.5.2, showed the mean mapping precision was approximately 75%. This result is comparable with the results from the previous user studies on ontology mapping tools, see section 6.5.5.2. There is still a considerable error rate but investigating the mappings indicated that the majority (73%) of the incorrectly classified mappings were generated by participants who were confused by certain suggested relations, in particular 'part of', see section 6.5.2.1. The user-defined relations were shown to be useful for defining relationships with 92% of these relations accurately defining the relationship of the correspondence, see section 6.5.2.2.

### 6.6.4 Mapping Process Engagement

The final question of this experiment examined *if showing the benefits from the mappings generated by the non-technical users will help in engaging them in the mapping process*. The seventh, eighth and ninth hypotheses (**H7**, **H8** and **H9**) addressed this question. The seventh hypothesis stated that non-technical users will finish the mapping sessions they enter. The results revealed that the number of mapping session answered once entered was quite high with a mean of approximately 60%, see section 6.5.4.2. Indeed it should be noted that there were technical issues with the tag-based mapping tool during the first week which caused the tool to sometimes not display a mapping task during mapping sessions, see section 6.5.4.2 and section 6.9. This could have had an adverse effect on the results but instead in the second week, once the issue had been fixed, the mean increased slightly. The eighth hypothesis stated that non-technical users will see the benefits that mappings bring. The engaging evaluation, see section 6.5.4.2, revealed that users were not identifying the benefits from the mappings (only 25% found the mapping to be tailoring the sport news shown). The ninth hypothesis stated that seeing the benefits, from **H8**, will then give non-technical users an incentive to continue participation in the semantic mapping process. Unfortunately due to the benefits not being recognised by the users there was no incentive available for them to continue participation. However the percentage of mapping sessions entered and finished was still quite high over the course of the experiment and the number of sports stories was also quite high. Furthermore it was indicated that the participants were engaged at the start of the experiment but when incorrect mappings were generated leading to no benefits shown or if mapping tasks not relating to their interests were asked, the participants became less engaged. It should also be observed that the tag-based mapping tool only produces a partial mapping and this could have also potentially affected the user's participation.

## 6.7 Overall Findings with respect to RO2

Although **RO2** is not specifically addressed in this experiment there were several conclusions and issues that emerged which impacted the design and implementation of the tag-based mapping tool. The research objective **RO2** is: Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts. This analysis is based on the results derived from the study.

### 6.7.1 Tag-based mapping tool was unintrusive but disrupted ordinary users when they were busy

It was concluded that the tag-based mapping tool was not hampering the daily work of the user and indeed was blending into the background although when the user was busy they could find it annoying and distracting when a mapping session prompt appeared, see section 6.5.3.3. It is envisaged that this issue would be resolved if the mapping tasks were **context-sensitive** to what the user is doing. It was also mentioned by the participants that they would like the opportunity to request mapping tasks to do at any time as well.

### 6.7.2 Lack of engagement

It was also demonstrated that if users were prompted to undertake mapping tasks that were not related to their interests or if the benefits from the mappings were minimal or unclear that users would become less engaged in the mapping process, see section 6.5.4.3. The participants mentioned the decomposition of the mapping process into small steps helped in engaging them as the time to complete a mapping session was quite small. The conclusion that was indicated is that it is important to keep the mapping tasks related to the users' interests and to keep showing the benefits of the mapping process to keep the user engaged over the course of the full mapping task. To keep showing the benefits to the user it is important to make sure the quality of mappings undertaken by the user is high. In the author's opinion, a solution to resolve the issue of incorrectly generated correspondences is for each correspondence to be **evolved** based on its usage. In this thesis, what is meant by mapping evolution is the changing of the mapping category of a correspondence over time based on its use. In addition, with users unsure of any benefit from mapping tasks, see section 6.5.4.3, it is vital that the benefits be made clear. In the experiment the fidelity of the SportsFlows application took an impact as only 25% of the users mentioned they saw any benefits of doing mapping tasks. This can potentially have an adverse effect on the engagement of users. Finally, in the experiment users were required to answer the same mapping tasks which caused problems as the users were asked tasks not related to their interests. This contributed in users becoming less engaged. It is imperative to keep engagement high, thus ordinary users should be **only asked tasks related to their interests**.

### 6.7.3 Tagging interaction was simple to achieve for ordinary users

The results also suggested that users can achieve rich mappings through the use of tagging although some of the tags may have different meaning to users so the personalisation of a user tag set may improve the quality of mappings, see section 6.5.5.3. It was also shown that the mapping tasks were quick to do and not time consuming for the participants. The participants also said the mapping tasks were straightforward and simple enough although on examining the results there is still a need to make the mapping tasks clearer. Some suggestions on how to make the mapping tasks clearer are to reduce the information to a **more readable form** and to **not to display candidate correspondences which have nothing in common**, e.g. basket and motor racing.

## 6.8 Updated Requirements for Ordinary Users

In this section the mapping requirements for ordinary users, detailed in section 2.8, are updated based on the results of the evaluation. Figure 6-7 displays the updated list of requirements with any modifications made highlighted in italics. Any requirements introduced specifically related to the personal information delivery tool, SportsFlows, will have an asterisk (*) at the beginning of the requirement.

| Eliminate | Reduce |
|---|---|
| Any Matching API Configuration by User<br>Visual Emphasis on Global View<br>**User-Specific Tasks not Related to their Interests* | Complexity of Mapping Process<br>*Tasks per Mapping Session* |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood<br>* *Visualisation of Information Gained Through Mapping* | Iterative Mapping Process<br>Familiar & Rich Mapping Interaction<br>Easy to Understand Interface<br>*Environments for Mapping Sessions*<br>*System Mapping Evolvement*<br>Evaluation Methods |

***Figure 6-7:*** *Updated Mapping Requirements for Ordinary Users*

It was revealed during experimentation that using a time-based approach to prompt mapping sessions led to sessions being prompted in inappropriate browsing contexts for the user, see section 6.7.1. Users found it annoying and distracting when a mapping session prompt appeared when they were busy, see section 6.5.3.3. This requires that mapping sessions only be prompted to users in acceptable browsing environments.

It was shown that users were not highly engaged in the mapping process which may have been due to important correspondences being incorrectly mapped by the user, see section 6.7.2. This requires that the correspondences be monitored and managed by the system to make sure they have been correctly mapped. The engagement rate may also have been impact by the level of work in each mapping session. The percentage of session not started and abandoned was high, see section 6.5.4.2. This suggests the level of work done in mapping sessions needs to be reduced.

The personal information tool, SportsFlows, needs to also be improved. It was shown in the evaluation that users were not seeing the benefits mappings were bringing which may have impacted their engagement in the mapping process, see section 6.7.2. SportsFlows needs to highlight to the user the benefits received from mappings. In addition, users were confused and became less engaged when asked tasks not related to their interests, see section 6.7.2. SportsFlows needs to be implemented to only require the user to answer tasks related to their interests.

## 6.9 Limitations

The number of subjects in this experiment was low at 12 but this number was comparable with other user studies undertaken related to mapping tools, see section 2.7. The low numbers meant that more detailed feedback could be gained and additional analysis from each participant through post experiment interviewing techniques could be achieved. However this impacted the numbers of users per user group and the sample size of the experiment.

There was bias in the selection of participants for the experiment, as the author selected the individuals to take part. This may have led to the participants being more favourable towards the tool being evaluated to please the author. Also it may have led to the participants continual participation during the course of the experiment. There was also potential bias in the selection of the sport domain which may be easier to understand than other domains, e.g. biology domain. However using the sports domain allowed for participants with basic computing skills to participate in the experiment. Additionally the ontologies used were devised by the author, see section 5.2. However this was only done as a search found no ontologies suitable for the scenario chosen. Indeed the author constructed the gold standard which may have also caused bias due to the point of view taken by the author. Finally the prompts were only displayed for 8 seconds to the participants which may have impacted their participation as it was revealed that several participants did not see multiple prompts.

Approximately four times during the first week of the experiment the tag-based mapping tool would not operate properly due to the server caching a huge volume of data which would then cause the server responses to become slow. This would cause a blank screen to appear when the user would click on a mapping prompt. This caused serious issues with and may have accounted for some of the lack of engagement with users. The error was fixed and the tag-based mapping tool operated correctly for the participants in the second week. An initial pilot study may have identified this error as well as any other issues there may have been, e.g. operating system incompatibilities.

## 6.10 Summary

In this chapter the second experiment is presented which evaluates whether it is possible for ordinary users to actively participate in the semantic mapping process within their own workplace over time. The study analyses the mapping interactions generated by three user groups with different levels of technical expertise.

The experiment highlighted that the semantic mapping process can be non-invasive to ordinary users when occurring over time within their computing environment. The results revealed there was no significant difference in the precision and utility of the mapping generated by the user groups. The time taken to complete a mapping task was also comparable across the user groups. However there was a difference in the number of mappings made by ordinary users but this can be attributed to the low number of browsing hours by these users. It was also shown ordinary users found using a ‘tagging’ approach for mapping to be easy and straightforward. However, it was revealed that there was a lack of engagement in the semantic mapping process with users possible due to the benefits of mappings being unclear. Overall in this study the results revealed that non-technical users can generate mapping results to the standard of ontologically aware users.

In the next chapter, several modifications to the tag-based mapping tool and SportsFlows application are detailed. These modifications are made as a result of the findings from the evaluations of the second experiment.

# 7 Modification of Tag-Based Mapping Tool

In this chapter, enhancements are made to the tag-based mapping tool, see section 5.2, based on the results of the user evaluation of the tool, see section 6.5. These improvements were made to allow ordinary users to interact more seamlessly with the tag-based mapping tool over time within their computing environment. First the modifications made to the tag-based mapping tool are described. Afterwards the modification made to the personal information delivery tool SportsFlows is detailed. SportsFlows uses the tag-based mapping tool to personalise sporting information for users.

## 7.1 The Updated Tag-Based Mapping Tool

In this section, the modifications made to the tag-based mapping tool are described. First the updated requirements for the tag-based mapping tool are specified. The modified tag-based mapping tool architecture is then presented. The tag-based mapping tool still has two major components: a browser plug-in and a mapping server. Next an overview of the changes in the architecture of the tag-based mapping tool is indicated. Subsequently the browser plug-in modifications are described, only emphasising the changes made in the implementation of each step in the incremental mapping process. Next the mapping server modifications are detailed, again only emphasising the changes of the implementation. As only the modifications made are described it might be useful to review the basics of the implementation by looking at section 5.1. A walkthrough is not given of how the act of mapping is achieved with the tag-based mapping tool as it is the same as the previous version, see section 5.1.7. The mappings the user develops with the tool are used to personalise sporting information shown to users, see section 5.2.

### 7.1.1 Requirements for Tag-Based Mapping Tool

Several mapping requirements have already been specified for modifying the mapping process and tools to allow ordinary users to engage in act of mapping, see section 2.8. These requirements have been updated throughout the thesis. Figure 7-1 displays the updated list of requirements. The specific requirements for updating the tag-based mapping tool are highlighted in bold.

| **<u>Eliminate</u>** | **<u>Reduce</u>** |
|---|---|
| Any Matching API Configuration by User<br>Visual Emphasis on Global View<br>User-Specific Tasks not Related to their Interests | Complexity of Mapping Process<br>**Tasks per Mapping Session** |
| **<u>Raise</u>** | **<u>Create</u>** |
| Visual Emphasis on Local Neighbourhood<br>Visualisation of Information Gained Through Mapping | Iterative Mapping Process<br>Familiar & Rich Mapping Interaction<br>Easy to Understand Interface<br>**Environments for Mapping Sessions**<br>**System Mapping Evolvement**<br>Evaluation Methods |

***Figure 7-1:*** *Specific Requirements for updating the Tag-Based Mapping Tool*

Firstly, in the previous experiment mapping session prompts were displayed to the user when they were busy doing another task. To combat this issue the tag-based mapping tool was made context-sensitive and only allowed prompts to be displayed in a browsing context agreeable with the user. These browsing contexts are called **mapping browsing contexts**, see section 7.1.4.2. Secondly, in the previous experiment certain correspondences were incorrectly mapped by users with the tag-based mapping tool. This may have been a cause for incorrect sporting information being shown to the user. To combat this problem the mapping category of correspondences will be monitored over time by the tag-based mapping tool and changed based on their use, see section 7.1.5.5. Thirdly, in the previous experiment a high portion of the mapping sessions were exited before they were completed, each session contained three mapping tasks. This may indicate that the level of work in each mapping sessions is too much for ordinary users. After analysing the results it was observed that the majority of participants exited a mapping session after answering one mapping task. To combat this issue the number of mapping tasks in each mapping session is reduced to only one task. Fourthly, in the previous experiment certain participants had trouble understanding the ontological information being displayed. To combat this problem the natural language used to represent the ontological information was refined to make it more readable, see section 7.1.4.3. Finally, in experiment two several of the suggested relations were misunderstood by the participants and lead to incorrectly mapped correspondences. To combat this issue, only the suggested relations which were fully understood by the participants will be used, see section 7.1.5.3.

## 7.1.2 Updated Tag-Based Mapping Tool Design

Figure 7-2 gives a high level architecture overview of the tag-based mapping tool and identifies where the steps of the incremental mapping process are integrated. As before, the tag-based mapping tool is split into two distinct parts: a browser plug-in and a mapping server. The step 'determine what mapping task to present' was modified to occur on the browser plug-in due to tasks being context sensitive, see section 7.1.4.2.

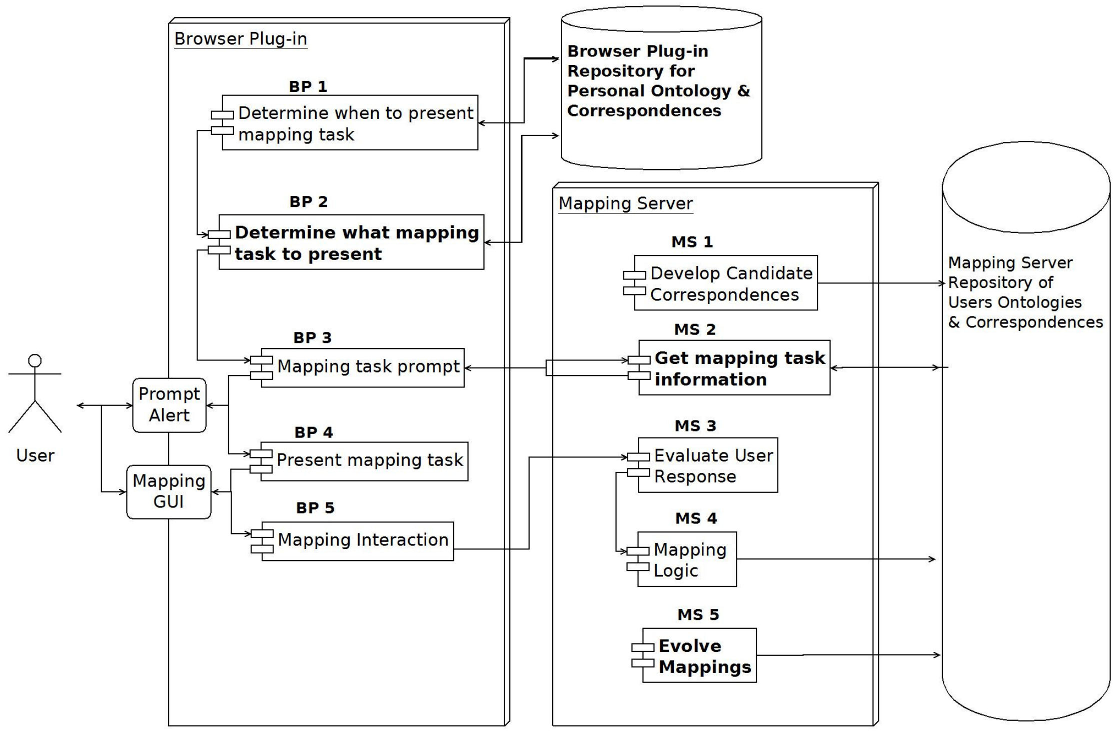



***Figure 7-2:*** *High Level Architecture Overview of Tag-Based Mapping Tool*

## 7.1.3 Updated Implementation Overview

The tag-based mapping tool uses a client-server architecture. A Firefox browser extension was the implementation of the client. The browser extension is compatible with all Firefox browsers up to version 3.3 and implements the user interface of tag-based mapping tool. In a modification to the implementation a SQLite database [Owens 2006] is used to store information in the user browser. The information stored includes correspondences and a log of users mapping actions.

The mapping server is a web application which again runs on Tomcat web server version 6. The mapping server again uses SPRING framework to allow for a plug-in architecture. In a modification made to the implementation, iBatis [Begin 2007] is used as the database persistent layer instead of Hibernate. This changed was made as in the previous evaluation there was an issue with caching with Hibernate which led to the mapping server running slower, see section 6.9. Another modification made was the implementation of Quartz threads [Cavaness 2006] on the mapping server which allowed for candidate correspondences to be generated daily, see section 7.1.5.1, as well as the mapping category of correspondences to be monitored daily, see section 7.1.5.5. A final change made was the addition of an instance based matcher. The instance based matcher used a Jaccard measure [Isaac 2007] to generated candidate correspondences. This measure compares the similarity and diversity of concepts by analysing the overlapping of instances between concepts.

### 7.1.4 Firefox Browser Extension Implementation

A Firefox browser extension is used to implement the user-interface of the tag-based mapping tool, see Figure 7-3. The user mapping interface is displayed transparently over the web page the user is viewing when the user enters a mapping session. There is an icon, which looks like a question mark in a circle that represents the tag-based mapping tool and is displayed in the bottom right-hand corner of the browser. The icon changes colour to indicate if something has occurred its default colour is blue. The sequence of action with the tag-based mapping tool is the same as the previous version of the tool, see section 5.1.4. During the implementation two volunteers were used to test and evaluate the extension and thus the approach taken. These volunteers also helped in devising the maximum numbers of mapping prompts displayed per day via trial and error.

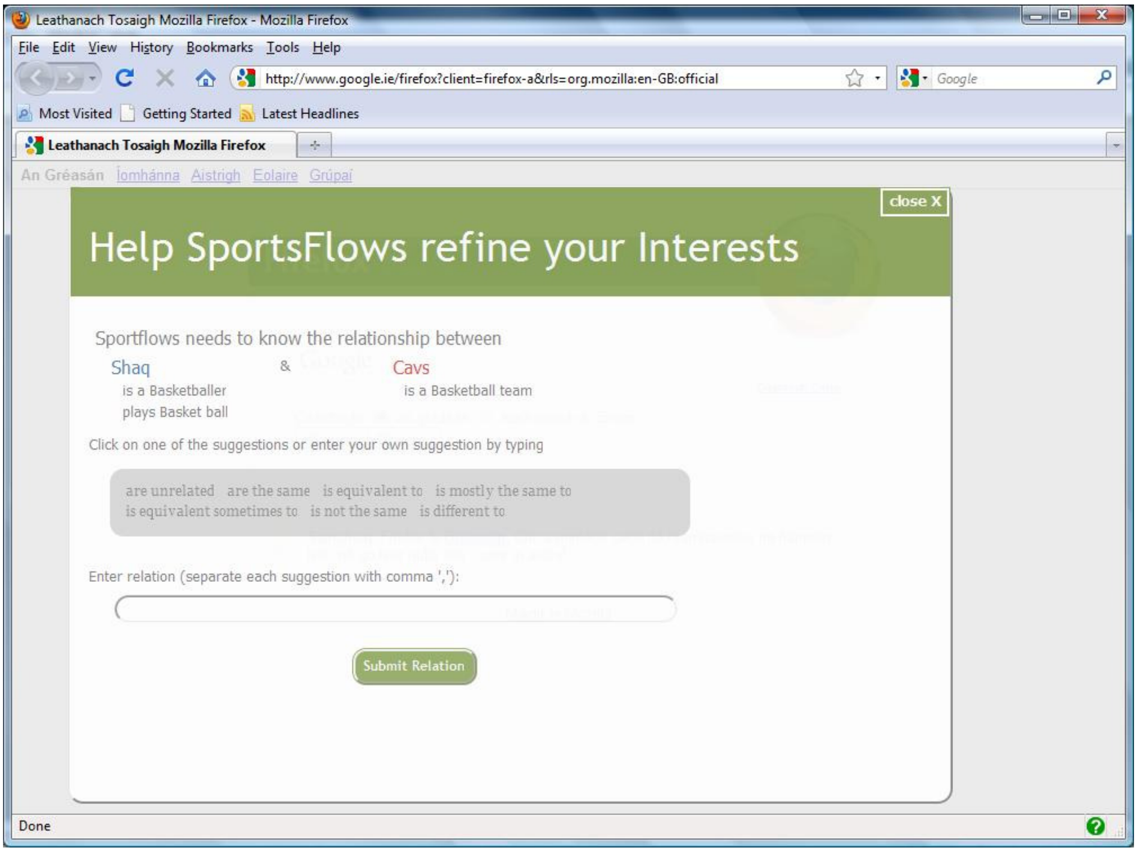



*Figure 7-3: Screenshot of the Mapping Interface of the Tag-Based Mapping Tool*

#### 7.1.4.1 Determine when to present mapping task (BP1 in figure 7-2)

Mapping tasks are only prompted to the users in a **mapping browsing context**. These mapping browsing contexts are only on web sites in which the user is willing to answer a mapping task and are not web sites in which they are busy and do not want to be disturbed, for example emailing. The only valid mapping browsing contexts available at first are the sporting RSS news websites used by SportsFlows, see section 7.2. However additional mapping browsing contexts are added over time. The additional mapping browsing contexts added are any web pages where the user clicks on a latest sporting news story message alert, see section 7.2.

The priority of the mapping browsing contexts will change over time based on the users' mapping activity within them. Table 7-1 displays the full list of priorities a mapping browsing context can have. The initial priority for a mapping browsing context is 3. The priority list allows for different level of mapping tasks to be asked but it is not implemented in this version of the tag-based mapping tool, e.g. only urgent tasks allowed for level 1 and level 2 while level 4 and level 5 can allow for less urgent tasks to be asked. The mapping browsing context also has a current index for the priority level. The current index ranges from 0 to 4 and initially starts at 2. If the user clicks on a prompt for the mapping browsing context the current index increases by one, if they don't click on the prompt the current index decreases by one. When the current index is 4 the priority level is increased by one, alternatively if it is 0 the priority level decreases by one. The current index will then resets to 2. If the mapping browsing priority level decreases to 0 the browsing context will be black-listed, i.e. no mapping task prompts will be asked in this mapping browsing context.

**Table 7-1:** Priorities of mapping browsing context

| **Level** | **0** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|---|
| **State** | Black-listed: No mapping prompts are to be shown | Valid mapping browsing context: Mapping prompts can be shown | | | | |

When a user navigates in their browser to a mapping browsing context, a mapping task prompt can be presented to the user. The mapping task prompt will only be presented to the user if they have not already been prompted with 3 mapping tasks that day or have not been prompted within the last hour. The number of mapping tasks prompted to the user is recorded in the SQLite database in the Firefox browser.

#### 7.1.4.2 Determine what mapping task to present (BP2 in figure 7-2)

The only mapping task asked is to categorise the relation of a candidate correspondence generated by the mapping server, i.e. map candidate correspondences. From the current list of candidate correspondences available to the user the decision on which mapping task to present is based on the following conditions in the following order: (1) context of the user and (2) the confidence of the correspondence. This modification was made due to instance information not being available through the correspondences as in the previous implementation, see section 7.1.5.1. The context of the user refers to their browsing context which could be specific to certain interests or sports. For example if the user is browsing a web page about 'Football' then only candidate correspondences which had 'Football' or a concept related to 'Football' would be asked. The confidence refers to the confidence measure given to the candidate correspondence by the matcher, see section 2.1. Figure 7-4 displays the pseudo-code for the process of determining which candidate correspondence.

```
Retrieve List of Candidate Correspondences
Initialise Response as null
Initialise Title as title of the webpage
Initialise Context as false
While List not Empty
      Get Candidate Correspondence from List
      If Context false
            If Concept in Candidate Correspondence contained in Title
                  Context is set to true
                  Response is set to Candidate Correspondence
                  Continue
            Else If Response is null
                  Response is set to Candidate Correspondence
                  Continue
            Else If Response Confidence < Candidate Correspondence
                  Response is set to Candidate Correspondence
                  Continue
      Else If Concept in Candidate Correspondence contained in Title
            If Response is null
                  Response is set to Candidate Correspondence
                  Continue
            Else If Response Confidence < Candidate Correspondence
                  Response is set to Candidate Correspondence
                  Continue

If Response null
      Don't Display Mapping Task Prompt
Else
      Display Mapping Task Prompt for Response
```

***Figure 7-4:*** *Pseudo-code for Selection of Candidate Correspondence*

The list of the current candidate correspondences available to the user is stored in the SQLite database in the Firefox browser. If no candidate correspondence is available then no mapping task prompt will be displayed to the user. The list of candidate correspondences is updated by the mapping server every time the user opens their Firefox browser and the extension logs onto the mapping server, see section 7.1.5.

#### 7.1.4.3 Mapping task prompt (BP3 in figure 7-2)

The mapping task is prompted to the user with a message alert, see Figure 7-5. The message in the task prompt is changed from the previous tool to ask for the user's help, i.e. 'Sport Interest Refinement Needed'. This change was made to better suit the message being asked that the application needs help. In the previous iteration of the mapping tool, the prompt would just tell the user that mapping tasks need to be answered, see section 5.1.4.3. The prompt asks the user for assistance in finding the relation between their interest and another user's interest, see section 7.1.5.1. This allows the user to view what the mapping task will be about before clicking on the prompt. Also the message in the prompt is asking the user for help. This will give the impression to the user that they are assisting the tool rather than being tested. Both the title and sentence of the prompt are placeholders which can be configured by applications using the tag-based mapping tool.

The user enters the mapping task by clicking the alert which will present the mapping task to the user, see next section. However the user can also choose to just ignore the message alert. The SportsFlows icon will then change colour to green to let the user know there is a mapping session available, see Figure 7-3. This allows the user to enter the session at a later time. After twenty minutes if the user has not entered the mapping task the task becomes unavailable and the icon changes colour back to blue.


***Figure 7-5:*** *Tag-Based Mapping Tool Mapping Task Prompt*

#### 7.1.4.4 Present mapping task (BP4 in figure 7-2)

As in the previous implementation, the mapping task is displayed transparently in a XUL panel over the web page the user is viewing. The mapping information is displayed to the user in natural language but the phrasing of the template is changed in this tool from the previous iteration to allow the information to be more clearly readable by users. Each ontological term is displayed by bullet point with each bullet point representing its structure, i.e. child, parent, properties. In the previous experiment some participants had trouble understanding the ontological information being displayed due to the volume on display, see section 6.5.5.3. To reduce the volume the parent concept is not chained back to the top level concept. Also the property template just displays the property next to the restricted concept which makes the information of the property more readable than with the previous tool.

In addition the phrasing of the question is changed, which emphasises that the tool needs help in evaluating what the relationship is and needs the user's help. Figure 7-6 displays the interface.

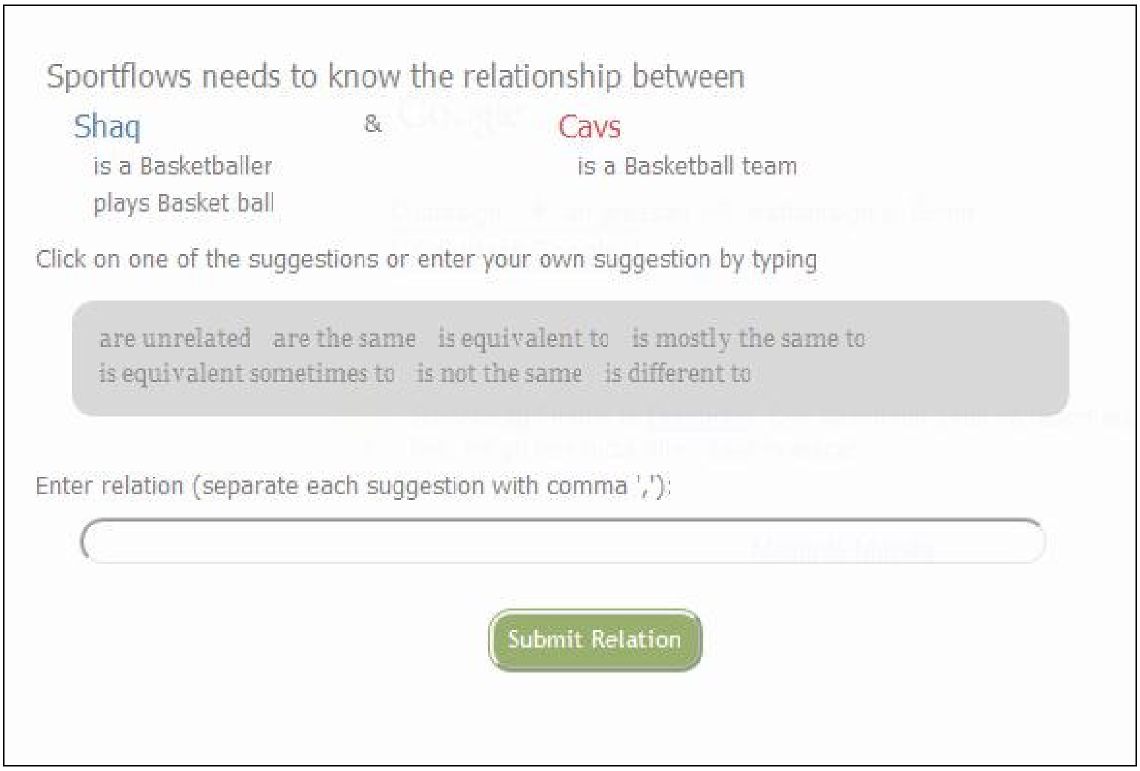



***Figure 7-6:*** *The Tag-Based Mapping Tool user mapping interface*

#### 7.1.4.4 Mapping task interaction (BP5 in figure 7-2)

As with the previous tag-based mapping tool, the mapping task requires the user to 'tag' the relation for a candidate correspondence which will be used to categorise the mapping category of the correspondence. As before the user can enter user-defined relations or tags from the suggested relation list, see Figure 7-6. However, in contrast to the previous iteration, the phrasing asks the user for the relation rather than to tag the relationship. The suggested relations are presented as portions of a sentence, see Figure 7-6. The suggestions are also limited to the relations which the majority of participants used in the previous evaluation, see section 6.4.5.2.

### 7.1.5 Mapping Server

The mapping server is again used to store information and to execute several system-specific mapping tasks, see section 5.1.5. This version of the mapping tool allows candidate correspondences to be generated at runtime. In the following sections, the modifications made to the mapping server are detailed.

#### 7.1.5.1 Develop candidate correspondences (MS1 in figure 7-2)

In contrast to the previous prototype, the candidate correspondences are developed at **runtime** using a hybrid approach of both schema and instance based matching. This approach was taken as it allows new candidate correspondences to be generated for the user to map over time. The INRIA Alignment API [Euzenat 2006] is the schema based matching tool used, the `NameAndPropertyAlignment` alignment method was selected again, while the instance based matcher was developed using the Jaccard measure [Isaac 2007]. The mapping server uses a plug-in architecture which makes the substitution of other matchers and alignment generators possible. The instances are the web addresses of the sport news stories that the user looks at, see section 7.2. The list of candidate correspondences, for each user, is kept up to date with a quartz thread running once every day to develop any new candidate correspondences if possible.

The candidate correspondences generated are between the users' personal ontologies. These candidate correspondences will help to enrich the user's ontology by annotating their interests with other terms taken from the ontologies of other users of the SportsFlow application, see section 7.2. For example if a user had an interest 'Lionel Messi' it might be enriched with added terms from correspondences such as 'Barcelona Players' and 'Argentine Players', which would be the interests of other users. As in the previous implementation the *Equivalent* and *Equivalent Sometimes* categories are used to categorise correspondences which serve as valid correspondences for interchange of information, see section 5.1.5.3.

#### 7.1.5.2 Get matching information (MS2 in figure 7-2)

All of the ontologies are stored on the mapping server. When presenting the mapping information for mapping tasks the local semantic neighbourhood for each concept in the candidate correspondence needs to be obtained.

#### 7.1.5.3 Evaluate user response (MS3 in figure 7-2)

The mapping category of the tags entered by the user is evaluated in this step. The implementation of this step is substantially the same as the previous version of the tag-based mapping tool, see section 5.1.5.3. The modification made is that several suggested relations which were unclear to the user in the previous experiment (or not used) were removed. This was to reduce the number of suggestions in the suggested relation list to make the list clearer to users. Unfortunately an oversight was the removal of the Subclass and Super class relations. This may have contributed to several correspondences being mapped incorrectly in the next experiment, see section 8.5.2.1. However the number was just 3% of the incorrectly constructed correspondences. Figure 7-7 displays the modified taxonomy of categories.

- **Equivalent**
    - *The same*
        - Are the same
        - Is equivalent to
    - *Super class*
    - *Intersection*
- **Equivalent Sometimes**
    - *General*
        - Is mostly the same to
        - Is equivalent sometimes to
    - *Subclass*
    - *One of*
    - *Union*
- **Different**
    - *Different from*
        - Is different to
        - Is not the same
        - Are unrelated
    - *Complement of*
    - *Disjoint*
- **Unknown**

***Figure** 7-7: The modified taxonomy of categories.*

#### 7.1.5.3 Mapping Logic (MS4 in figure 7-2)

The mapping logic step evaluates the mapping category of correspondences from the tags that the users enter and is the same implementation as the previous version of the tag-mapping tool, see section 5.1.5.4. Figure 7-8 shows the mapping category cycle of correspondences. The top half of the diagram represents the implementation of this step while the bottom half is how the mapping category of correspondences evolves which is discussed in the next section.

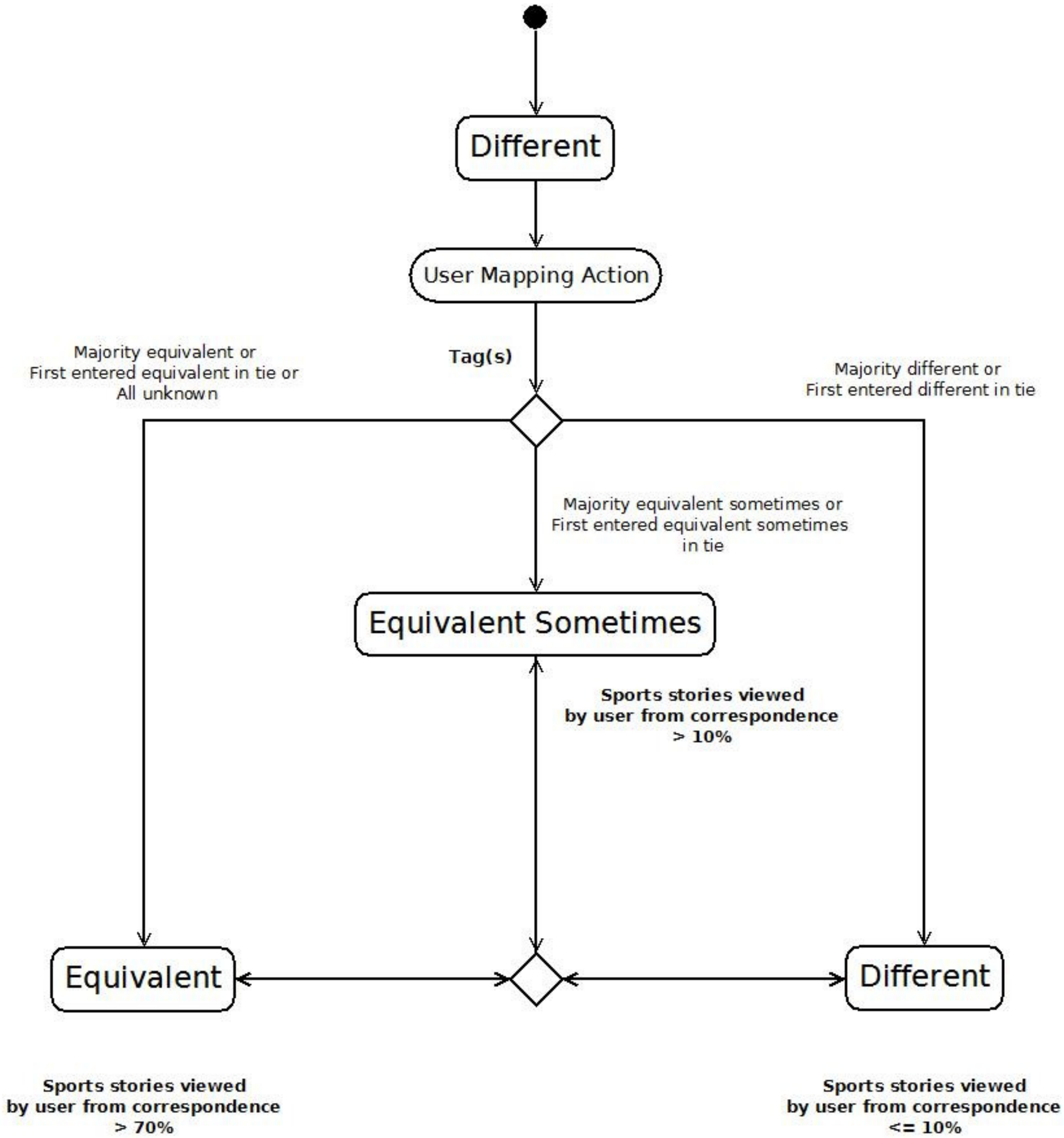



***Figure 7-8:*** *Mapping Category Cycle of Correspondences*

#### 7.1.5.4 Mapping Evolution (MS5 in figure 7-2)

The usage of correspondences generated by the user is monitored. This involves calculating the percent of use of the correspondences, i.e. the percentage of sporting news stories from correspondences which are viewed by the user. This percentage is then used to check the legitimacy of the mapping category for the correspondence, see Figure 7-8. If a correspondence has a percentage which does not satisfy their mapping category then their mapping category is changed, see Figure 7-8. For example if a correspondence was categorised as equivalent but the user was generally ignoring all of the suggested sport news stories generated from the correspondence, e.g. the percentage would be less than 10%, then the correspondence would be re-classified as different. The percentages were calculated from trial and error when testing the system and can be modified for other implementations. The approach taken is measure in the evaluation of the tag-based mapping tool.

## 7.2 Updated SportsFlows

This section gives the details of the modifications made to the personal information delivery tool SportsFlows based on the results of the previous experiment. Firstly the requirements evaluated from the previous experiment are given. Then the updated design of SportsFlows is presented. Afterwards the modifications made to the personal ontology generation are detailed. Finally the modifications are given on how SportsFlows personalises the sporting news for users.

### 7.2.1 Updating Requirements for SportsFlows

Figure 7-9 displays the list of updating requirements for SportsFlows. The specific requirements are highlighted in bold.

| **Eliminate** | **Reduce** |
| --- | --- |
| Any Matching API Configuration by User | Complexity of Mapping Process |
| Visual Emphasis on Global View | Tasks per Mapping Session |
| **User-Specific Tasks not Related to their Interests** | |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | Iterative Mapping Process |
| **Visualisation of Information Gained Through Mapping** | Familiar & Rich Interaction |
| | Easy to Understand Interface |
| | Browsing Environments for Mapping Sessions |
| | System Mapping Evolvement |
| | Evaluation Methods |

***Figure 7-9:*** *Specific Requirements for updating SportsFlows*

Firstly, the previous evaluation showed participants were not seeing the benefits of mappings which led to a lack of incentive for the users to answer mapping tasks. To combat this concern the benefits of mappings will be shown to the user by highlighting which information offered to the user has been gained as a result of mapping tasks, see section 7.2.4. Secondly, the majority of the participants in the evaluation responded that they found being asked about mapping tasks which were not related to their interests unnecessary and confusing and thus they would exit the task mapping session. This problem was resolved by only prompting with mapping tasks related to the user's own interest. To keep the mapping tasks related to the user's interests, a personal ontology was developed independently for each user describing their own interests rather than using the same personal ontology for each user, see section 7.2.3.

### 7.2.2 Updated SportsFlows Design

There are two responsibilities that the SportsFlows application has: (1) The generation of the personal ontology, and (2) The personalisation of sporting news from RSS feeds. Figure 7-10 displays the updated design of SportsFlows. The updated version has the user develop a personal ontology to suit their needs. In the previous version the users had the same ontology represent their personal ontology, see section 5.2. The ontology generation process required users to enter and categorise their sporting interests which are then used to generate the personal ontology for the user. Once the user's personal ontology is generated, the user can then update their ontology by either entering or removing sporting interests. In addition, the RSS reader is implemented on the client. Unlike the previous version of SportsFlows, there is no central sport ontology to represent sport news feeds. Instead mappings will be generated between each of the user's personal ontologies. Thus, sporting knowledge will be shared between the users. This scenario requires multiple users to use the tool for the benefits of mappings to be shown effectively.

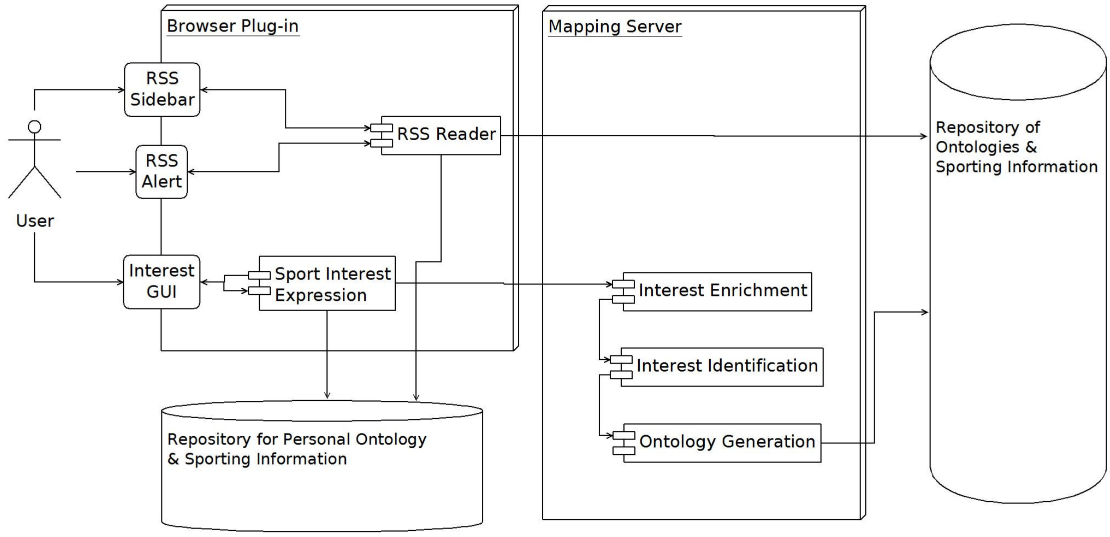



***Figure 7-10:*** *Overall architecture of SportsFlows*

## 7.2.3 Personal Ontology Generation Implementation

This section details the steps of how a personal ontology is generated for the users. The steps follow the process presented in [Angeletou 2008].

### 7.2.3.1 Sport Interest Expression

This step details how the user's sport interests are entered into the system. The sporting interests will form the concepts of the user's personal ontology. There are two ways user's interests are entered into SportsFlows: (1) through user input and (2) through user monitoring. Users can input their own sporting interests with a graphical interface in the Firefox browser extension of SportsFlows. The users first enter which sports they are interested in, see Figure 7-11, and then adds any interest they have to the sport, see Figure 7-12. For example a user may select the sport 'Football' and add the interests 'Pele' and 'Brazilian Players'. The users can also remove any sporting interests they have by clicking on the 'x' icon. If the user does not append an interest to the sport then it is considered the user has a broad interest in the sport and the sport itself is chosen as the sporting interest.

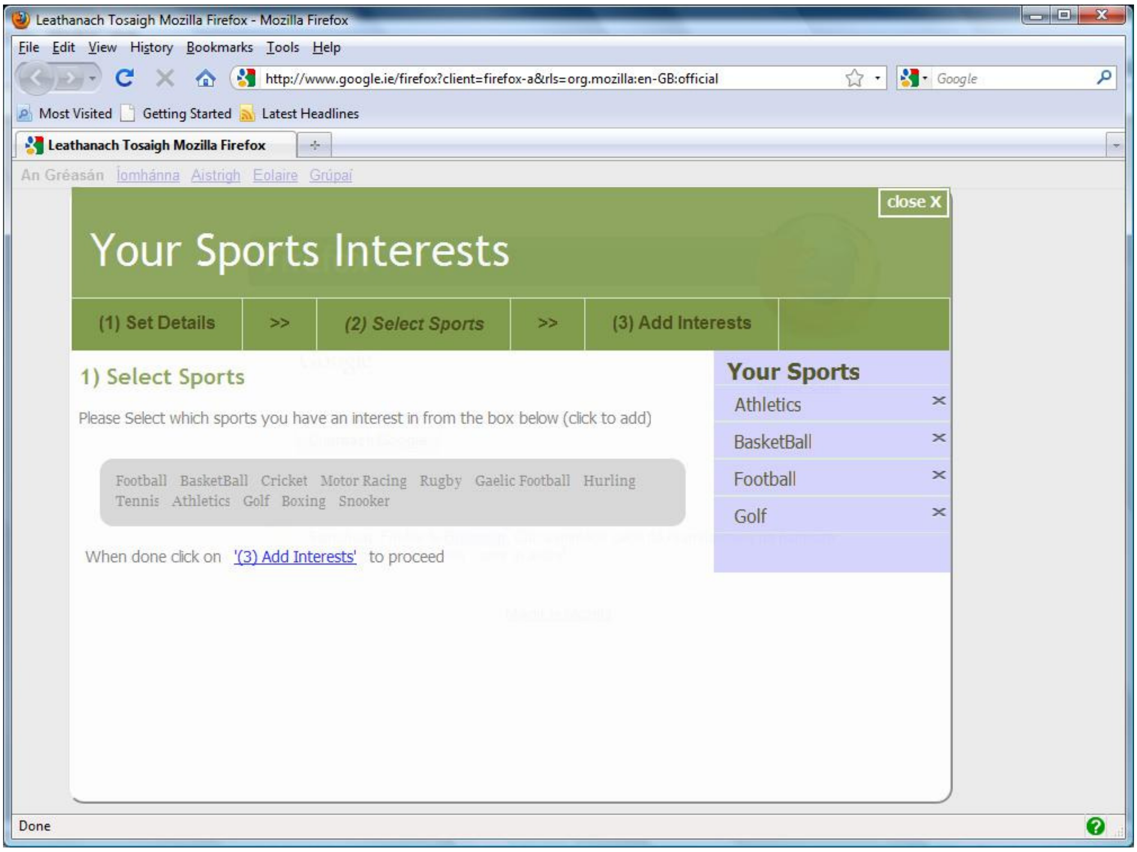



***Figure 7-11:*** *Add Sport as an interest in SportsFlows*

In addition, sporting interests can also be added by monitoring what sporting stories, for each individual sport, the user is looking at. How this works is as follows,

- Firstly, the words in the titles of the stories the user views are extracted;
- Secondly, words which are stop-listed are removed from this list. For example the words ‘from’, ‘at’ and ‘on’ are stop-listed. The resultant list now contains keywords;
- Thirdly, the keywords are entered into a table on the mapping server. If the keyword already exists its count is increased for the user;
- Fourthly, an administrator verifies if the keyword is valid as an interest or if it should be blacklisted. In the experiment carried out in the next chapter the author verified the keywords at the end of each day of experimentation;
- Fifthly, if the count of a valid keyword is 5 or more then it is append as a sporting interest of the user.

This method allows a way for users who do not use the graphical interface to express their interests without user interaction. However, this would be a slow start method and the users would not see the personalisation and benefits until interests have been added. Instead the method is used to increase the knowledge of the user’s interests by appending any missed interests. Once sporting interests have been added or removed, the personal ontology for the user is generated (or updated) with the following steps.

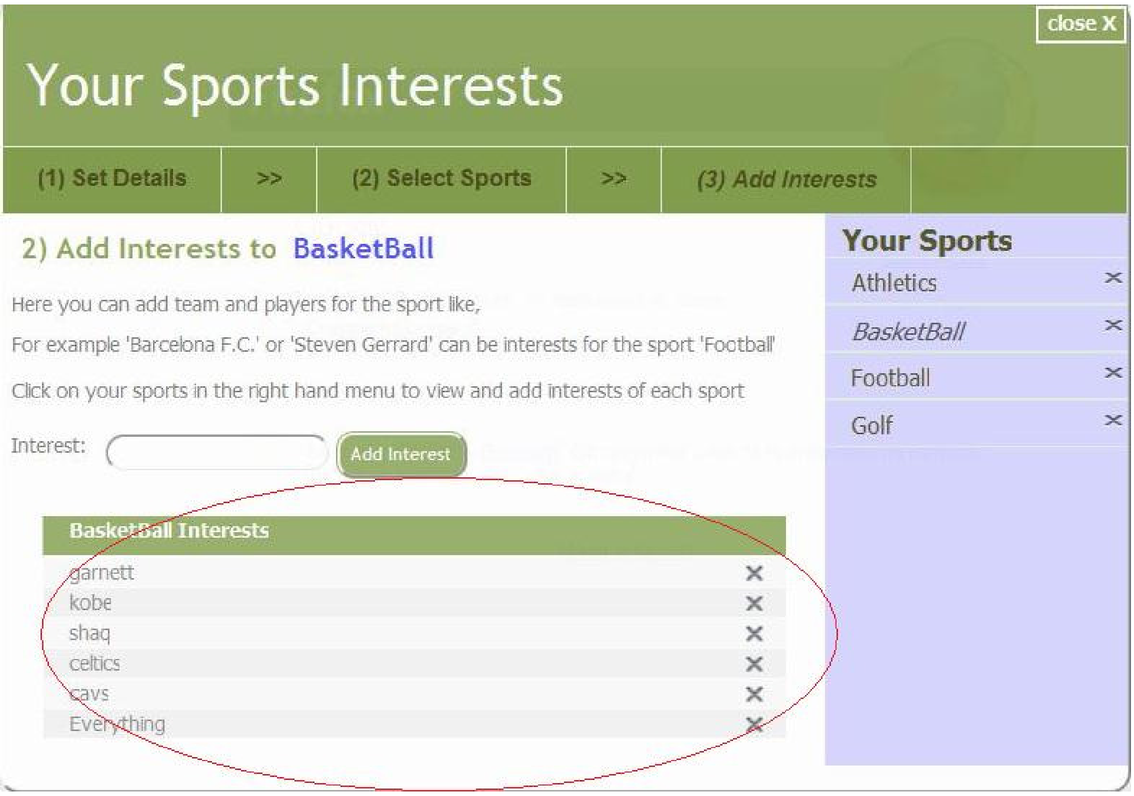



***Figure 7-12:*** *Add interests to sports in SportsFlows. Red circle highlights the list of interests for the sport.*

#### 7.2.3.2 Interest Enrichment

After the user finishes entering or removing sport interests, the interests are sent to the mapping server where they are enriched. Firstly the name for the sporting interest appended with the sport, for example 'Pele Football', is used in an http search query on Wikipedia[32] search box by the mapping server. The resulting http response is screen scraped to find the corresponding Wikipedia article for the sporting interest, i.e. the link to the first result in the result list. This link is then filtered to find the corresponding Wikipedia name for the sport interest, e.g. the interest 'messi' for football would have Wikipedia name 'Lionel_Messi'. This name is then used in a remote SPARQL query on DBpedia [Auer 2007] to enrich the interest with *added terms*, see Figure 7-13. The added terms are the 'type' and 'subject' of the corresponding DBpedia article for the interest. These added terms are used in the next step to help rationalise the parent concept of the interest.

[32] http://en.wikipedia.org/wiki

```
SELECT ?type WHERE {<http://dbpedia.org/resource/"+ interest +">
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> ?type}

SELECT ?subject WHERE {<http://dbpedia.org/resource/"+ interest
+"> <http://www.w3.org/2004/02/skos/core#subject> ?subject}
```

***Figure 7-13:*** *The remote DBpedia SPARQL query*

#### 7.2.3.3 Interest Identification

In this step, the parent concept of the interest and the relationship between the interest and the sport are calculated, e.g. the relationship between 'Lionel Messi' and 'Football'. The author developed a sport domain ontology which is located on the mapping server and encapsulates the structure of the sports domain. The sport domain ontology is provided in the accompanying DVD media under Experimental Datasets - Experiment 3. All of the sports suggested to the user by SportsFlows, see Figure 7-11, have a corresponding concept in the sport domain ontology. Each concept in the sports domain ontology is annotated with rdf-label properties. These labels are the added terms found in DBpedia for the concept, see section 7.2.3.2. A corresponding concept is found for the interest by comparing the added terms found for the interest against the rdf-label properties for the concepts. The concept that best matches the interest, i.e. has the most overlapping terms to labels, is then chosen to be the interests parent concept. The relationship between the interest and sport is the relationship between the parent concept and sport concept in the sport domain ontology. If a concept in the sports domain ontology cannot be found for the interest then the parent concept and relationship will be undefined. An administrator is then tasked with entering both the parent concept and relationship with the sport for the interest. In the experiment carried out in the next chapter the author did this task at the end of each day of experimentation.

#### 7.2.3.4 Ontology Generation

The final step is to generate the personal ontology from the information found for the sporting interests in the previous step. Formal concept analysis (FCA) is the method chosen and is a way of automatically deriving an ontology from a collection of objects and their properties. The technique used was first presented in [Quan 2004] and requires a set of triples (O1, R, O2) where R is the relationship between O1 and O2. Using this technique both O1 and O2 will be generated into concepts. The set of triples used are the following:

1. *(S,R,I)*: where *S* is the sport, *I* is the interest, and *R* is the relationship between the sport and the interest;
2. *(I,R',S)*: where *S* is the sport, *I* is the interest, and *R'* is the relationship between the interest and the sport;

3. *(S,R',P)*: where *S* is the sport, *P* is the parent concept of the interest, and *R'* is the relationship between the sport and the parent concept;
4. *(P,R,S)*: where *S* is the sport, *P* is the parent concept of the interest, and *R* is the relationship between the parent concept and the sport;
5. *(I,sub,P)*: where *P* is the parent concept of the interest, *I* is the interest, and *sub* is the subclass relationship;
6. *(P,super,I)*: where *P* is the parent concept of the interest, *I* is the interest, and *super* is the superclass relationship;
7. *(P,Rel,C)*: where *P* is the parent concept of the interest, *C* is a concept in the sport domain ontology, and *Rel* is the relationship between *P* and *C*;
8. *(C,Rel',P)*: where *P* is the parent concept of the interest, *C* is a concept in the sport domain ontology, and *Rel'* is the relationship between *C* and *P*;
9. *(I,Rel,C)*: where *I* is the interest, *C* is a concept in the sport domain ontology, and *Rel* is the relationship between *I* and *C* [Corresponding to number 7];
10. *(C,Rel',I)*: where *I* is the interest, *C* is a concept in the sport domain ontology, and *Rel'* is the relationship between *C* and *I* [Corresponding to number 8].

#### 7.2.3.4 Limitations

The approach taken only allows for the generation of simple ontologies. The ontologies generated will have concepts that have very simple relationship between one another. This may lead to ambiguity around some of the concepts generated. The approach also relies on the sport domain ontology to encapsulate the entire sport domain.

SCARLET [Sabou 2008] is an application which discovers the relationship between two concepts by making use of online available ontologies. SCARLET could potentially be used to enrich concepts by finding property relationships between different concepts. Unfortunately, at present SCARLET is only implemented to find simple concept relationships between concepts, e.g. 'is a' and 'subclass', and is not implemented to return the property relationships between concepts.

### 7.2.4 Personalisation of Sporting News

The `RSS Reader` in the Firefox extension is responsible for personalising the sport news and displaying the sport news to the user. The RSS feeds used and the timing of when the feeds are searched are the same as the previous version, see section 5.2.2. The RSS sport news stories are stored in the database on the browser.

#### 7.2.4.1 Personalisation of Sporting News

The sports news is personalised using a similar method to the previous version of SportsFlows, see section 5.2.2. Firstly, the user's sporting interests are enriched with additional terms from the correspondences generated by users. The correspondences which have been mapped as *equivalent* or *equivalent sometimes* will provide these additional terms from the names of the target concepts. For example, if a user had an interest in 'Lionel Messi' and they mapped the correspondence 'Lionel Messi' to 'Barcelona Footballer' with an *equivalent* relation, then 'Barcelona Footballer' would be an additional term for 'Lionel Messi'. However 'Lionel Messi' would not be an additional term for 'Barcelona Footballer' unless the converse correspondence 'Barecelona Footballer' to 'Lionel Messi' is mapped with an *equivalent* or *equivalent sometimes* relation. The rationale behind this decision is that every correspondence generated by the user is from their perspective and cannot be shared with other users as they may have differing views.

Afterwards, the sports news stories from RSS feeds are categorised as following:

1. The title of sport news story RSS item is search for the additional terms of the sporting interests (the terms include the sporting interest name);
2. The sport news story is categorise to the sporting interest which had the most additional terms in the title;
3. If there was a tie between interests then the story would be categorised to the interest which had the first term in the title, i.e. the term which is closest to the start of the title;
4. If no interest had any terms in the title then the sport news story would be categorised as general.

As more correspondences are developed by the user more additional terms will be appended to their interests, which will then categorise additional sport news stories to their interests. Therefore in contrast to the previous tool, the act of mapping will be a branching out method that enriches the user sporting interests with additional terms. This should allow the benefits from mappings to be more visible to the users, as the sport news stories for interests will increase as users complete mapping tasks. In addition, unlike the previous version of SportsFlows the users should not see stories that are not relevant to their interests due to this branching out method.

#### 7.2.4.2 Sport stories prompts

Latest sports news stories for the users' sporting interests are prompted to the user. The latest sport news stories are stories which have occurred within the last hour and have not already been prompted to the user. The stories are displayed to the user in a message prompt in the right hand bottom corner of the screen, see Figure 7-14. If the sports news story was added using a correspondence from a mapping then a '*' would be appended to the start of the story title. This was done to make the benefits of mappings more easily identifiable to the users and to help encourage the user to use the tag-based mapping tool.


***Figure 7-14:*** *Latest sport news story prompt, displayed in the bottom right-hand corner of the Firefox browser*

The prompt is displayed for 8 seconds before closing, giving the user enough time to see the story without it becoming too intrusive. This is the same time used in the previous tag mapping tool, see section 5.4.4. If the user clicks on the prompt the sport news story will be opened in a new tab on the Firefox browser. Also a sidebar displaying all the sports news stories will be opened, see section 7.2.4.3. If the user chooses instead not to click on the prompt, the SportsFlows icon in the bottom right hand corner of the Firefox browser will change colour to orange to indicate to the user that new sports news stories are available. If the user double clicks on the icon the sport new story sidebar will open displaying the new sport news stories and the icon reverts back to its original colour. After twenty minutes if the user has not opened the sport news story sidebar the icon will revert back to its original colour.

#### 7.2.4.3 Sports Stories Sidebar

SportsFlows has a sport story sidebar which is used to display all the sports news stories that the application has stored in the browser database. Figure 7-15 displays the sport story sidebar.

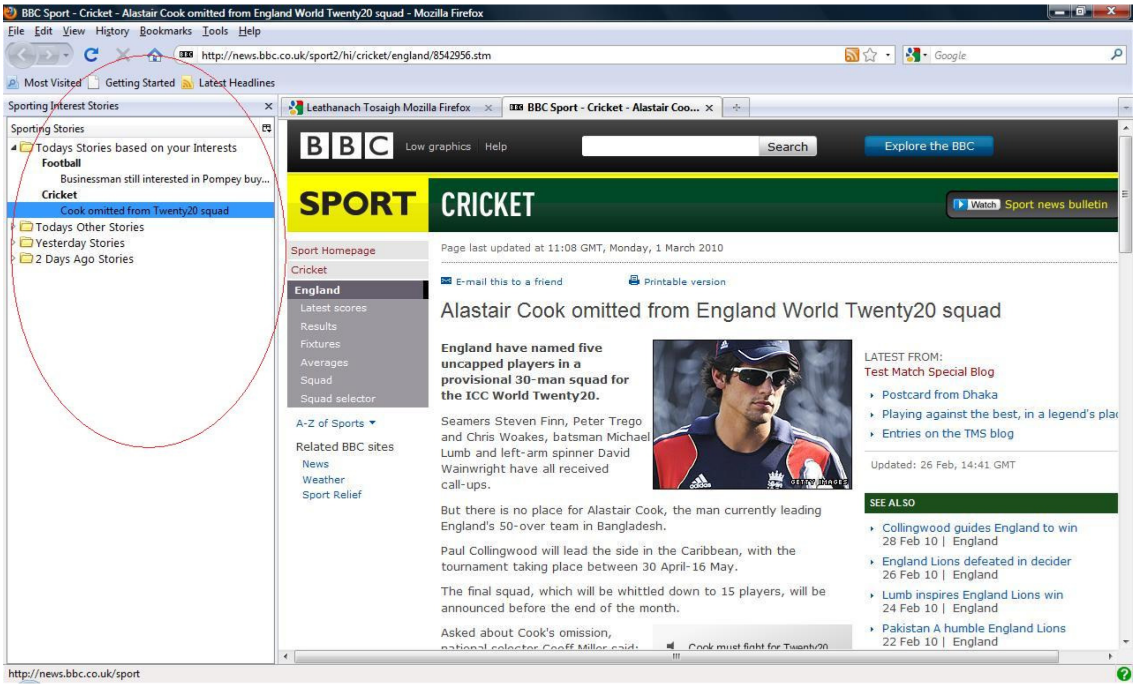



***Figure 7-15:*** *The sport news story sidebar, highlighted with a red circle*

The sidebar separates the sporting news into four different categories: *Today's stories based on your interests* the sport news for today that are based on the user's interests, *Today's other stories* the sports news for today that are not based on the user's interests, *Yesterday* for the sport news from yesterday, and *2 days ago* for the sports news from two days ago. The modification made to sports stories sidebar from the previous implementation is how the sport stories are arranged in these categories. In each category the sports news stories are segmented into stories for the user's interest, see figure 7-15. Sport news stories that are not based on the user's interests are displayed in the sidebar under the title "other stories" however they are not prompted to the user, see Figure 7-14. These sport stories are available in the sidebar so the user can view them and develop new interests by opening them, see section 7.2.3.1. By double clicking on any sport news story the story will open on a new tab in the Firefox browser. Each category can be collapsed or opened. When the sidebar is opened only the latest and today category are opened.

## 7.3 Summary

This chapter describes the updates made to the tag-based mapping tool and the personal information delivery tool SportsFlows. In the next chapter, the modified tag-based mapping tool is evaluated with an open field trial of ordinary users using the tool over time within their own work environment.

# 8 Updated Tag-Based Mapping Tool Evaluation

In this chapter the last of the three experiments is presented. The goal of experiment three was to address research objectives **RO1**: Determine whether ordinary users can actively participate in the construction of semantic mappings to a similar standard as knowledge engineers, and **RO5**: Evaluate the mapping tool with ordinary users over a long term period. The study evaluates whether ordinary users can actively participate in the semantic mapping process within their own workplace over a long time period by analysing the mapping results generated by three different groups of users each with different technology expertise. In contrast to experiment two: the call for participation was open allowing for anybody within the college to participate, the purpose of the experiment was concealed from the participants and the experiment was carried out over a five week period. The rationale behind concealing the purpose of the experiment was to allow real world data with ordinary users using the tag-based mapping tool to be recorded without them being unfairly influenced. Instead the participants were told the SportsFlows application was being evaluated. The participants were required to use the tool for at least two weeks out of the five weeks. In particular the experiment **solicits ordinary users to participate in the mapping process** with the tag-based mapping tool within their browsing environment **for at least two weeks and up to five weeks**.

The structure of this chapter is as follows: first requirements for the evaluation of the tag-based mapping tool are specified, secondly the design of the study is discussed, thirdly the hypotheses of the experiment are stated, fourthly the methods for the experiment are explained, fifthly the results of the experiment are presented, sixthly the analysis of the experiment is detailed, seventhly the findings of the experiment are described, finally the limitations of the experiment are given.

## 8.1 Requirements for Evaluation

Several mapping requirements have been specified for modifying the mapping process and tools to allow ordinary users to engage in the act of mapping, see section 2.8. These requirements have been updated throughout the thesis. Figure 8-1 displays the updated list of the requirements. The specific requirement for the evaluation of the tag-based mapping tool is highlighted in bold.

| **Eliminate** | **Reduce** |
| --- | --- |
| Any Matching API Configuration by User | Complexity of Mapping Process |
| Visual Emphasis on Global View | Tasks per Mapping Session |
| User-Specific Tasks not Related to their Interests | |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | Iterative Mapping Process |
| Visualisation of Information Gained Through Mapping | Familiar & Rich Mapping Interaction |
| | Easy to Understand Interface |
| | Environments for Mapping Sessions |
| | System Mapping Evolvement |
| | **Evaluation Methods** |

***Figure 8-1:*** *Specific Requirements for the Evaluation of the Tag-Based Mapping Tool*

Evaluation methods and metrics have been devised to evaluate whether the tag-based mapping tool is convenient, accessible, simple, efficient, effective and satisfactory for ordinary users, see section 6.5. These methods and metrics are expanded upon in section 8.5 to include evaluation over a long time period.

## 8.2 Study design

In this study, the mapping interactions of four different user groups are observed, recorded and evaluated. The user groups were:

- **Ontologically aware** users are those who have experience using or designing ontologies. These users are representative of the current users of mapping tools (knowledge engineers) and the participants typically used in other user mapping studies, see section 2.7. Thus this group reflects the skill level of knowledge engineers and was chosen to be used as a control group for the study.
- **Technically aware** users are those who have database/UML modelling experience but no ontology experience. This group have modelling skills which can be translated into the skills needed for mapping tools. In the author's opinion, this group also represents a high percentage of web users who are the early adopters of new technologies. This group was chosen to examine how significant the differences are the interactions between ordinary users and users with modelling skills.
- **Non-technical** users are those who have basic computer or web experience with no

database/UML modelling or ontology experience. These users have the lowest level of technical experience and have no relevant modelling skills. However some of these users may be capable of/trained in abstract thinking which gives them necessary experience needed for modelling skills. This group is taken to be a proxy for ordinary users.

- **Other** users are the participants in the study who did not enter their email address into SportsFlows, see section 8.4.3, and thus could not be contacted to answer the post-experiment questionnaire and give their experience. The users in this group could be ontologically aware, technically aware or non-technical users.

However due to the recruitment of participants for the study being open to anybody within the college and participants not being preselected, see section 8.3.2, the number of **active**[33] technically aware users was too small to statistically compare with the other groups, see section 8.4. Thus, only the results of the ontologically aware and non-technical user groups are statistically compared to each other. For convenience the outline for the evaluation approach is displayed again in table 8-1, from section 1.5.2.

**Table 8-1:** Experiment evaluation outline with current experiment (3) highlighted

| **Evaluation Category** | | **Experiment 1** | **Experiment 2** | **Experiment 3** |
|---|---|---|---|---|
| **Efficiency** | | Yes | Yes | Time Taken (section 8.5.1) |
| **Effectiveness** | | Yes | Yes | Precision & Utility (section 8.5.2) |
| **Accessibility & Convenience** | ***Unintrusive*** | No | Yes | Specific Metrics (section 8.5.3) |
| | ***Engaging*** | No | Yes | Specific Metrics (section 8.5.4) |
| **Simple** | | No | Yes | Specific Metrics (section 8.5.5) |
| **User Satisfaction** | | Yes | Yes | Post-Study Questionnaire |

This study is designed to investigate the behaviours and experiences of the four user groups using SportsFlows within their own browsing environment over time. The study is a *field experiment* where the experiment takes place in a naturally-occurring environment for participants[34]. This allowed the experiment to occur within the **user's own computing**

[33] The number of technically aware users who answered a mapping task was only three with one of these users only answering one task over the course of the experiment, see Table 8-3
[34] http://en.wikipedia.org/wiki/Field_experiment

**environment without supervision**. Specifically, in this experiment the questions being investigated are as follows:

> **Q1:** Can non-technical users complete mapping tasks to the standard of ontologically aware users over time within their own work environment given appropriate tool support and an awareness of the domain? (i.e. the time taken for the task, the precision of task, the number of tasks entered, the number of tasks answered once entered)
>
> **Q2:** Does using context-sensitive mapping tasks aid in the non-invasiveness and engagement of the semantic mapping process for non-technical users?
>
> **Q3:** Will non-technical users continue to participate in the semantic mapping process over time?
>
> **Q4:** Does showing the benefits from developing correspondences help to engage non-technical users in the semantic mapping process?
>
> **Q5:** Will the correspondences generated by non-technical users be reliable enough to support the personalisation of sporting information?
>
> **Q6:** Does restricting mapping tasks to those of the non-technical users' own interest and background knowledge reduce the complexity of the mapping tasks for non-technical users?

## 8.3 Hypotheses

From the research questions stated above, several hypotheses were formulated to be evaluated. The mapping performance of users refers to their performance with the tag-based mapping tool: the number of correspondences mapped, the precision of the mapped correspondences, the time taken to complete a mapping task, the percentage of mapping tasks entered and the percentage of mapping tasks answered

> **H1**: There will be no statistically significant difference in mapping performance between the non-technical users and the ontologically aware users, derived from **Q1**;
>
> **H2**: The mapping process will not disrupt non-technical users from their daily work, derived from **Q2**;
>
> **H3**: The mapping tasks will not be time-consuming for non-technical users, derived from **Q2**;

**H4**: There will be no statistically significant drop in the level of mapping performance of non-technical users in the experiment, derived from **Q3**;

**H5:** Non-technical users will see the benefits generated from the correspondences they have developed, derived from **Q4**;

**H6**: Seeing the benefits, from **H5**, will then give non-technical users an incentive to continue participating in the semantic mapping process, derived from **Q4**;

**H7**: Non-technical users will be able to achieve a high mapping precision, derived from **Q5**;

**H8**: Mapping tasks will be viewed as simple to accomplish for non-technical users, derived from **Q6**.

Unlike the other hypotheses, which are evaluated from the user interactions and feedback with the tag-based mapping tool, hypothesis 5 is evaluated from the user interactions with the personal information delivery tool SportsFlows and feedback with the participants.

## 8.4 Method

This section gives the details of how the experiment was carried out. Firstly the details of the participants of the experiment are given. Then the materials used in the experiment are described. Afterwards the procedure of the experiment is detailed. Then a summary is given of the data collected in the experiment. Finally the analysis methods used in the evaluation are explained.

### 8.4.1 Participants

SportsFlows was available for download online to any student or staff member of Trinity College Dublin. A prize draw for an iPod Touch was used as incentive for people to download and use SportsFlows. Recruitment took place by advertising posters in buildings around the college, an advertisement was placed on the college online notice board, and call for participation emails were sent to college sports groups and the computer science mailing list, see Appendix E. This recruitment method guarded against selection bias. A total of 389 people visited the website. 36 people downloaded and installed SportsFlows. 6 people who downloaded SportsFlows did not use the tool after installing it. This left 30 participants who actively used SportsFlows during the experiment. However, only 17 of the participants undertook mapping tasks this was due to the number of mapping task prompts being displayed to the other 13 participants being very low at a mean of approximately one prompt being shown to the 13 participants. The low number of tasks prompted was due to the low number of sport stories viewed by these participants at a mean of less than one story. Due to the low number of stories viewed there were very few prompts shown to these participants which made it hard for them to participate in the mapping process, see section 8.5.1 for more details.

The participants' skills were evaluated through questions asked in the post-experiment questionnaire, see Appendix F. The participants were segmented into three different users groups, based on the answers given to categorise their skills in post-experiment questionnaire. However, only 20 participants entered their email address and thus responded to the survey. So a fourth user group was added for the participants who did not respond to the survey. The user groups are: *Ontologically aware* users who have ontology work experience. There were 6 participants that belonged to this group. This group is labelled in diagrams as onto aware. *Technical aware* users who have database/UML modelling experience but no ontology experience. There were 5 participants who belonged to this group. This group is labelled in diagrams as tech aware. *Non-Technical* users who have basic computer experience with no database/UML modelling or ontology experience. There were 9 participants that belonged to this group. This group is labelled in diagrams as non tech. *Other* users who did not respond to the survey and skills could not be assessed. There were 10 participants that belonged to this group. The evaluation of this experiment analysed the results of all 30 participants and measured the differences in the results between the ontologically aware user group and the non-technical user group. However, only the results of the 17 participants who answered a mapping task could be used in evaluating the mapping results, e.g. the time taken to answer a mapping task.

### 8.4.2 Materials

The experiment involved the participants using the SportsFlows Firefox extension within their browser on their own personal computer. The tool allowed the participants to view sporting stories from sports RSS feeds on the web, see section 7.2.4. The mappings were used to personalise the RSS content shown to them.

The average number of candidate correspondences generated for each participant was 800 by the matching component of SportsFlows, see section 7.1.5.1. Thus, no participant was able to validate all of their candidate correspondences during the experiment. The mapping evolution component of SportsFlows, see section 7.1.5.4, modified the mapping category of correspondences based on their usage which should help to re-categorise incorrectly categorised mappings. The author measured the mapping precision by validating the mappings generated by users.

A post study questionnaire was provided to the participants to evaluate the user satisfaction with SportsFlows. In particular the questionnaires focussed on measuring whether the mapping tool was unintrusive, engaging and simple for users.

### 8.4.3 Procedure

Ethical approval was granted for the experiment in early March 2010, after which participants were sought through posters and emails, see section 8.4.1. The experiment was carried out over a five week period which started in March 2010 and finished in April 2010[35]. Each participant was required to install and use SportsFlows over **at least** a two week period[36] on their own personal computer. This allowed the results of the participants to be compared to the standard set in the previous experiment. The experiment included several steps. The participants were first asked to download the SportsFlows mapping tool from the *SportsFlows website* which required them to accept a disclaimer. The website screenshots are provided in the accompanying DVD media under Experimental Datasets - Experiment 3. An *introduction to SportsFlows* was given on the website describing how to use the tool. However, although it was requested and required that participants view the introduction it was impractical to enforce this and some participants may have decided not to view the instructions. This potentially could have a negative effect on the mapping results. After downloading SportsFlows, each participant was asked by SportsFlows to enter their email address, used to contact participants, and their personal sporting interests, used to generate the participants own personal ontology, see section 7.2.3. The participants were also asked to *use the tool* over the experimentation time period. While the participants were using the tool all their actions with the tool were logged through Ajax http requests to the mapping server. During the experiment mapping task prompts would be displayed to each participant asking them to do a mapping task depending on their browsing context, see section 7.1.4.1. Participants were asked when downloading SportsFlows to answer the mapping tasks prompted but this could not be enforced. They were **only required** to enter their email and use SportsFlows to view sport stories to be in the prize draw. As mentioned previously, this was to allow real world data with the participants to be recorded without them being unfairly influenced to answer tasks just to be in the prize draw. After the experiment, participants who had provided email addresses were requested to fill out a *post-study questionnaire* to evaluate their final thoughts. The questionnaire focused on the usability of the

[35] The call for participation ended after the third week
[36] The participants could have used SportsFlows for up to five weeks

tool with a four point scale used to allow participants express different levels of satisfaction. The same questions as the previous experiment were asked as well as a System Usability System (SUS) questionnaire [Brooke 1996] to score the usability of SportsFlows, see Appendix F. The SUS questionnaire was used to assess the users' satisfaction with the tool as interviews were not carried out in this experiment. There were also questions in the survey which were used to measure the technical expertise of the participants such as 'Do you know what an ontology is?' and 'What department do you belong to (Arts/Computer Science/etc...)?'.

### 8.4.4 Data Collection

The following data was collected: (a) *Post Questionnaire* – used to examine the long-term usability and usefulness of the tool. (b) *Log of users actions* – used to track the participant's behaviour with the tool in particular how active they are in the mapping process. The results for the five weeks were recorded. (c) *Resultant Mapping Precision* – the precision of the user mappings was evaluated by the author. The full log of each user's actions and the set of user mappings are provided in the accompanying DVD media under Experimental Datasets - Experiment 3. (d) *Resultant Evolvement of Mapping Precision* – the mappings developed by the users were also analysed by the tag-based mapping tool and the mapping category for the mapping would change over time based on mapping usage, see section 7.1.5.4.

### 8.4.5 Analysis Approach

The ontologically aware group are the control group for the experiment as they reflect the technical experience of current users of mapping tools. It should be noted that the tag-based mapping tool is not designed and tailored for ontologically aware users as other current state of the art mapping tools are. The tag-based mapping tool implements the mapping process to occur over time within the user's own browsing environment which is not an environment ontologically aware users would be familiar with for mapping and uses a tagging mapping interaction which would also be unfamiliar to these user. However the ontologically aware group have the relevant skills for mapping which should be able to translate into skills needed for the tag-based mapping tool. Furthermore, three of the ontologically aware users were participants in the last experiment and should have gained the necessary skills for the tool. There were no other participants in this experiment that were participants in the previous experiment.

The main quantitative data analysed in this experiment were the mapping task time, mapping precision, percentage of mapping tasks entered and percentage of mapping tasks completed.

Several statistical tests, see section 1.5.3, were carried out on the data. Any significant result from these tests will be highlighted while the full results can be found in Appendix A. A post study questionnaire was used to measure the user satisfaction with SportsFlows. In addition, the qualitative data from the survey helped in revealing the participants' opinion of the tag-based mapping tool.

## 8.5 Evaluation

This section outlines the results from the experiment. The evaluation is split into the five areas, displayed in table 8-1, as follows: First, the efficiency of the tag-based mapping tool is calculated. Next the effectiveness of the tag-based mapping tool is measured. Following this the non-invasiveness of the tag-based mapping tool is evaluated. Then the engagement of both the tag-based mapping and SportsFlows tool is measured. Finally the analysis of whether the tag-based mapping tool is simple to use is given.

The number of participants per user group was as follows: 6 participants for the ontologically aware user group, 5 participants for the technically aware user group, 9 participants for the non-technical user group and 10 participants for the other user group. Table 8-2 and Table 8-3 give a snapshot of the experiment results for these user groups. Furthermore, the entire set of anonymised data records are provided in the accompanying DVD media under Experimental Datasets - Experiment 3. Table 8-2 gives the results for both the non-technical and ontologically aware user group.

**Table 8-2:** Overall mapping results of Non-Technical and Ontologically Aware user groups. A double dash, --, indicates there is no value in the cell.

| | Days Used | RSS Viewed | Mappings Done | Incorrect Mappings | Total Mapping Task Time (Seconds) | Task Prompted | Task Entered | Task Answered |
|---|---|---|---|---|---|---|---|---|
| **Non-technical user 1** | 27 | 181 | 31 | 11 | 554 | 64 | 41 | 31 |
| **Non-technical user 2** | 16 | 2 | 23 | 5 | 617 | 37 | 24 | 23 |
| **Non-technical user 3** | 21 | 3 | 1 | 0 | 49 | 15 | 1 | 1 |
| **Non-technical user 4** | 27 | 27 | 16 | 8 | 290 | 135 | 17 | 16 |
| **Non-technical user 5** | 18 | 4 | 37 | 9 | 480 | 89 | 43 | 37 |
| **Non-technical user 6** | 7 | 0 | -- | -- | -- | 2 | 0 | 0 |
| **Non-technical user 7** | 8 | 0 | -- | -- | -- | 2 | 0 | 0 |
| **Non-technical user 8** | 1 | 3 | -- | -- | -- | 0 | 0 | 0 |
| **Non-technical user 9** | -- | -- | -- | -- | -- | -- | -- | -- |
| | | | | | | | | |
| **Onto aware user 1** | 20 | 1 | 37 | 6 | 419 | 176 | 38 | 37 |
| **Onto aware user 2** | 6 | 2 | 3 | 0 | 39 | 14 | 5 | 3 |
| **Onto aware user 3** | 21 | 5 | 10 | 0 | 207 | 114 | 10 | 10 |
| **Onto aware user 4** | 24 | 14 | 57 | 1 | 637 | 63 | 58 | 57 |
| **Onto aware user 5** | 8 | 5 | 11 | 1 | 168 | 42 | 11 | 11 |
| **Onto aware user 6** | 9 | 5 | -- | -- | -- | 2 | 0 | 0 |

Table 8-3 gives the results for both the technically aware and others user group. Unfortunately due to the low number of technically aware users who completed a mapping task, it was decided the group was too small for statistical comparison with the ontologically aware and non-technical user groups. The other user group was also not used in statistical comparison as the users in this group had varying levels of technical experience, see section 8.1.

**Table 8-3:** Overall mapping results of Technically Aware and Others user groups. A double dash, --, indicates there is no value in the cell.

| | Days Used | RSS Viewed | Mappings Done | Incorrect Mappings | Total Mapping Task Time (Seconds) | Task Prompted | Task Entered | Task Answered |
|---|---|---|---|---|---|---|---|---|
| **Tech aware user 1** | 20 | 11 | 3 | 0 | 87 | 13 | 5 | 3 |
| **Tech aware user 2** | 16 | 17 | 51 | 1 | 816 | 93 | 56 | 51 |
| **Tech aware user 3** | 6 | 8 | 1 | 0 | 61 | 15 | 1 | 1 |
| **Tech aware user 4** | 11 | 5 | -- | -- | -- | 2 | 0 | 0 |
| **Tech aware user 5** | 3 | 2 | -- | -- | -- | 0 | -- | -- |
| | | | | | | | | |
| **Others user 1** | 20 | 11 | 13 | 5 | 566 | 79 | 16 | 13 |
| **Others user 2** | 15 | 4 | 12 | 1 | 222 | 42 | 13 | 12 |
| **Others user 3** | 11 | 6 | 35 | 1 | 169 | 70 | 35 | 35 |
| **Others user 4** | 31 | 13 | 9 | 4 | 198 | 110 | 9 | 9 |
| **Others user 5** | 4 | 5 | -- | -- | -- | 0 | 0 | 0 |
| **Others user 6** | 16 | 2 | -- | -- | -- | 0 | 0 | 0 |
| **Others user 7** | 23 | 0 | -- | -- | -- | 0 | 0 | 0 |
| **Others user 8** | 8 | 0 | -- | -- | -- | 0 | 0 | 0 |
| **Others user 9** | 20 | 0 | -- | -- | -- | 0 | 0 | 0 |
| **Others user 10** | 11 | 0 | -- | -- | -- | 0 | 0 | 0 |

The participants who did not answer a mapping task had a mean of 12 days using the SportsFlows application during the experiment. So their lack of participation was not as a result of not using SportsFlows. However, the mean for the number of mapping tasks prompted to these participants over the course of the experiment is less than 0.62. In comparison, the participants who completed at least one mapping task had a mean of 68.9.

The cause of this low number was investigated and shown that the inactive participants had only a mean of 1.69 sports stories viewed over the course of the experiment. In contrast, the active mapping participants had a mean of 18.5 sports stories viewed. The lack of sporting stories viewed would have contributed to the lack of any browsing contexts being recorded as mapping contexts, see section 7.1.4.1, and thus mapping task prompts being displayed to these users. So the lack of use of SportsFlows would explain the lack of mapping participation by these participants.

### 8.5.1 Efficiency

The metric taken to measure the efficiency of the tag-based mapping tool is the time taken to map a correspondence. This is the commonly used metric to evaluate the efficiency of mapping tools, see section 2.7. Figure 8-2 below gives the results for the mean time taken to map a correspondence for each user group.

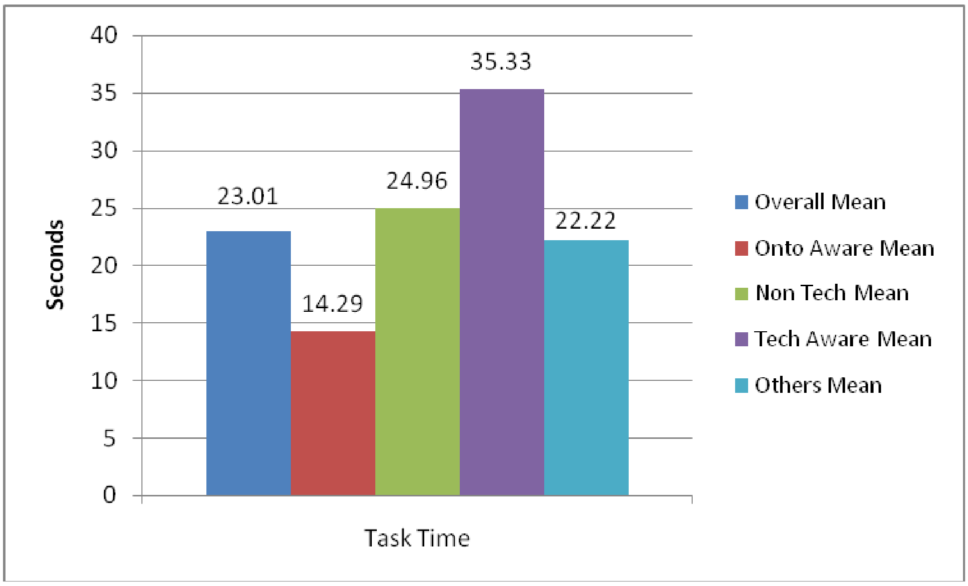



***Figure 8-2:*** *Mean time for Mapping Task*

The ontologically aware group had a mean of approximately 14 seconds for completing a mapping task, technically aware group had a mean of approximately 35 seconds and the non-technical group a mean of approximately 25 seconds. There was no significant difference in the time taken by non technical users and ontologically aware users, see Appendix A. This suggests that non technical users do not take longer to complete a mapping task than ontologically aware users. Furthermore the time taken by non technical users did not significantly increase over the course of the experiment. In fact the mean time for each week is always below 33 seconds for non-technical users which is considerably low for a complex task as mapping. Additionally the results of a one-sample t-test showed that the time taken by the non-technical user group significantly decreased from the standard set by the non-technical user group in the previous experiment ($t = -2.634$, $p = 0.05$). Overall the results from this section suggest that mapping tasks are not time-consuming for non-technical users with the mapping task taking on average approximately 25 seconds to complete.

### 8.5.2 Effectiveness

Since the mappings generated by users can have both correspondences mapped with suggested relations and with new user-defined relations, the effectiveness is evaluated with two metrics: (1) the precision of correspondences that were generated with suggested relations or a combination of user-defined relations and suggested relations, (2) the utility of tags for correspondences mapped with **only** user-defined relations. In addition the mapping evolvement component of the tag mapping tool is also evaluated to assess whether the component improved the precision of the mappings. In total there were 355 correspondences generated by the participants. 76% of these mappings were generated with suggested tags while 24% were generated with user-defined tags.

#### 8.5.2.1 Precision of Suggested Relations

Figure 8-3 shows the mapping precision results, for the correspondences mapped with suggested relations, for each group. The precision was calculated by the author evaluating the validity of each correspondence constructed by users. The user mapping records are provided in the accompanying DVD media under Experimental Datasets - Experiment 3. If the mapping category selected by the user for a correspondence is not the same as the author selected, then the correspondence would be regarded as incorrect. It should be observed that this is the author's own opinion of the correspondences and the mapping category of **certain** correspondences may be viewed differently by multiple people. For example the correspondence between 'Lionel Messi' the footballer and 'Barcelona' the football team was deemed to belong to the different mapping category by the author, as they are different concepts, but it could equally be considered as equivalent sometimes due to the overlap of instances, i.e. sports stories. The results showed that overall 76% of the correspondences mapped were deemed to be accurate.

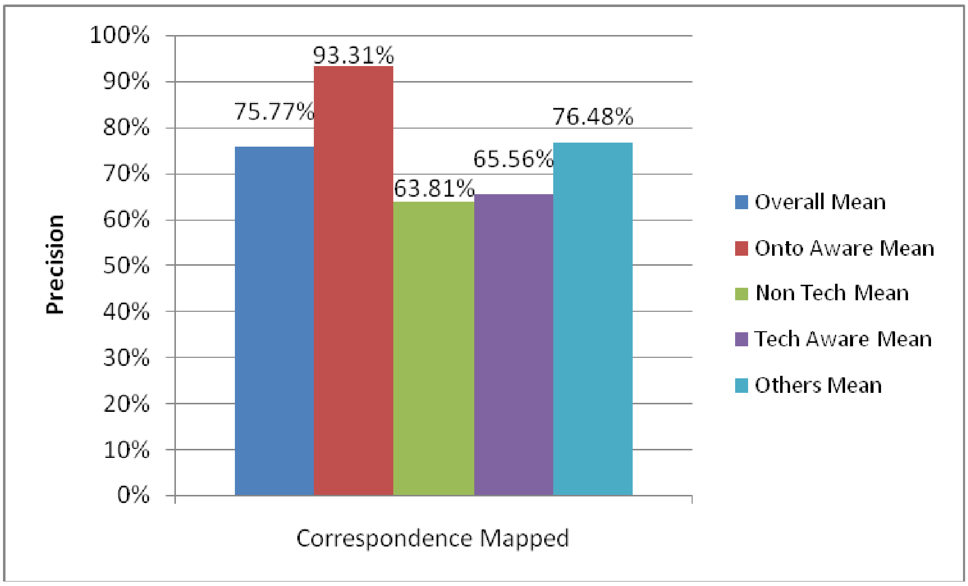



***Figure 8-3:** Mapping precision results for user groups*

Reviewing the precision for the user groups revealed the ontologically aware group achieved the highest precision at 93% while the non-technical group achieved the lowest precision at 64%. The results of an unpaired two-tailed t-test indicated that there was a statistically significant difference between the precision achieved by the ontologically aware user group and the non-technical user group ($t = 2.5578$, $p = 0.05$), see Appendix A. However the precision of the non-technical user group was compared to the overall standard set in the previous experiment and the results revealed there was no statistically significant difference in the precision achieved ($t=-1.141$, $p=0.32$). Furthermore the mapping precision did not decrease significantly over the course of the experiment for non-technical users.

The incorrectly mapped correspondences were reviewed by the author. It should be recalled that although personal ontologies were generated for the users, the ontologies were very simply constructed which may have led to ambiguity about certain generated concepts, see section 7.2.3.4. The following was revealed:

- *28% were correspondences where the concepts have the same parent concept (34% for non-technical users)*: These correspondences were mapped as either equivalent or equivalent sometimes when they should have been mapped as different. An example of this type of correspondence would be 'Rodger Federer' to 'Rafa Nadal' where they would both have 'Tennis Player' as a parent concept;
- *26% were correspondences where the parents of the concepts have a relationship between them (21% for non-technical users)*: These correspondences were mapped as either equivalent or equivalent sometimes when they should have been mapped as

different. An example of this type of correspondence would be ‘Munster’ to ‘Six Nations’. The concepts have no overlapping relationship but Munster has a parent of ‘Rugby Team’ while Six Nations has a parent of ‘Rugby Competition’ and the relationship between these parents is ‘competes in’;

- *24% were correspondences which have no relationship between the concepts (30% for non-technical users)*: These correspondences were mapped as either equivalent or equivalent sometimes when they should have been mapped as different. An example of this type of correspondence would be ‘Celtic Football Club’ to ‘Rodger Federer’ where the concepts belong to different sports and have no overlapping relationship;
- *15% were correspondences where one of the concepts was incorrectly defined (3% for non-technical users)*: These correspondences were mapped as either equivalent or equivalent sometimes when they should have been mapped as different. An example of this type of correspondence would be ‘Leinster’ to ‘Leinster’ where the first concept has the parent rugby club and the second concept has parent football club. However there is no Leinster football club so the user constructed an incorrectly defined concept. After reviewing these correspondences, if the concepts had been constructed properly then the correspondences relation entered by users would have been correct. This either suggested that the users did not look in depth at the concept information given in the tasks or that the users realised the mistake of the other user;
- *4% were correspondences where the concepts have a relation between them (6% for non-technical users)*: These correspondences were mapped as either equivalent or equivalent sometimes when they should have been mapped as different. An example of this type of correspondence would be ‘Rob Kearney’ to ‘Leinster’ which could have had a relationship of ‘plays for’;
- *3% were correspondences which should have been mapped as equivalent sometimes (6% for non-technical users)*: These correspondences were mapped as either equivalent when they should have been mapped as equivalent sometimes. An example of this type of correspondence would be ‘Newcastle United’ to ‘United’. These concepts have the same parent concept of ‘Football Team’ and can even be referred to as the same football club however there are multiple football teams which are also called united and referred to as united. Thus the mapping category should be equivalent sometimes as the term united may refer to multiple teams. This error may have occurred due to the oversight of not having a superclass tag in the suggested relation list, see section 7.1.5.3.

The above list suggests that the users may have been trying to suggest relationships or instance information overlap between concepts by using the equivalent and equivalent sometimes. However in several cases the users was just incorrect in their mapping decision.

#### 8.5.2.2 Utility of User-Defined Relations

Investigating how many user-defined relations were added revealed that: 34% of the mappings generated by the ontology aware group were mappings with user-defined tags, 38% of the technical aware group, 21% of the non technical group, 0% of the unknown group. This indicates a more even spread of the user groups adding user-defined relations than the previous evaluation, see section 6.4.2. The author reviewed the set of user-defined relations added by the participants, the user mapping records are provided in the accompanying DVD media under Experimental Datasets - Experiment 3. The results revealed that 97% of the user-defined relations added made sense for defining a relationship and 91% accurately defined the relationship between the concepts. Furthermore the user-defined relations were used on correspondence where there was ambiguity between the concepts with no exact suggested relation which the user could use, see Figure 8-4 for the breakdown of correspondences mapped. The user-defined relations can be categorised into the following categories:

- *Meaning of a suggested tag (1%)*: tags which have the same meaning as a suggested tag, for example 'can be equivalent',
- *Tag giving semantic meaning (59%)*: tags which correspond to a property, for example 'has player',
- *Incorrect Mappings (3%)*: tags users have used to indicate that the mapping is wrong, for example 'don't know',
- *Other (37%)*: tags which indicate a general correspondence, for example 'is similar to'.

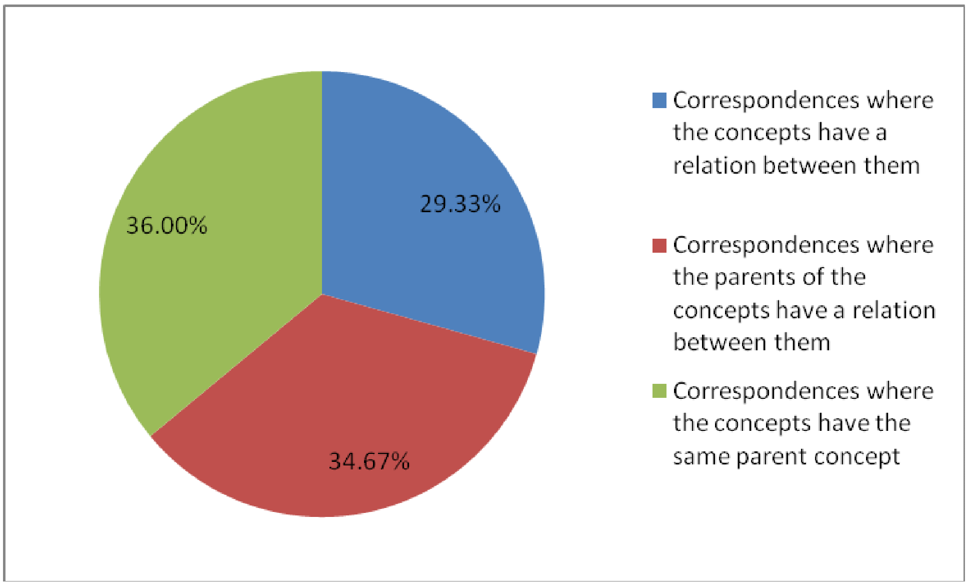



***Figure 8-4:** Breakdown of correspondences mapped with user-defined relations*

#### 8.5.2.3 Evolvement of Correspondence

The mapping evolvement component of the tag-based mapping tool was used to improve the precision of the correspondences mapped by the users and to also define the mapping category for correspondences mapped with user-defined relations, see section 7.1.5.4. The component evolves the mapping category of the correspondences based on the user's use of the correspondence over time. The author reviewed the resultant mapping categories of the correspondences after the mapping evolvement component changed the category to assess the gain in performance. The results revealed that instead of increasing the precision of the correspondences mapped with suggested relations it actually decreased the precision to 47% overall. Furthermore, the precision of the correspondences with user-defined relations was only 11%. These results are extremely low but upon investigation the cause of these results was due to the lack of sport stories viewed by the users. Only 4% of the sporting news stories that were added by correspondences were looked at by the participants. The decrease in precision could have additionally affected the participation and thus engagement of the users.

### 8.5.3 Unintrusive

In this section, first the analysis of the unintrusive results is presented. Subsequently, some conclusions from this evaluation are given. The quantitative metrics and qualitative feedback for a task to be unintrusive were first defined by the author in section 6.5.3.1. Any changes made are highlighted below. The quantitative metrics and qualitative feedback are:

- **Quantitative metrics**
    - *Percentage of mapping tasks entered*: an unintrusive task will have a high rate of response to task prompts **and** the percentage of tasks entered will not decrease significantly over time;
    - *Mean mapping task time*: an unintrusive task will be quick to complete **and** the task time will not increase significantly over time;
- **Qualitative feedback**
    - *Response to the questions about "How disruptive were the mapping tasks to your work?" measured through post-study questionnaire*;
    - *Response to the questions about "How interfering were the mapping tasks to your work?" measured through post-study questionnaire*.

#### 8.5.3.1 Analysis

The mean mapping task time was evaluated in section 8.5.1 and the results revealed that mapping tasks were quick to complete for non-technical users. Figure 8-5 below gives the results for the number of mapping task prompts clicked upon (mapping tasks entered). The results give an indication of the participation rate of each user group and to a lesser extent the perceived utility of the task for participants within each group.

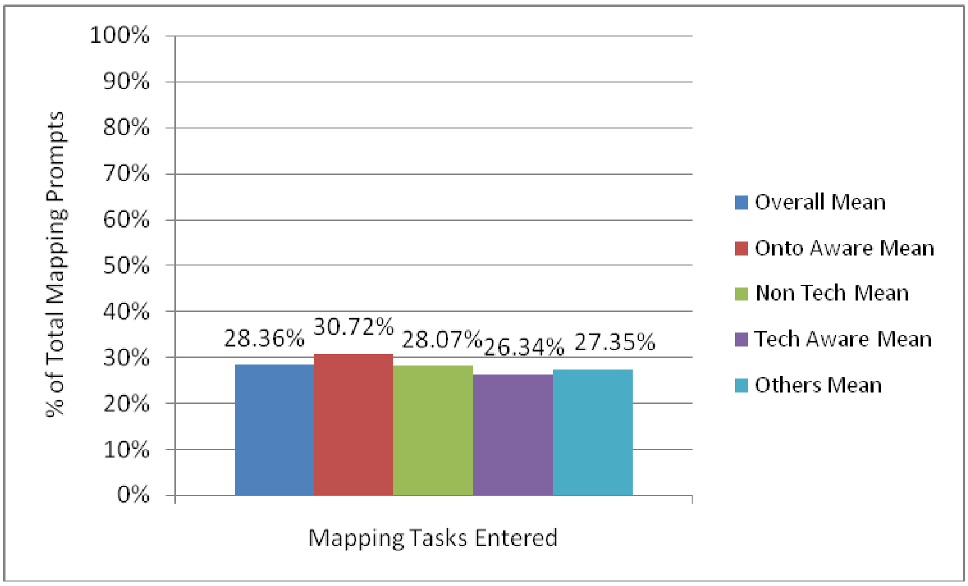



***Figure 8-5:*** *Mapping tasks entered results for each user group*

The results reveal that overall just over 26% of the mapping task prompts were entered. The ontologically aware group had a mean of approximately 31%, technically aware group had a mean of approximately 26% and the non-technical group a mean of approximately 28%. There was no significant difference in the percentage of mapping tasks entered between user groups, see Appendix A, despite the low entered rate of just 28% for the non-technical user group. However the entered rate is higher at 38.81% when the non-technical users who did not answer any tasks are excluded (the excluded users only saw a mean of 1.33 prompts in total over the course of the experiment). Additionally the percentage of mapping tasks entered by non technical users did not significantly decrease over the course of the experiment, see Appendix A. To assess why the participation rate was low, the users were asked in the post-study questionnaire the reasons why they did not click on a mapping task prompt. The users could choose multiple answers from a list. Figure 8-6 below displays both the list of answers allowed and the results given by the users.

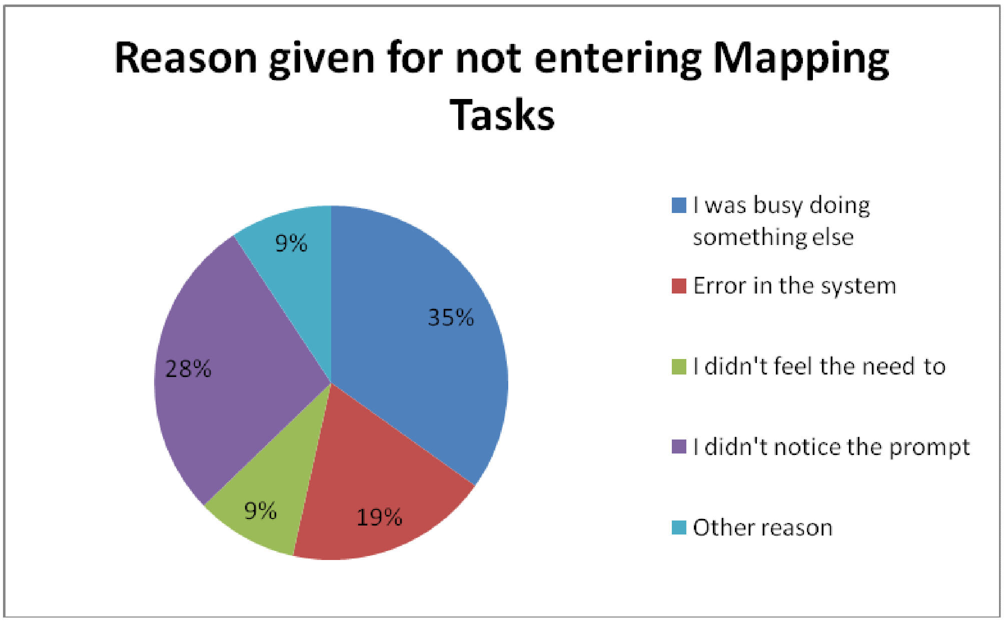



***Figure 8-6:*** *Survey results for ignored mapping tasks*

Only 9% of the response selected "they didn't feel the need to" which is quite a low response and highlights there may be a willingness to undertake the mapping tasks. 28% highlighted that they did not notice the prompt for the mapping tasks. This highlights how easy it can be for participants to miss the prompt. Alarmingly, 19% of the response mentioned that when they clicked on the prompt nothing happened. Upon investigation it was discovered that there was a software error with the Firefox extension of the tag-based mapping tool. The problem occurred when the browser was opened by the users, the extension would then log onto the mapping server, see section 7.1.5, but on certain occasions the extension would not log onto the mapping server. The extension would then appear to work correctly as if it had logged on and would display personalised sporting stories and log the user's interactions. The issue would then occur when the user would click on a mapping prompt and the call to the server would be rejected, as the extension had not logged on, then nothing would happen in the browser. The author investigated two of the ontologically aware users, who had encountered this error, and it was observed that the issue of the extension failing to logon was due to their browser having multiple web pages saved on different tabs which would load every time the user opened the browser. This would lead to multiple website requests which led to the call to the mapping server being occasionally lost due to the volume of requests by the browser and thus the extension would fail to logon. Unfortunately this was spotted late in the experiment and an updated version of the tool could not be sent to the users. This could have led to a loss of trust with the users and the tag-based mapping tool which may have had a negative effect on the mapping results. However the results of the experiment were up to the standard set in the previous experiment, see table 8-7, which suggest it may have had minimal effect on the users.

Worryingly, 35% of the responses said that they were busy doing something else which indicates that the time and browsing context in which the mapping task was prompted was undesirable. This may suggest that certain browsing contexts which were registered for the participants as mapping contexts should not have been. This might be due to the aggressive browsing context registration by SportsFlows, i.e. any time a participant viewed a sports news story the previous browsing page would be registered, see section 7.1.4.1. On the other hand, the browsing contexts could have been fine but the user may have been busy at the time the prompt occurred.

In the post-study questionnaire a set of questions was included to evaluate how unintrusive the tag-based mapping tool was perceived to be. These questions focused on gathering the user's opinion on how disruptive and distracting the mapping tool was. The same questions were also asked in the previous experiment, see Appendix D and Appendix F. The response by the non-technical users were to the same standard set in the previous experiment, see Appendix A. In fact the responses were quite high in both experiments and indicated that the non-technical users did not find the tag-based mapping tool to be disruptive or distracting to them in their workplace.

#### 8.5.3.2 Conclusion

Overall the results indicate that the mapping tasks were not hampering the non-technical users' daily work given that the mapping tasks were quick to complete and their feedback suggested that the tool was not disruptive or distracting. The percentage of mapping tasks entered for non-technical users was low in this experiment. However, there was no statistically significant decrease in performance compared to the standard set in the previous experiment. The feedback from the questionnaire indicated that the mapping prompts were still displayed at times when the participants were busy. These results show that there was no evidence of context-sensitive mapping task prompts, which were introduced in section 7.1.4.1, making any improvement over time-interval based mapping tasks prompts, which were introduced in section 5.2.2. On the other hand, there is also no evidence that context-sensitive mapping task prompts decreased the performance. The feedback also revealed there were a high number of users who did not notice the prompt. This indicates that the prompt may be too subtle and this may have affected the percentage of mapping tasks entered. Overall the feedback and results indicated that the mapping tasks were unintrusive to the non-technical users.

### 8.5.4 Engaging

In this section, first the analysis of the engaging results is presented. Subsequently, some conclusions from this evaluation are given. The specific quantitative metrics and qualitative feedback for an engaging task were defined by the author in section 6.5.4.1. Any changes made are highlighted below. The quantitative metrics and qualitative feedback are:

- **Quantitative Metrics**
    - *Number of correspondences mapped*: an engaging task will have a high number of correspondences mapped **and** the number of correspondences mapped will not decrease significantly over time;
    - *Percentage of mapping tasks entered*: an engaging task will have a high percent of response to task prompts **and** the percentage of tasks entered will not decrease significantly over time;
    - *Percentage of mapping tasks finished*: an engaging task will have a high percent of tasks answered once entered **and** the percentage of tasks finished will not decrease significantly over time;
    - *Number of RSS sport stories viewed*: an engaging task gives benefits to the user in a functional, social or emotional dimension. With SportsFlows the user will view a high number of RSS sport stories **and** the number viewed will not decrease significantly over time;
- **Qualitative Feedback**
    - *Response to the questions about "How appropriate were the sports information and did this help motivate you to answer more tasks?" measured through a post-study questionnaire.*

#### 8.5.4.1 Analysis

The number of correspondences mapped by the users will give an indication of how engaged the users were with the tag-based mapping tool. The overall number of mapping tasks completed by the users was 20.58 with approximately 1.27 mapping task done per day by users. This is quite high and suggests users were participating in the mapping process over time. The non-technical users had a mean of 1.07 mapping tasks completed per day while the ontologically aware users had a mean of 1.32 mapping tasks completed per day. There was no significant difference in the number of mapping tasks completed by non technical users and ontologically aware users, see Appendix A. Indeed the number completed by non technical users did not significantly decrease over the course of the experiment.

The percentage of mapping tasks entered was evaluated in section 8.5.3.1 and the results revealed that the percentage was low but still comparative to the results of the previous experiment. Figure 8-7 below gives the results for the percentage of mapping tasks answered by the users once the mapping task was entered. The results showed that overall 90% of the mapping tasks entered by the users were answered.

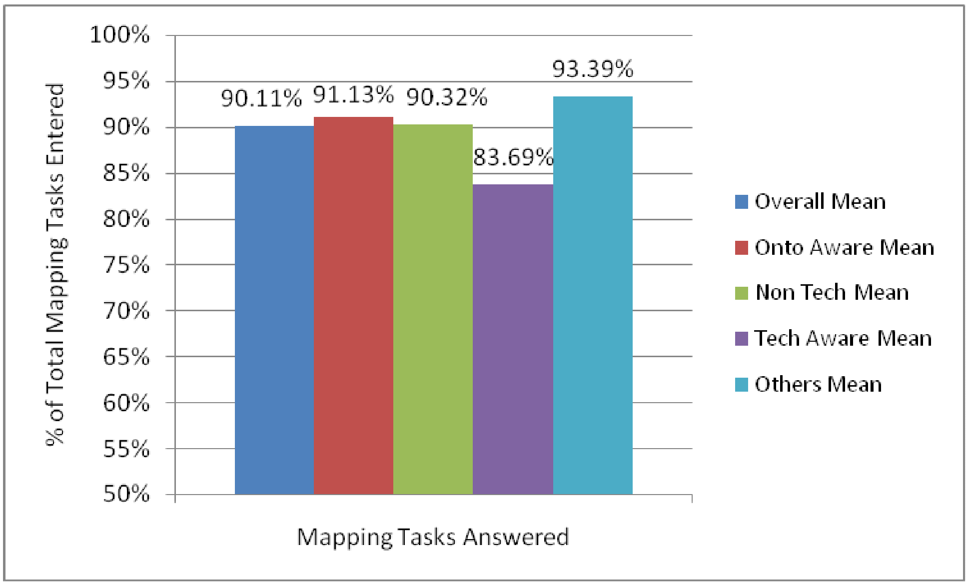



***Figure 8-7:*** *Mapping tasks answered results for each user group*

The non-technical users had a mean of 90% mapping tasks answered while the ontologically aware users had a mean of 91% answered. This reveals a high number of mapping tasks were answered by the non-technical users once the task were entered. There was no significant difference in the percentage of mapping tasks answered by non technical users and ontologically aware users, see Appendix A. Additionally the percentage did not significantly decrease over the course of the experiment for non technical users. Furthermore there is a statistically significant increase in the percentage of mapping tasks answered by non-technical users in this experiment compared with the standards set in the previous experiment (t=4.477, p =0.01). This is slightly surprising as this experiment is more of a real case study than the previous experiment. Although the improvement may have been due to the reduced number of mapping tasks per mapping sessions and the tag-based mapping tool being fixed from a serious issue of the mapping server running slow, see section 6.9. However the high percentage of mapping tasks entered by non-technical users indicates a high participation rate for non-technical users once the task has been entered.

The number of RSS stories viewed by the users is quite low at only 0.72 stories being viewed per day by the users. This suggests that the SportsFlows application was not providing the users with appropriate benefits. The non-technical users had a mean of 1.4 stories while the ontologically aware users had a mean of 0.40 stories. There was no significant difference in the number of RSS stories viewed by non technical users and ontologically aware users, see Appendix A. Additionally the number viewed did not significantly decrease over the course of the experiment for non technical users. Examining the percentage of RSS stories viewed by the participants showed that only 11% of the stories were viewed (336 stories out of 3133). This is an extremely low number and indicates that the user were finding little to no benefit with the SportsFlows application. However it could also be possible that the prompts for the sport stories were too subtle as a large percentage of the feedback for why users did not click on a mapping session prompt indicated they did not notice the prompt, see section 8.5.3.1. Furthermore it is possible the user may have already viewed the story being prompted. Additionally the headline of the prompt may have contained enough information of the story for the user. In the post-study questionnaire a set of questions was included to evaluate how appropriate the sport news information was for the users. These questions were also asked in the previous experiment. The response by the non-technical users were to the same standard set in the previous experiment, see Appendix A. The ratings for the questions were low and also indicate that the users are not finding any benefits with the SportsFlows application.

Upon further examination of the percentage of the stories viewed by the users, it was revealed that only 4% of the sporting news stories which were added by mapped correspondences were viewed by the participants. This is again an extremely low number and indicates that the user were finding little to no benefit with the correspondences they constructed. However as discussed above this was a general problem with the SportsFlows application rather than just with the correspondences generated. In the post-study questionnaire a set of questions was included to evaluate the benefits of the mapping tasks to the users. These questions focused on evaluating if the participants felt the mappings were bringing benefits and if it helped in motivating them. These questions were also asked in the previous experiment. The response by the non-technical users were to the same standard set in the previous experiment, see Appendix A. In fact the responses were quite low and supported the assumption that the users were not seeing any benefits from the SportsFlows application.

#### 8.5.4.2 Conclusions

It was a clear in this experiment that the SportsFlows application was bringing little or no benefits to the users. As there were no benefits viewed from the correspondences constructed by the users, they had less incentive to complete mapping tasks so they could have become less engaged. Additionally another potential reason for the participants' lack of motivation to use the tool may have been due to their lack of interest in sports. However, although the percentage of mapping tasks entered is low both the number of correspondences generated and the percentage of mapping tasks answered by non-technical users is high. This indicates that non-technical users can become engaged in the mapping process. It should also be observed that the tag-based mapping tool only produces a partial mapping, which is continually updated by the user over time, and this could have also potentially affected the user's participation.

### 8.5.5 Simple

In this section, first the analysis of the results is presented. Subsequently, some conclusions from this evaluation are discussed. The specific quantitative metrics and qualitative feedback for a simple task were defined by the author in section 6.5.5.1. Any changes made are highlighted below. The quantitative metrics and qualitative feedback are:

- **Quantitative Metrics**
    - *Mean mapping task time*: a simple task will be quick to complete **and** the task time will not increase significantly over time;
    - *Number of correspondences mapped*: a simple task will have a high number of correspondences mapped **and** the number of correspondences mapped will not decrease significantly over time;
    - *Precision of correspondences mapped*: a simple task will have a high percentage of correspondences mapped correctly **and** the percentage will not decrease significantly over time;
- **Qualitative Feedback**
    - *Responses to the questions about "Are the questions and supplied information confusing?" measured through the post-study questionnaire.*

#### 8.5.5.1 Analysis

The mean mapping task time was evaluated in section 8.5.1 and the results revealed that mapping tasks were quick to complete for non-technical users with a mean of just 24.96 seconds. The number of correspondences mapped was evaluated in section 8.5.4.1 and the results revealed that the number was relatively high for non-technical users with a mean of 1.07 mapping tasks completed per day. The precision of correspondences mapped was evaluated in section 8.5.2 and the results revealed that the precision of correspondences, mapped with suggested relations, for non-technical users was low at a mean of 63.81%, but the precision was still comparative to the standard set in the previous experiment. The correspondences could also be mapped with user-defined relations and the results showed that 97% of the user-defined relations added by the users were useful for defining the relationship of the correspondence.

In the post-study questionnaire a set of questions was included to evaluate how simple the task of mapping correspondences was for users. These questions focused on evaluating if the users felt the mapping task was confusing. The same questions were also asked in the previous experiment. The response by the non-technical users were to the same standard set in the previous experiment, see Appendix A. In particular, the responses from the non-technical users were highly positive and indicated that these users did not find the task of mapping to be confusing.

The post-study questionnaire included the questions from the SUS questionnaire which was used to evaluate the user satisfaction with the tag-based mapping tool. Scores in SUS range from 0 to 100, a higher score being better. The tag-based mapping tool had an average score of 79 (64 for non-technical users). As usability is not an absolute criterion the resulting score can only be understood relatively to others [Nielsen 1993][Preece 2002]. Another SUS evaluation on two semantic mapping tools resulted in average SUS scores of 73 for CoGZ and 46 for PROMPT [Falconer 2009]. In the authors' opinion, the SUS scores of the non-technical users in this experiment is relative to the SUS scores of these mapping tools and indicate that non-technical users found the tag-based mapping tool appropriate for mapping tasks. This is a very good result considering that the non-technical users were unfamiliar with ontology mapping tools before the experiment and that "they were thrown in the deep end of ontology mapping tasks".

#### 8.5.5.2 Conclusions

The results illustrated that non-technical users can complete the mapping tasks to a high standard, although the precision achieved was not as high as the ontologically aware user group.

The results also revealed that the time spent by the users doing the mapping tasks was quite acceptable and was not time consuming. The responses of non-technical users from the post-study questionnaire indicate that these users found the mapping tasks straightforward and easy to achieve.

## 8.6 Overall Conclusions

This section gives a summary of the analysis of the experiment. Table 8-4 displays the evaluation of the experiment hypotheses and the subsequent subsections discuss. The mapping performance of users refers to their performance with the tag-based mapping tool: the number of correspondences mapped, the precision of the mapped correspondences, the time taken to complete a mapping task, the percentage of mapping tasks entered and the percentage of mapping tasks answered.

**Table 8-4:** Summary of the Analysis of the Hypothesises

| **Hypothesis** | | **Suggests support for** | **Evidence** |
|---|---|---|---|
| **H1** | There will be no statistically significant difference in mapping performance between the non-technical users and the ontologically aware users | Maybe | Section 8.5.1 – 8.5.5 |
| **H2** | The mapping process will not disrupt non-technical users from their daily work | Yes | Section 8.5.3 |
| **H3** | The mapping tasks will not be time-consuming for non-technical users | Yes | Section 8.5.1 |
| **H4** | There will be no statistically significant drop in the level of mapping performance for non-technical users in the experiment | Yes | Section 8.5.1 – 8.5.5 |
| **H5** | Non-technical users will see the benefits generated from the correspondences they have developed | No | Section 8.5.4 |
| **H6** | Seeing the benefits, from **H5**, will then give non-technical users an incentive to continue participating in the semantic mapping process | Maybe | Section 8.5.4 |
| **H7** | Non-technical users will be able to achieve a high mapping precision | Maybe | Section 8.5.2 |
| **H8** | Mapping tasks will be viewed as simple to accomplish for non-technical users | Yes | Section 8.5.5 |

### 8.6.1 Mapping Performance of Non-Technical Users

The first question of the experiment to be investigated was *whether non-technical users can complete mapping tasks to the same standard as ontologically aware users*. The first hypothesis **H1** was designed to address this question. This hypothesis stated that there would be no significant difference in mapping performance between the non-technical and the ontologically aware users. The only significant difference in performance between the user groups was the mapping precision achieved by the ontologically aware users. However, the ontologically aware user group's mapping precision result is extremely high while the non-technical user group mapping precision was comparative to standard set in the previous experiment. Additionally, the percentage of mapping tasks answered by non-technical users was statistically significantly better than the standard set in the previous experiment.

The third question of the experiment investigated *whether non-technical users continue to participate in the semantic mapping process over time*. The fourth hypothesis **H4** was aimed at addressing this question. The fourth hypothesis stated that there will be no statistically significant drop in the level of mapping performance for non-technical users in this experiment. The results revealed that there was no drop in mapping performance for the non-technical users over the course of the experiment. Thus the non-technical users continued to participate in the mapping process over time. Furthermore, the results showed that mapping performance of the non-technical user group was up to the standard set by the previous experiment.

### 8.6.2 Mapping Process Intrusiveness

The second question of the experiment aimed at evaluating *whether context-sensitive mapping tasks aid in the non-invasiveness of the semantic mapping process for non-technical users*. The second and third hypotheses (**H2** and **H3**) were designed to address this question. The second hypothesis stated that context-sensitive mapping tasks will not disrupt non-technical users from their daily work. The evaluation of how unintrusive the tool is, see section 8.5.3.1, indicated that the mapping tasks were not hampering the web users daily work from the feedback given by the participants. However, the results revealed that the context-sensitive mapping prompts were shown at times when the participants were busy which caused a reduction in the number of mapping tasks entered. The third hypothesis stated that the mapping tasks will not be time-consuming for non-technical users. In the evaluation the mapping task time was revealed to be very low at a mean of approximately 25 seconds for the non-technical user group, see section 8.5.1. Overall, there are indications that the mapping process, using a context-sensitive approach, is non-invasive for non-technical users.

### 8.6.3 Mapping Process Engagement

The fourth question of the experiment aimed at examining *whether the benefits from developing correspondences help to engage non-technical users in the semantic mapping process*. The fifth and sixth hypotheses (**H5** and **H6**) were aimed at addressing this question. The fifth hypothesis stated that the non-technical users would see the benefits generated from the correspondences they have developed. The responses given by non-technical users for the post-study questionnaire, see section 8.5.4.1, revealed that from the users' perspective the correspondences they generated were bringing little if any benefits. Additionally, the response also revealed that the users found the appropriateness of the sports stories shown to be low. This explained the low number of stories viewed by the participants at just 11% of the stories prompted, see section 8.5.4.1. This also indicated that the non-technical users were finding little benefits from the SportsFlows application. It should also be observed that the tag-based mapping tool only produces a partial mapping which could have also potentially affected the user's participation with SportsFlows. The sixth hypothesis stated that seeing the benefits, from **H5**, will give non-technical users the incentive to continue participation in the semantic mapping process. As there were no benefits being viewed from the mappings the participants had less incentive to complete mapping tasks and should have become less engaged. However, the number of correspondences generated and the percentage of mapping tasks answered are quite high which indicates the non-technical users were engaging in the mapping process. Overall, although the SportsFlows application brought little benefits to the non-technical users, these users still participated in the mapping process over time.

### 8.6.4 Mapping Process Simplicity

The fifth question of the experiment investigated *whether the correspondences generated by non-technical be reliable enough to support the personalisation of sporting information*. The seventh hypothesis **H7** was aimed at addressing this question. This hypothesis stated that non-technical users will be able to achieve a high mapping precision. The mapping precision achieved by the non-technical users was shown to be statistically significantly lower than what the ontologically aware users achieved, see section 8.5.2.1. However the ontologically aware users precision results was extremely high while the non-technical users precision results was relative to the results achieved in the previous experiment, see section 8.5.2.1. The majority of the user-defined tags added by the users were also shown to be useful for defining the relationship of the correspondences, see section 8.5.2.2. The percentage of sports stories viewed from correspondences is extremely low at only 4%, see section 8.5.4.1. This suggests the correspondences generated by the users were not reliable enough. However the mapping precision was decreased by the mapping evolvement component of the tag-based mapping tool incorrectly re-categorising the mapping category of correspondences, see section 8.5.2.3. In addition, the prompts were also shown to be too subtle, see section 8.5.3.1, and as mentioned before the users were not using the SportsFlows application.

The final question of the experiment aimed to *examine if restricting the mapping tasks to the non-technical users own interests and background knowledge reduce the complexity of the mapping task for non-technical users*. The eighth hypothesis **H8** was aimed at addressing this question. This hypothesis stated that the mapping tasks will be viewed as simple to accomplish for non-technical users. The tag-based mapping tool restricted the mapping tasks to tasks that were related to the users sporting interests. The response given to the post-study questionnaire in the evaluation, see section 8.4.5, revealed that the non-technical users found the mapping tasks to be straightforward and easy to achieve. The mean time taken to complete a mapping task was also very low adding weight to this assertion. Overall, the results indicate that non-technical users found the mapping task straightforward and mapped correspondences to a high standard.

## 8.7 Limitations

The number of subjects in the experiment was low, even though the experiment was designed to have a large participation. The low uptake of SportsFlows can be possibly explained with the following: (1) SportsFlows required the individual to use the Firefox browser, (2) Participants needed to download and install software in their browser which may have discouraged people from taking part, (3) The experiment required the person to use SportsFlows for at least two weeks which is a long time period and again may have discouraged people from taking part, (4) The experiment started three weeks before the teaching term ended and exams begun, (5) The teaching semesters were changed in the year of experimentation which led to a reduction in the number of teaching weeks and thus an increase in the volume of work the students had to perform in the final weeks of term. This impacted the numbers of users per user group and the sample size of the experiment.

Although there was no bias in the selection of participants for the experiment, as the participation was open to anyone within the college, there was potential bias in the selection of the sport domain which may be easier to understand than other domains, e.g. biology domain. However using the sports domain allowed for a greater number of people to participate than with a more complex domain. Additionally the ontologies used were simple and did not contain any complex relationships such as the disjoint relationship due to how they were devised, see section 7.2.3. However a large portion of the correspondences generated were cognitively tough to answer. Indeed the author validated the correspondences which may have also caused bias based on the point of view of the author. Finally the prompts were only displayed for 8 seconds to the participants which may have impacted their participation as it was revealed that several participants did not see multiple prompts.

There was an intermittent error with the Firefox extension of the tag-based mapping tool logging onto the mapping server. The issue occurred if the browser was opened with multiple web pages which could have led to the extension not logged on. The participant would then be unable to answer a mapping task. Unfortunately this was spotted late in the experiment and a fix could not be sent to the participants. This could have led to a loss of trust with the participants and SportsFlows which may have had a negative effect on the mapping results. An initial pilot study may have identified this error as well as any other issues there may have been, e.g. operating system incompatibilities.

## 8.8 Summary

In this chapter the final user evaluation is presented. This experiment evaluates whether it is possible for ordinary users to actively participate in the semantic mapping process within their own workplace over a long time period. The study analyses the mapping results of four different user groups with different technical expertise.

The experiment highlighted that the mapping process can be non-invasive to ordinary users within their own workplace over time. However there was a lack of engagement from ordinary users which was possibly due to there being few benefits gained from the SportsFlows application and the prompts being too subtle. Restricting the mapping task to the ordinary users' own background knowledge helped in simplifying the mapping task for these users.

In the next chapter, a summary of the three user evaluation studies presented in this thesis is given with emphasis on if the research question of the thesis was achieved.

# 9 Evaluation Summary

This chapter presents in summary the key points from analysis of the results arising from the experiments carried out in the course of writing this thesis. First, for convenience, a summary of each experiment is provided. Then the subsections of the chapter present the key points that have been identified arising from the experiments with respect to the key characteristics of the mapping process that should be supported, as argued in the research question.

The first experiment was a laboratory study on the approaches of two mapping tools with three groups of users. The three user groups were: (1) *Ontologically Aware* users who have experience of working with ontologies and represent the skill level of the current users of mapping tools, knowledge engineers, (2) *Technically Aware* users who have modelling skills but no ontology experience, and (3) *Non-technical* users who only have basic computing experience with no modelling skills and are taken to be a proxy for ordinary users. The two tools used were COMA++ which reflected the mapping approach taken by the state of the art in mapping and NL prototype mapping tool which was an initial prototype mapping tool designed for ordinary users. This experiment was primarily a behaviour study on ordinary users using mapping tools, as it was critical to observe these users tackle the mapping problem to gain an understanding of where they found problems with the process, approaches and tools. The evaluation results contributed to the design and implementation of a tag-based mapping tool for ordinary users.

The second and third experiments address whether the tag-based mapping tool satisfied the requirements from the research question. These experiments were field studies of users using the tag-based mapping tool within their own work environment over time. The second experiment separated the users into the same user groups as the first experiment and these users were required to use the tool over a two week period. The users were chosen by the author and the number was small to allow for interviews to be carried out post-experimentation. Modifications were made to the tag-based mapping tool based on the evaluation of the second experiment. The third experiment separated the users into the same user groups as the preceding experiments and the users were required to use the modified tag-based mapping tool over at least a 2 week period during the 5-week period of the experiment. In contrast to the other experiments, this experiment was open for participation by anybody within Trinity College Dublin. Thus the number of users was unknown before experimentation. The users needed to answer a post-study questionnaire so their skills could be assessed and assigned to a user group.

However several users did not answer the questionnaire and could not be assigned which led to another user group being formed with these users called the *other* user group.

The research question stated:
*How can the act of semantic mapping become accessible, convenient, simple, effective, efficient and more satisfactory for ordinary users to achieve semantic mappings gradually and over time between semantic models of interest to the user?*

The mapping task that users were required to complete in the experiments was to map correspondences. However the task was implemented in a different way in each experiment. The first experiment required users to validate candidate correspondences with COMA++ and the NL prototype mapping tool. The second and third experiment required the user to enter the relationship of the correspondence with the tag-based mapping tool. The summary of how well the tools developed during this thesis satisfied the requirements of the research question is discussed in the following sections. Table 9-1 displays how each experiment was evaluated.

**Table 9-1:** Thesis evaluation outline

| **Evaluation Category** | | **Experiment 1** | **Experiment 2** | **Experiment 3** |
|---|---|---|---|---|
| **Efficiency** | | Section 3.5.1 | Section 6.5.1 | Section 8.5.1 |
| **Effectiveness** | | Section 3.5.2 | Section 6.5.2 | Section 8.5.2 |
| **Accessibility & Convenience** | ***Unintrusive*** | Not evaluated in this experiment | Section 6.5.3 | Section 8.5.3 |
| | ***Engaging*** | Not evaluated in this experiment | Section 6.5.4 | Section 8.5.4 |
| **Simple** | | Section 3.5.3 | Section 6.5.5 | Section 8.5.5 |
| **User Satisfaction** | | Section 3.5.3 | Section 6.5.3 – 6.5.5 | Section 8.5.3 – 8.5.5 |

## 9.1 Efficiency

Efficiency was measured by the time taken to complete a mapping task, in this case mapping correspondences. The results from the first experiment showed that non-technical users took 24 seconds to complete a mapping task with NL prototype mapping. Indeed with COMA++ it took non-technical users 36 seconds to complete a mapping task. However the majority of this time was spent traversing the candidate correspondence as the feedback revealed users were just label matching, see section 3.5.4. This indicates that non-technical users did not use the proper approach for answering mapping tasks with COMA++. In experiment two, the time taken to complete a mapping task with the tag-based mapping tool increased to 42 seconds for non-technical users. This is alarming as the mapping task time nearly doubled from the prototype version. However after the tool and approach had been refined, in experiment three the time taken to complete a mapping task was reduced to 25 seconds for non-technical user. This was a statistically significant decrease from the previous experiment. In addition, there was no statistical significant difference between the mapping task time taken for the non-technical user and for the ontologically aware users in each of the experiments. This suggests that the speed to complete a mapping task with the tag-based mapping tool is not dependent on the technical skills of the users. Furthermore, the mapping task time taken for non-technical users in experiment three is extremely low considering that mapping is a complex task.

In the author's opinion, the time spent answering a mapping task by ordinary users is not a lot of time and should not take up too much time of the ordinary users work day. In fact it can take up to 20 seconds to open a Firefox browser[37]. In addition, research has shown that users sometimes browse to alleviate boredom, fill time or provide entertainment [Thakor 2004]. In the author's opinion, in these states ordinary users could be willing to take the time to undertake mapping tasks. To add weight to this argument, when Google modified the logo on their search page to play the Pac-Man game people spent on average 36 more seconds on the search page than they normally would playing the game [Wright 2010]. This suggests that ordinary users can be encouraged to spend at least 30 seconds of their time on tasks that interest them.

The time taken to complete a mapping task was also long enough to suggest that users were actually reading the task information and making an informed decision rather than choosing a random answer. In the interviews and questionnaires for both experiments the users responded that they read and understood the information for the mapping task, see section 6.5.5.2 and 8.5.5.1. The mapping precision results also showed that the choices made by users were not random, see section 9.2. In addition, if the users were just choosing random answers then the time taken would have been nearer to 4 or 5 seconds. Overall the results showed that ordinary users can efficiently map a correspondence with the tools.

[37] http://www.technoslot.com/2011/02/how-to-speed-up-your-firefox-browser/

## 9.2 Effectiveness

The effectiveness of the mapping tools used in experiment one, NL prototype mapping tool and COMA++, was measured through the mapping precision and recall of the mapping generated by the users. However the tag-based mapping tool had users map the relationship of correspondences using tags. The mapping precision can be evaluated on correspondences mapped with tags that are suggested by the tool which are called suggested relations. However the mapping precision cannot be calculated on correspondences that are mapped with non-suggested tags, i.e. user-defined relations. Effectiveness was measured by: (1) the mapping precision of correspondences mapped with suggested relations, and (2) the utility of the user-defined relations.

The results from the first experiment revealed that the non-technical users achieved a mapping precision of 72% with the NL prototype mapping tool, see section 3.5.2. The non-technical users achieved a precision of 78% with COMA++. However as mentioned above the non-technical users did not use the proper approach for answering mapping tasks with COMA++. The recall results revealed that the non-technical users achieved a mapping recall of 55% with COMA++ which indicated that these users found it hard to find any missing correspondences with the tool which the feedback from the users confirmed. The mapping recall result for the NL prototype mapping tool was 51% for the users, however the users were unable to append any missing correspondences with the tool. In experiment two, the mapping precision increased to 78% for non-technical users with the tag-based mapping tool, see section 6.5.2. While in experiment three, the mapping precision decreased to 64% for the non-technical users with the tool, see section 8.5.2. There was no statistical significant difference in the mapping precision achieved by non-technical users in the second and third experiment. However there was a statistically significant difference in the mapping precision achieved by non-technical users and ontologically aware users in the first and third experiments. It is worth observing that in the first experiment the tool used was an initial prototype which was not designed for non-technical users and that the precision achieved by ontologically aware in the third experiment was extremely high. In addition, the results from these experiments may suggest that non-technical users cannot achieve the same high level of mapping precision as ontologically aware users.

In the third experiment, the tag-based mapping tool implemented a mapping evolution component which was used as to improve the mapping precision. However the evaluation results revealed that the component was reducing the precision instead of improving it, see section 8.5.2.3. Investigating the results indicated that this reduction was primarily due to the lack of use of the SportsFlows application by users.

The utility of the user-defined tags added by the users was also investigated for the second and third experiments and the results indicated that these tags were useful in defining the relationships for the correspondences. In the third experiment, the mapping evolution component was used to assign the mapping category for the correspondences which had user-defined tags, and thus the mapping category of the user-defined tags. However due to the lack of use of SportsFlows by users, only 11% of these correspondences were correctly assigned. In addition although the mapping category of the user-defined tag may be able to be computed using this method, there is still research needed to define the concept relation category of the user-defined tag, see section 10.2.1.3.

Although the mapping precision results for non-technical users were moderately high in each experiment, there was still a significant error rate which may be unacceptable in certain mapping scenarios. This gives rise to a requirement to improve the precision once the mapping has been generated by the user. One such method is evolving mappings based on their usage. In the third experiment an approach to this method was used, however an issue occurred in miscategorising the correspondences mapping category due to the lack of use of the SportsFlows application by users. Other methods which could improve the precision include the collaborative sharing and/or validation of correspondences [Zhdanova 2006]. Furthermore, in certain mapping scenarios it could be beneficial to allow an administrator to validate the correspondences generated. Overall the results indicated that ordinary users may be able to effectively map correspondences with the tag-based mapping tool depending on the standard required.

## 9.3 Accessibility & Convenience

To allow ordinary users to use a mapping tool in their own work environment the tool must be accessible and convenient for the users to use. For the tool to be both accessible and convenient it needs: the tool and process to be non-invasive to the user as they are carrying out their daily work as the work may be voluntary making it important not to disrupt the user, and engaging enough so that the user will want to interact in the mapping process. The NL prototype mapping tool and COMA++ could not be evaluated under these evaluation areas due to experiment one being carried out in a controlled lab situation. In the second and third experiment, the evaluation of the tag-based mapping tool investigated whether the tool was unintrusive and engaging to ordinary users. The summary of these results is given in the following sections.

### 9.3.1 Unintrusive

*Unintrusive* means that the mapping tool will not disrupt the user from their daily work and the mapping process will blend into the background while they work, i.e. when doing mapping tasks regularly over the time period it does not hamper their work and that in fact they see it as an everyday task. The results of the evaluations revealed that the mapping tool was not disruptive or distracting to the non-technical users. The participation level of the non-technical users did not significantly drop in either experiment two or three over time which indirectly indicates that the mapping tool was not disruptive or distracting otherwise there would have been a significant drop. There was no statistically significant difference between participation results for the non-technical users and the ontologically aware users in both experiments which indicate that technical skill is not an issue. In addition, the time taken to complete a mapping task was low and suggests that the task was not taking too much time for the users, see section 9.1. Furthermore, the feedback from the users indicates that the users did not find the tool to be distracting or interfering. The results also revealed that the participation level of non-technical users did not increase with the context-sensitive mapping task approach in experiment three over the time-based mapping task approach in experiment two. Although this may have been due to the implementation of aggressive context-sensitive mapping tasks which were displayed in contexts which may not have been suitable for mapping. Overall, in the author's opinion, these results suggest that the mapping process can be embedded in the ordinary users browsing environment in a seamless way which will not disrupt the daily work of the users.

### 9.3.2 Engaging

*Engaging* means users will enter and finish mapping tasks regularly over time and will see the benefits from the correspondences generated, i.e. personalised sporting news stories. The number of mapping tasks completed per day by non-technical users is high in both experiments and also shows no statistical difference to the number achieved by the ontologically aware users. The percentage of mapping tasks answered in experiment two was low with only 48% for non-technical users. However the percentage statistically significantly increased to 90% in experiment three after the tool and approach had been refined. In addition, there was no statistical significant difference between the percentage of tasks answered by non-technical users and ontologically aware users in both experiments. As mentioned in the previous section the participation rate did not drop over time. These results indicate that non-technical users were engaged in the mapping process.

However the number of RSS sport news stories viewed by non-technical users was low in both experiments. The feedback from the users revealed that they were not finding any benefits with the SportsFlows application. This indicates that the users did not see the benefits the mappings were bringing and should have negatively affected the engagement. It should also be observed that the tag-based mapping tool only produces a partial mapping, which is continually updated by the user and tool over time, and this could also potentially have impact the personalised sporting information being displayed to the user and thus possible contribute to why users where not finding any benefits. However the results above showed that users were engaging in the mapping process. As mentioned before, a reason may be due to users sometimes browsing to alleviate boredom, fill time or provide entertainment. Overall the results indicated that ordinary users were engaging in the mapping process but, in the author's opinion, there is a need for the results of mapping to be displayed clearly to the user to both maintain and give incentives for the user to participate in the mapping process.

## 9.4 Simple

*Simple* means users will be able to achieve rich mappings in a non time-consuming way that is not confusing but instead is viewed as straightforward to the user. In experiment the feedback from the non-technical users revealed that they found the NL prototype mapping tool easier to use than COMA++, see section 3.5.3. The opinion from the users was that the act of mapping was straightforward with NL prototype mapping tool but was much harder with COMA++. The feedback from the evaluation of the tag-based mapping tool in experiment two and three revealed that the non-technical users did not find the act of mapping to be hard. Furthermore, many suggested it was quite easy and straightforward to complete mapping tasks with the tag-based mapping tool. The evaluation results of both experiments also showed the time taken to complete a mapping task was low, see section 9.1, the number of mapping tasks completed was high, see section 9.3.2, and the mapping precision was moderately high, see section 9.2. Several participants in experiment three did not answer a mapping tasks but that was due to the lack of mapping prompts being displayed to these users as they did not use the SportsFlows application, see section 8.5. Overall, in the author's opinion, these results indicated that the act of mapping can become simple enough for ordinary users to achieve with the tag-based mapping tool.

## 9.5 User Satisfaction

The user satisfaction was measured through questionnaires and interview given to users. In experiment one, the interviews with users revealed that the non-technical users found COMA++ hard to use while they found the NL prototype mapping tool to straightforward and much easier to use. In experiment two, the response from questions asked in the interviews revealed that the non-technical users found mapping very straightforward and not difficult at all. Finally in experiment three, a SUS questionnaire was given to assess the user satisfaction and the non-technical users had an average score of 64 which was relative to the results of other mapping case studies and indicates that the users found the tag-based mapping tool appropriate for answering mapping tasks, see section 8.5.5.1. Overall, in the author's opinion, the results indicate that users found the tag-based mapping tool quite appropriate for answering mapping tasks.

## 9.6 Summary

The evaluation of experiment one indicated that ordinary users could potentially answer mapping tasks if the mapping tool is designed specifically for them, NL prototype mapping tool, but the current state of the art mapping tools are too confusing for them to use properly. The results of the experiment two and three revealed that using the tag-based mapping tool the act of mapping can become efficient, effective, unintrusive, engaging, simple and satisfactory for ordinary users. These results are very surprising and emphasis that ordinary users can participate in the mapping process. In particular, the results from experiment two and three revealed that ordinary users continued to answer mapping task over time which was not expected and highlights that ordinary users may be inclined to continue to contribute to the mapping process over time. Furthermore the level of participation of ordinary users was relative to the participation of ontologically aware users which was greatly surprising. The user evaluation carried out in this thesis is unique in relation to the user studies carried out by the state of the art and it is hoped that due to the positive results that it will encourage other researchers to develop ordinary user support in their mapping tools. However, although the mapping precision results for ordinary users was moderately high, the precision was still significantly lower than ontologically aware users in two of the experiments. In addition although the engagement results revealed that although ordinary users were engaging in the mapping process they were also not seeing any benefits to mapping. In the author's opinion, there is a need to clearly show the benefits of mappings to give incentives for users to continue to participate in the mapping process.

In the next chapter, the contributions made in this thesis are discussed, suggestions for future work are detailed and finally some conclusions are given.

# 10 Conclusions

In this chapter, the research objectives of the thesis are presented with discussion on the degree to which they have been achieved. This discussion is followed by the contributions achieved. The primary contribution of the thesis was devising and implementing an approach which allows ordinary users to participate in the mapping process within their own work environment. Following this a number of future directions for this research are presented. Finally some final remarks are given.

## 10.1 Research Objectives

In chapter one, five research objectives were derived from the research question, see section 1.3. The extent to which the research objectives were accomplished is discussed in the following sections.

### 10.1.1 Ordinary users mapping involvement

The first research objective **RO1** stated: *Determine whether ordinary users can actively participate in the construction of semantic mappings to a similar standard of knowledge engineers in areas of interest to the ordinary users*. Three experiments were carried out in this thesis and the evaluation results were used to examine whether ordinary users can complete mapping tasks to the standard of knowledge engineers. The first experiment examined the mapping performance achieved by the users with a current state of the art mapping tool and the initial prototype of the tag-based mapping tool. The second and third experiments examined the mapping performance achieved by the users with the tag-based mapping tool. The evaluation results revealed that the only difference in mapping performance was the mapping precision achieved for ordinary user which was significantly worse in experiment one and three than the precision achieved by knowledge engineers, see section 9.2. The mapping precision result in experiment one is understandable as the users were using a state of the art mapping tool and only a prototype of the tag-based mapping tool. However in experiment three the users were using the final version of the tag-based mapping tool. The precision result achieved in experiment three by the knowledge engineers at 93% was extremely high in comparison to the other experiment results. In contrast, the precision result achieved by the ordinary users at 64% was more comparative to the other experiment results. Overall, the results in this thesis indicate that although ordinary users can participate in the construction of mappings, the mapping precision achieved by them was not to the same standard as knowledge engineers. This indicates a need to supplement and improve the mappings undertaken by ordinary users post-interaction, especially in areas where high precision is needed. This assistance could potentially come from a variety of methods including: (1) ordinary users managing and maintaining their mappings, (2) assistance from collaborative sharing with other users, (3) assistance from an administrator with knowledge engineer skills and (4) mapping evolvement, see section 10.3.1.

### 10.1.2 Mapping requirements for ordinary users

The second research objective **RO2** stated: *Determine the parts of the semantic mapping process which are difficult for ordinary users and identify techniques to reduce the difficulty of these parts*. In section 2.8, several requirements were specified after a review of the state of the art. These requirements have been updated throughout the thesis based on experimental analysis and the ongoing research. Figure 10-1 displays the updated list.

| **Eliminate** | **Reduce** |
|---|---|
| Any Matching API Configuration by User | Complexity of Mapping Process |
| Visual Emphasis on Global View | Tasks per Mapping Session |
| User-Specific Tasks not Related to their Interests | |
| **Raise** | **Create** |
| Visual Emphasis on Local Neighbourhood | Iterative Mapping Process |
| Visualisation of Information Gained Through Mapping | Familiar & Rich Mapping Interaction |
| | Easy to Understand Interface |
| | Environments for Mapping Sessions |
| | System Mapping Evolvement |
| | Evaluation Methods |

***Figure 10-1:*** *Mapping Requirements for Ordinary Users*

These are the requirements that need to be addressed to allow ordinary users to participate in the mapping process. The techniques used in this thesis to address these requirements are described in the implementation chapters for the tag-based mapping tool, see section 5.1.1 and section 7.1.1. It is possible there can be other methods/techniques that can solve these requirements.

### 10.1.3 Incremental mapping process

The third research objective **RO3** stated: *Develop a mapping process that will aid ordinary users in the act of mapping over time, called the incremental mapping process*. The incremental mapping process was presented in section 4.3 and was designed to: (1) reduce the difficulty of the mapping process by eliminating several user-specific mapping tasks with high cognitive load, (2) allows the mapping process to occur intermittently when required within the computing environment of the users, and (3) shortens the workload for users by both decomposing the process to occur over multiple sessions and removing unnecessary user-specific mapping tasks. An implementation of the incremental mapping process for ordinary users was presented in section 5.1 which reduced the user-specific mapping tasks to just mapping candidate correspondences. This implementation was used in the design of the tag-based mapping tool and helped to allow the mapping process to occur within the browsing environment of ordinary users.

### 10.1.4 Tag-based mapping tool

The fourth research objective **RO4** stated: *Develop a tool that is based on the incremental mapping process from **RO3** and the techniques identified in **RO2***. The tag-based mapping tool was implemented using the incremental mapping process and the techniques identified for the requirements, see section 5.1 and section 7.1. The tool allows the mapping process to take place in the users own browsing environment over time. The design of the tag-based mapping tool allows it to be used by applications in different mapping scenarios, as both the personal ontology and correspondences are stored locally on the user's computer in a database, see Figure 7-2. A personal information delivery tool was implemented using the tag-based mapping tool in order to support evaluation of the approach. The tool called SportsFlows allowed users to personalise sporting information from RSS feeds, see section 5.2 and section 7.2. SportsFlows was used by users in the experiments to evaluate the tag-based mapping tool and the proposed approach.

### 10.1.5 Evaluation

The final research objective **RO5** stated: *Evaluate the tool with ordinary users over a long term period (at least 2 weeks and up to 5 weeks)*. Experiments two and three were designed to evaluate the mapping performance of ordinary users with the tag-based mapping tool. There has been a distinct lack of user studies carried out on semantic mapping tools, see section 2.7. In addition, the following list details the differences between the user studies carried out on the tag-based mapping tool and the user studies carried out by the state of the art,

- *Field study*: Although there have been other field mapping studies none, as of writing, have evaluated the mapping performance of users using mapping tools within their own work environment. The only published field studies have been a study which evaluated the candidate correspondences generated by PROMPT [Noy 2002] and a study investigating what mapping tools users use and why they use them via a questionnaire and interviews [Falconer 2009];
- *Long-term study*: The previous state of the art user mapping studies have been carried out in a one-shot session. Experiment two and three evaluated the tag-based mapping tool over multiple weeks. This allowed the evaluation to assess the impact of doing mapping tasks over time on users;
- *Real world study*: The majority of the previous mapping studies have not shown the outcome of mapping to the users. Furthermore, the mapping tasks may not have been related to the users own personal interest model. In experiment two and three the impact of mapping was shown through the personalisation of sporting information. In addition, the mapping tasks asked were also related to the users own interests, i.e. personal interest model;
- *Group study*: Only one other user mapping study has evaluated the difference in mappings performances between users with different technical experience [Lambrix 2003], see section 2.7.2. However the study focussed primarily on the performance of the mapping tools rather than the user groups. Experiment two and three measured the mapping performance between different technical user groups.

The previous user mapping studies carried out by the state of the art did not have the above characteristics. These studies used precision and recall as the methods for evaluation, see section 2.7.6. In addition, the usability of the mapping tools was evaluated through interviews and questionnaires. However, none of these studies assessed the accessibility and convenience of using mapping tools within the users own work environment.

Evaluation areas need to be devised to evaluate the usability, convenience and accessibility of mapping tools with users. Six different evaluation areas were suggested by the author, see section 1.5.2, and are: *efficiency*, *effectiveness*, *unintrusive*, *engaging*, *simplicity* and *user satisfaction*. Both methods and metrics were defined for these evaluation areas in experiment two and three, see table 6-1 and table 8-1. Experiment two was designed to allow users to be interviewed post-study to gain further insight into the evaluations results. However in experiment three the users could not be interviewed post-study which lead to the metrics and methods being modified. In addition, the metrics and methods were also modified in experiment three to take into account the longer experimentation time. Thus, a couple of different ways of carrying out mapping studies with these metrics and methods have been detailed.

## 10.2 Contributions

In the following sections the contributions of this thesis are discussed.

### 10.2.1 Major Contribution: Mapping Approach for Ordinary Users

The approaches taken in the state of the art of mapping have been to design the tools and processes for knowledge engineers or users who have received substantial training. These approaches are not suitable for ordinary users. However, ordinary users participation is needed in the mapping process due to the proliferation of semantic models on the web, see section 1.1. In this thesis, a mapping approach was devised which allowed ordinary users to participate in the mapping process over time within their browsing environment. The steps used to devise this approach were as follows: (1) analyse the state of the art in mapping and identify issues which need to be addressed, (2) conduct a behaviour study with ordinary users using mapping tools which will aid in highlighting the problematic areas, (3) from the previous steps devise a list of mapping requirements needed for ordinary users using an eliminate-reduce-raise-create (ERRC) grid and identify techniques to solve these requirements, (4) design an incremental mapping process which will reduce the number of user-specific mapping tasks and in addition allow the process to occur over time, and (5) implement the approach through a tag-based mapping tool developed using the incremental mapping process and the techniques identified to address the requirements. The tag-based mapping tool was evaluated and the results indicated that the approach taken is suitable for ordinary users. These results highlighted to the state of the art that it is indeed possible that ordinary users can participate in the mapping process and it is hoped that this will encourage other researchers when designing their mapping tools to include support for ordinary users. The mapping requirements devised can be used by other researchers when the designing their mapping tools. Additionally the techniques used to address these requirements can also be used by other researches. Furthermore the steps used to devise the approach can be used by other researchers when developing their mapping tools. The work in this thesis has been contributed to the community through a series of academic publications:

> Conroy C., Brennan R., O'Sullivan D., Lewis D. (2011) "Semantic Mapping for Ordinary Users", submitted to Journal of Web Semantic

Conroy C., Brennan R., O'Sullivan D., Lewis D. (2009) "User Evaluation Study of a Tagging Approach to Semantic Mapping", In Proceedings of the 5th European Semantic Web Conference (ESWC 09)

Lanzenberger M, Sampson J, Kargl H, Wimmer M, Conroy C, O'Sullivan D, Lewis D, Brennan R, Gargantilla J.A.R, Gómez-Pérez A, Fürst F, Trichet F, Euzenat J, Polleres A, Scharffe F, Kotis K, (2008) "Making Ontologies Talk: Knowledge Interoperability in the Semantic Web", IEEE Intelligent Systems 23(6): 72-85

Conroy C. (2008) "Towards Semantic Mapping for Casual Web Users", In proceeding of the International Semantic Web Conference (ISWC 2008)

Conroy C., (2008) "Towards Ontology Mapping for Ordinary People " at the Phd Symposium at 5th European Semantic web Conference (ESWC 2008) (poster)

Conroy C., O'Sullivan D., Lewis D., (2008) "Ontology Mapping through Tagging", at the International Workshop on Ontology Alignment and Visualisation (OnAV 08), CISIS, 2008

Conroy C., O'Sullivan D., Lewis D., (2007) "A tagging approach to ontology mapping", at the Ontology Matching Workshop (OM 2007) at the 6th International Semantic Web Conference (ISWC 2007)

### 10.2.2 Minor Contribution: Incremental mapping process

The current state of the art mapping processes lack the ability of allowing the act of mapping to occur incrementally over time, see section 2.3. Instead they rely on "one-shot" session for users to develop mappings which tend to be time-consuming. In addition they lack the ability to intertwine the mapping process with the users own work environment over time. Furthermore they do not allow for changes to be implemented on the fly at runtime when required. The incremental mapping process is designed to allow the process to occur over time by prompting mapping tasks when required. This process allows the act of mapping to be intertwined into the users computing environment. The incremental mapping process can be used by other researchers to help guide the design of their own mapping tools. Additionally the process also allows the number of user-specific mapping tasks to be reduced which allows mapping tools to be tailored towards different types of users. Furthermore the process identifies how the mapping process can be personalised for users, see section 4.3. Indeed there can be no 'one-size fits all' mapping approach as everybody has different skills and tastes which need to be taken into account and used to tailor the mapping process for the individual. The incremental mapping process has been contributed to the community through a series of academic publications:

Conroy C., Brennan R., O'Sullivan D., Lewis D. (2009) "User Evaluation Study of a Tagging Approach to Semantic Mapping", In Proceedings of the 5th European Semantic Web Conference (ESWC 09)

Lanzenberger M, Sampson J, Kargl H, Wimmer M, Conroy C, O'Sullivan D, Lewis D, Brennan R, Gargantilla J.A.R, Gómez-Pérez A, Fürst F, Trichet F, Euzenat J, Polleres A, Scharffe F, Kotis K, (2008) "Making Ontologies Talk: Knowledge Interoperability in the Semantic Web", IEEE Intelligent Systems 23(6): 72-85

Conroy C. (2008) "Towards Semantic Mapping for Casual Web Users", In proceeding of the International Semantic Web Conference (ISWC 2008)

Conroy C., (2008) "Towards Ontology Mapping for Ordinary People " at the Phd Symposium at 5th European Semantic web Conference (ESWC 2008) (poster)

Conroy C., O'Sullivan D., Lewis D., (2008) "Ontology Mapping through Tagging", at the International Workshop on Ontology Alignment and Visualisation (OnAV 08), CISIS, 2008

### 10.2.3 Minor Contribution: Evaluation Approach for Mapping Tools

There has been no real world experimentation on users using mapping tools over time within their own work environment. This type of experimentation is necessary if mapping tools are to gain wide scale adoption as it is important to gain an understanding of the user behaviour with mapping tools in a real world and non laboratory situation. The approach taken in this thesis allowed the tag-based mapping tool to be analysed with ordinary users over time in their computing environment and had (1) several author defined evaluation areas with corresponding methods and metrics to evaluate the users performance with the tag-based mapping tool over time and (2) an RSS sports personalisation application SportsFlows to display the results of mapping which replicated a real world case study mapping tools would encounter.

There is no standard defined benchmark on how to perform user evaluation of mapping tools, see section 2.7.6. Furthermore there has been no evaluation undertaken on mapping tools over time and thus no methods or metrics have been defined for how to evaluate mapping tools over time. In this thesis, six evaluation areas with corresponding methods and metrics are suggested and used to evaluate the tag-based mapping tool over time, see section 10.1.5. These evaluation areas are not specific to ordinary users. The evaluation areas suggested in this thesis, with corresponding methods and metrics, can be used by other researchers to evaluate the performance of users using their mapping tools.

There have been no real world user case studies on mapping tools by the state of the art in mapping. Instead the experiments have been conducted in laboratory setting where the users cannot view the full effects of their mapping decisions. There needs to be real world user case studies on mapping tools to analyse the users usage of the tool. In this thesis, an RSS sports personalisation application SportsFlows was used to replicate a real world case study. The evaluation results reveal that despite users not finding SportsFlows to be much use they were still participating in the mapping process with the tag-based mapping tool, see section 9.3.2. It is hoped that these results will encourage other researchers to conduct real world user case studies on their mapping tools. The design of SportsFlows was generic and can potentially be modified to be used in other domains such as financial domain. This allows other researchers to develop their own applications using this design which can then be used to evaluate their mapping tools. The evaluation approach has been contributed to the community in the following academic publication:

> Conroy C., Brennan R., O'Sullivan D., Lewis D. (2009) "User Evaluation Study of a Tagging Approach to Semantic Mapping", In Proceedings of the 5th European Semantic Web Conference (ESWC 09)

## 10.3 Future Work

While it has been shown that ordinary users can interact in the mapping process within their own browsing environment there is still a lot of work to be done in this area. In the following sections, a number of potential research directions this work could follow are discussed.

### 10.3.1 Mapping evolution

After the mapping has been generated it will also need to be evolved over time based on the following issues: (1) the ontologies being altered leading to correspondences being invalidated, (2) the need of the mapping being changed, and (3) the mapping generated having too low a precision. In the final version of the tag-based mapping tool a mapping evolution component was developed, see section 7.1.5.4. The mapping evolution component was meant to improve the mapping precision but instead decreased the precision, see section 8.5.2.3. This was due to the SportsFlows application use, i.e. how often the user looked at any information given by

SportsFlows, which was very low and negatively affected how the correspondences were evolved. The overall application use of the mapping will need to be taken into account when evolving correspondences based on their usage. In addition, correspondences can affect other correspondences. For example if there were two correspondences 'Male->Man' and 'Male->Woman' between the same ontologies and 'Man' and 'Woman' are disjoint concepts then if 'Male->Man' relation was equivalent then the 'Male->Woman' relation would be disjoint. This allows a hierarchical structure to be constructed for the correspondences which can be used to propagate the modifications made to correspondences to other correspondences through the structure. Furthermore the hierarchical structure can allow for correspondences to spawn several child correspondences or several correspondences to derive a parent correspondence. For example if all the subclasses of concept A have a correspondence to concept B with the equivalence relation then a 'parent' correspondence can be derived between concept A and concept B with the equivalence relation. Alternatively if there is a correspondence between concept A and concept B with an equivalent sometimes relation then several 'child' correspondences can be derived between the subclasses of concept A and concept B with the same equivalent sometimes relation, indeed these child correspondences could eventually change relations to identify which part of the structure of concept A is equivalent to concept B.

### 10.3.2 Tag processing

The tag-based mapping tool allowed users to enter the relation of the correspondence in their own words using tags. This led to ambiguity in the system as some of the tags could not be identified. These tags were called user-defined relations. In the final version of the tag-based mapping tool a mapping evolution component was implemented to reason the mapping category of the correspondences mapped with user-defined relations, see section 7.1.5.4. Although not implemented, this component can potentially be used to reason the mapping category of user-defined relations. However the component would not be able to identify the concept relation category of the user-defined relation. In addition, in the evaluation of the experiments the user-defined relations entered by the user were analysed by the author and the results revealed that the user-defined relations were both useful and accurate at defining the relationship of the correspondence. Using natural language based techniques it might be possible to make use of these user-defined relations to define relationships between the concepts. An example of an approach would be to extend the system SCARLET [Sabou 2008], which currently returns the relationship between two concepts by searching multiple ontologies online, to included searching for relations which would return the relationship category of the relation by searching multiple ontologies online for the use of the relation in the ontologies. In addition, the user-defined relations of correspondences can potentially be used to devise new correspondences. For example if the correspondence 'Golf' to 'Harrington' was tagged with the user-defined relation *has player*, then if 'Golf' had a property relation has player to 'Golfer' then a correspondence can be derived between 'Golfer' and 'Harrington'. Furthermore, the user-defined relations may be used as a way of developing complex correspondences between ontologies, for example many-to-many correspondences. For example if the correspondences 'first name' to 'person's name' and 'surname' to 'person's name' are both mapped with the user-defined relation *is part of* then with some reasoning or additional human input it may be possible to devise the mapping 'first name' & 'last name' is equivalent to 'person's name'.

### 10.3.3 User-Specific Mapping Tasks

If the mapping process is to see wide scale adoption it will be necessary to identify what mapping tasks ordinary users can accomplish and to what standard. The research in this thesis investigated how and whether ordinary users can map correspondences. The level of ordinary user involvement in the mapping process was not addressed. However several of the mapping tasks were identified as being too complex for ordinary users and instead were made into mapping system specific tasks, see section 4.2. Further investigation is needed to identify which mapping tasks ordinary users can accomplish and to what standard. In addition the more complex mapping tasks which ordinary users would be unable to accomplish could potentially have user input through other mapping tasks which have not been defined yet. For example the users could be tasked with entering which concepts in their ontology are the most important which can be then be used to order the candidate correspondences. Of particular interest is how ordinary users can manage and maintain the correspondences they generated. In the author's opinion this is a critical mapping task that needs to have some level of user involvement as it is important to give the users control to change their minds over the mapping decision they make. Furthermore this mapping task should allow the users to see the feedback of their mapping actions more clearly which was an issue in the user studies carried out in this thesis.

### 10.3.4 Personalisation of Mapping Process

There should be no 'one-size fits all' mapping approach. Each person has different skills or tastes and they must be taken into account to tailor the mapping process for the individual. The incremental mapping process decomposed the mapping process into a series of steps for how users accomplish user-specific mapping tasks. It was also identified how each step could be personalised for the individual, see section 4.3. However the implementation of the tag-based mapping tool only allowed the suggested relations list to be personalised by appending any user-defined relation the user enters, see section 5.1.4.4. Investigation is needed into how each step can be personalised for users through either passive or active feedback. Furthermore, a full user-model could potentially be used to personalise the mapping process. Additionally as the users skills increase so should their control over the process, i.e. the number of different mapping tasks available to the user. As Kathy Sierra noted on her famous creating passionate users blog about the level of control an application or tool should give to the user *as user capability (knowledge, skill, expertise) increase, so should control*[38].

### 10.3.5 Use-Case for Mapping

There has been little experimentation done on mapping tools especially in real world situations. In particular, there has been a lack of investigation on identifying the areas where mappings are needed for ordinary users and indeed whether the current solutions provided are good enough, i.e. are mappings needed at all. For example, although if the user's personal ontology model is mapped to a search engine ontology model it may provided increased precision in the search results from the engine, the results provided without the mapping may be good enough for the user to not warrant the hassle of constructing a mapping, i.e. the current solution is good enough. In addition, it will also need to be investigated whether partial mappings can suffice for each area as the tag-based mapping tool only provides a partial mapping which is continually updated over time by the user and tool. The experiments carried out in this thesis were on the personalisation of sporting information. However, no investigation was done by the author on whether mappings are needed in this area. Indeed the benefits of personalisation were not seen by users which indicated that the partial mappings generated were unnecessary. However there was distinct lack of participation with the SportsFlows tool from users with feedback revealing that the users could not identify the benefits given from the partial mappings. Thus there can be no conclusion given on whether mappings are necessary in this area. Further investigation is needed on where mapping could potentially be needed and experimentation to evaluate whether it is needed. Some potential areas are stock & shares and the financial area. The requirements for mapping for these areas would be that it provides a solution that is both fast and reliable. These requirements were not needed when evaluating the tag-based mapping tool in the personalisation of sporting information area, so it remains to be seen if the approach taken by the tool can provided mappings that are both fast and reliable for the areas. Furthermore since the tag-based mapping only provides a partial mapping it will also need to be investigated whether partial mappings can provide a solution in those areas.

[38] http://headrush.typepad.com/creating_passionate_users/2007/02/how_much_contro.html

## 10.4 Final Remarks

The vision of the semantic web is to bring structure and formal semantics to information on the web so that the information can be exchanged and shared in machine-processable formats. However, similar information can be represented with a range of different semantic models. This leads to a need for mappings to be developed between semantic models to resolve the model interoperability issue. Traditionally, the mapping process is carried out offline by knowledge engineers with the assistance of domain experts if needed. However, if the mapping process is to scale to the web, for example where every ordinary user will have a personal semantic model representing their interests, the mapping process will need to be performed at runtime with the involvement of ordinary users. Since current mapping tools target knowledge engineers, both the usability and accessibility of these tools will need to be significantly improved if ordinary users are to be able to interact within the mapping process. The work in this thesis addresses how an approach can be developed to allow ordinary users to become involved in the mapping process within their own work environment over time.

In the mapping research community, there have only been limited usability experiments and user-based evaluations. In addition, the few experiments that have been carried out have focussed on improving the usability of tools for knowledge engineers. Thus there has been no approach or requirements being identified on how ordinary users can participate in the mapping process. Furthermore both the accessibility of mapping tools and how the mapping process and tool are to be incorporated into the work environment of users have not been addressed.

An initial experiment was carried out to isolate the problem areas of the act of mapping for ordinary users using both standard tools and a novel natural language question and answer-based tool. The evaluation results, together with a literature review of the state of the art, identified several requirements and techniques needed to allow the act of mapping to become more accessible, convenient and usable. An incremental mapping process was developed to allow the mapping process to become accessible and occur within the work environment of users. Furthermore, a tag-based mapping tool was developed based on the incremental mapping process and the techniques identified. The evaluation of the tool indicated that the mapping process can be incorporated into the ordinary user's work environment. Several modifications were made to the tag-based mapping tool to refine the approach. The final experiment evaluated the tag-based mapping tool with ordinary users over a long term period within their browsing environment. The mapping results from this evaluation demonstrated that ordinary users were using the tag-based mapping tool within their browsing environment during the course of the experiment.

The overall motivation of this thesis is that ordinary users need to be involved in the mapping process. Through three user studies it was shown that ordinary users were able to map correspondences using the tag-based mapping tool. In the author's opinion, by helping ordinary users to take part in the mapping process it will allow the mapping process to scale to the web. Overall the results from the experiments revealed that ordinary users using the tag-based mapping tool can take part in the mapping process within their own work environment over time.

# References

[Aberdeen 2003] "University Ontology", (2003) University of Aberdeen, Available at: http://www.csd.abdn.ac.uk/~cmckenzi/playpen/rdf/akt_ontology_LITE.owl, [Accessed: September 2006].

[Angeletou 2008] Angeletou S., Sabou M., Motta E., (2008) "Semantically enriching folksonomies with FLOR", Workshop Collective Intelligence & the Semantic Web at the 5th European Semantic Web Conference (ESWC 2008)

[Anthony 2008] Anthony S.D., Johnson M.W., Sinfield J.V., Altman E.J (2008) "The Innovator's Guide to Growth", Harvard Business Press, Boston, MA

[Auer 2007] Auer S., Bizer C., Kobilarov G., Lehmann J., Cyganiak R., Ives, Z., (2007) "DBpedia: A Nucleus for a Web of Open Data", In Proceedings of 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference, (ISWC 2007 & ASWC 2007), Lecture Notes in Computer Science, Volume 4825, Springer, Heidelberg, pp. 722 – 735.

[Aumuller 2005] Aumuller D., Do H., Massmann S., Rahm E., (2005) "Schema and Ontology matching with COMA++", Proceedings of the ACM SIGMOD conference on management of data (SIGMOD 2005),ACM Press, pp. 906 – 908.

[Bauer 2005] Bauer C., King G., (2005) "Hibernate in Action", Manning, ISBN 1932394-15-X

[Begin 2007] Begin C., Gooding B., Meadors L., (2007) "iBATIS in Action", Manning, 1st Edition, ISBN 1932394825

[Bernstein 2006a] Bernstein A. and Kaufmann E., (2006) "GINO – A Guided Input Natural Language Ontology Editor", In Proceedings of the 5th International Semantic Web Conference (ISWC 2006), Lecture Notes in Computer Science, Volume 4273, Springer, Heidelberg, pp. 144 – 157.

[Bernstein 2006b] Bernstein P.A., Melnik S., Curchill J.E., (2006) "Incremental Schema Matching", In Proceedings of the 32nd International Conference on Very Large Databases (VLDB), ACM Press, pp. 1167-1170.

[Bernstein 2007] Bernstein P.A. and Melnik S., (2007) "Model management 2.0: Manipulating richer mappings", In Proceedings of the 2007 ACM SIGMOD international conference on Management of data, ACM Press, pp. 1–12

[Bizer 2009] Bizer C., HeathT., Berners-Lee T., (2009) 'Linked Data – The Story So Far', Special Issue on Linked Data, International Journal on Semantic Web and Information Systems (IJSWIS), 5(3), pp. 1-22.

[BizTalk 2005] Microsoft Visual Studio, Version 8.0, (2005) Available at http://msdn2.microsoft.com/en-us/ie/bb188238.aspx, [Accessed September 2008]

[Bogdan 2008] Bogdan A., Wang-Chiew T., Yannis V., (2008) "STBenchmark: towards a benchmark for mapping systems", In Proceedings of the VLDB Endowment, 1(1), pp. 230 – 244.

[Brooke 1996] Brooke J., (1996) "Usability evaluation in industry", In SUS: a 'quick and dirty' usability scale, Taylor & Francis, pp. 184–194

[Bruijn 2005] deBruijn J., Martin-Recuerda F., Ehrig M., Polleres A., Predoiu L., (2005) 'Ontology Mediation Management', IST SEKT project deliverable, 4.4.1, February 2005.

[Cavaness 2006] Cavaness C., (2006) "Quartz Job Scheduling Framework: Building Open Source Enterprise Applications", Prentice Hall, 1st edition

[Chan Kim 2005] Chan Kim W., Mauborgne R., (2005) 'Blue Ocean Strategy: How to Create Uncontested Market Space and Make Competition Irrelevant', Cambridge, MA: Harvard Business School Press

[Chopra 2004] V. Chopra, A. Bakore, J. Eaves, B. Galbraith, S. Li, C. Wiggers, (2004) "Professional Apache Tomcat 5", published by Wrox, ISBN 0764559028.

[Christensen 1998] Christensen CM, (1998) "The innovator's dilemma", Boston: Harvard Business School Press, Boston, MA

[Conroy 2007] Conroy C., O'Sullivan D., Lewis D., (2007) "A tagging approach to ontology mapping", at the Ontology Matching Workshop (OM 2007) at the 6th International Semantic Web Conference (ISWC 2007)

[Conroy 2008a] Conroy C., O'Sullivan D., Lewis D., (2008) "Ontology Mapping through Tagging", The International Workshop on Ontology Alignment and Visualisation (OnAV 08), In Proceedings of the 2008 International Conference on Complex, Intelligent and Software Intensives Systems (CISIS), ACM Press, pp. 886 – 891.

[Conroy 2008b] Conroy C., (2008) "Towards Ontology Mapping for Ordinary People " at the Phd Symposium at 5th European Semantic web Conference (ESWC 2008) (poster)

[Conroy 2008c] Conroy C. (2008) "Towards Semantic Mapping for Casual Web Users", In proceeding of the 7th International Semantic Web Conference (ISWC 2008), Lecture Notes in Computer Science, Volume 5554, Springer, Heidelberg, pp. 623 – 637

[Conroy 2009] Conroy C., Brennan R., O'Sullivan D., Lewis D. (2009) "User Evaluation Study of a Tagging Approach to Semantic Mapping", In proceedings of the 5th European Semantic Web Conference (ESWC 2009), Lecture Notes in Computer Science, Volume 5318, Springer, Heidelberg, pp. 907 – 913.

[Creswell 2003] Creswell J.W., (2003) "Research Design: Qualitative, Quantitative, and Mixed Methods Approaches", SAGE Publications, 2nd edition

[Crichton 2000] Crichton N., (2000) "Mann-Whitney test", Journal of Clinical Nursing, 9(4), pp. 583.

[Crockford 2009] Crockford D., (2009). "Introducing JSON", Available at http://json.org. [Accessed September, 2011].

[Davis 2008] Davis B., Iqbal A.A., Funk A., Tablan V., Bontcheva K., Cunningham H., Handschuh S., (2008) "Roundtrip ontology authoring", In Proceedings of the 7th International Semantic Web Conference (ISWC

2008), Lecture Notes in Computer Science, Volume 5318 Springer, Heidelberg, pp. 50 – 65.

[Dimitrova 2008] Dimitrova V., Denaux R., Hart G., Dolbear C., Holt I., Cohn A.G., (2008) "Involving Domain Experts in Authoring OWL Ontologies", In Proceedings of the 7th International Semantic Web Conference (ISWC 2008), Lecture Notes in Computer Science, Volume 5318 Springer, Heidelberg, pp. 1 – 16.

[Do 2002] Do H.H., Rahm E., (2002) 'COMA: A system for flexible combination of schema matching approaches', In Proceedings for the 28th international conference on Very Large Data Bases (VLDB 2002), ACM Press, pp. 610 – 621.

[Doan 2003] Doan A., Domingos P., Halevy A., (2003) "Learning to match the schemas of data sources: A multistrategy approach", Machine Learning, 50(3), pp. 279 – 301.

[Euzenat 2004a] Euzenat J., (2004) "State of the art on ontology", deliverable d2.2, Technical report, Knowledge web, Available at http://people.aifb.kit.edu/meh/publications/euzenat04state.pdf [Accessed Feburary 2008]

[Euzenat 2004b] Euzenat J., Loup D., Touzani M., Valtchev P., (2004) "Ontology alignment with ola", In Proceedings of the 3rd EON Workshop, 3rd International Semantic Web Conference (ISWC 2004)

[Euzenat 2006] Euzenat J., (2006) An API for ontology alignment (version 2.1), Available at http://gforge.inria.fr/docman/view.php/ii7/251/align.pdf, [Accessed February 2006]

[Falconer 2007] Falconer et al. (2007) 'A Cognitive Support Framework for Ontology Mapping' In Proceedings of 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference, (ISWC 2007 & ASWC 2007), Lecture Notes in Computer Science, Volume 4825, Springer, Heidelberg, pp.114 – 127.

[Falconer 2009] Falconer S., (2009) "Cognitive support for semi-automatic ontology mapping", PhD Thesis, University of Victoria.

[Feldt 2007] Feldt K., (2007) “Programming Firefox: Building Rich Internet Applications with XUL”, O'Reilly Media Inc.

[Fisher 1925] Fisher RA., (1925) “Statistical Methods for Research Workers” Edinburgh: Oliver & Boyd, first edition

[Gal 2005] Gal A., Anaby-Tavor A., Trombetta A., Montesi, D., (2005) ‘A framework for modelling and evaluating automatic semantic reconciliation’, In VLDB Journal, 14(1), pp. 50-67

[Gal 2006] Gal A., (2006) ‘Why is Schema Matching Tough and What Can We Do About It?’, In SIGMOD Record, 35(4), pp. 2 – 5.

[Gauch 2003] Gauch S., Chaffee J., Pretschner A., (2003) “Ontology-based personalized search and browsing”, Web Intelligence and Agent Systems, 1(3-4), pp. 219-234.

[Giunchiglia 2004] Giunchiglia F., Shvaiko P., Yatskevich M., (2004) “S-Match: an algorithm and an implementation of semantic matching”, In proceedings of the European Semantic Web Conference (ESWC 2004), Lecture Notes in Computer Science, Springer, Heidelberg, pp. 61-75

[Hampton 1979] Hampton J.A., (1979) “Polymorphous concepts in semantic memory”, Journal of Verbal Learning and Verbal Behaviour, 18, pp. 441–481

[Heflin 2000] Heflin J., Hendler J., (2000) “Semantic Interoperability on the Web”, In Proceedings of Extreme Markup Languages 2000, Graphic Communications Association, 2000. pp. 111-120.

[Hertling 1999] Hertling P., (1999) ‘An effective Riemann mapping theorem’, In Journal of Theoretical Computer Science – Special issue on computability and complexity in analysis, 219(1-2), pp. 225 – 265.

[Huhns 2002] Huhns M.N., (2002) “Software Agents: The Future of Web Services”, Agent Technologies, Infrastructures, Tools, and Applications for E-Services, Springer, Heidelberg, pp. 1 – 18.

[Hull 1997] Hull R., (1997) ‘Managing semantic heterogeneity in databases: A theoretical perspective’, In Proceedings of the ACM SIGACT-

SIGMOD-SIGART Symposium on Principles of Database Systems (PODS), ACM Press, pp. 51 – 61.

[Isaac 2007] Isaac A., van der Meij L., Schlobach S., Wang S., (2007) "An empirical study of instance-based ontology matching", Proceedings of the 6th International Semantic Web Conference (ISWC 2007), Lecture Notes in Computer Science, Volume 4825, Spinger, Heidelberg, pp. 253-266.

[Iswc 2007] "ISWC Ontology", The International Semantic Web Ontology, (2007) Available at: http://annotation.semanticweb.org/iswc/iswc.owl, [Accessed January 2007].

[Jameson 2006] Jameson A., (2006) "Usability and the Semantic Web", In proceedings of the 3rd European Semantic Web Conference (ESWC 2006), Springer Lecture Notes in Computer Science, Volume 2011, Spinger, Heidelberg, pp. 81 – 101.

[Kalfoglou 2003] Kalfoglou, Y., Schorlemmer, M., (2003) "Ontology mapping: the state of the art", The Knowledge Engineering Review, 18 (1). pp. 1-31.

[Kanzaki 2007] "Music Ontology", Kanzaki Music Ontology, (2007) Available at: http://www.kanzaki.com/ns/music [Accessed January 2007]

[Kaufmann 2007] Kaufmann E. and Bernstein A., (2007) "How Useful are Natural Language Interfaces for Casual End-users?", In Proceedings of the 6th International Semantic Web Conference (ISWC 2007), Lecture Notes in Computer Science, Volume 4825, Spinger, Heidelberg, pp. 281 – 294.

[Klein 2001] Klein M., (2001) "Combining and relating ontologies: an analysis of problems and solutions", In Proceedings of IJCAI-2001 Workshop on Ontologies and Information Sharing, pp. 53-62.

[Konstantinou 2008] Konstantinou N., Spanos D.M., Nikolas M., (2008) 'Ontology and database mapping: a survey of current implementations and future directions', In Journal of web Engineering, 7(1), pp. 001-024

[Lambrix 2003] Lambrix P. and Edberg A. (2003) “Evaluation of ontology merging tools in bioinformatics”, In Proceedings Pacific Symposium on Biocomputing, pp. 589–600.

[Lanzenberger 2006] Lanzenberger M. and Sampson J., (2006) “Alviz - a tool for visual ontology alignment”, In IV ’06: Proceedings of the conference on Information Visualization, IEEE Computer Society, pp. 430–440.

[Lanzenberger 2008] Lanzenberger M, Sampson J, Kargl H, Wimmer M, Conroy C, O'Sullivan D, Lewis D, Brennan R, Gargantilla J.A.R, Gómez-Pérez A, Fürst F, Trichet F, Euzenat J, Polleres A, Scharffe F, Kotis K, (2008) “Making Ontologies Talk: Knowledge Interoperability in the Semantic Web”, IEEE Intelligent Systems 23(6): 72-85

[Lewotin 1982] Lewontin R. (1982) ‘Human Diversity’, New York: Scientific American Books

[Libby 1999] Libby D.,. (1999) "RSS 0.91 Spec, revision 3", Netscape Communications, Archived from the original at, http://web.archive.org/web/20001204093600/http://my.netscape.com/publish/formats/rss-spec-0.91.html [Accessed September 2011]

[Lowgren 1993] Lowgren J., (1993) “Human-computer Interaction. What every system developer should know”, Studentlitteratur, Lund

[Maedche 2002] Maedche A., Motik B., Silva N., Volz R., (2002) ‘MAFRA – A Mapping Framework for Distributed Ontologies’, In Proceeding of Knowledge Engineering and Knowledge Management: Ontologies and Semantic Web, Volume 2473, pp. 68-75

[Manchester 2003] "Knowledge Acquisition Ontology", University of Manchester, (2003) Available at: http://protege.stanford.edu/plugins/owl/owl-library/ka.owl, [Accessed September 2006]

[Maciej 2007] “Music ontology”, Maciej Music ontology, (2007) Available at http://maciej.janik/test [Accessed January 2007]

[Meilicke 2009] Meilicke C., Stuckenschmidt H., Svab-Zamazal O., (2009) *“*A Reasoning-Based Support Tool for Ontology Mapping Evaluation*”*, In proceedings of the 5th European Semantic Web Conference (ESWC

2009), Lecture Notes in Computer Science, Volume 5318, Springer, Heidelberg, pp. 878 – 882.

[McGuinness 2000] McGuinness D.L., Fikes R., Rice J., Wilder S., (2000) "The chimaera ontology environment", In Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence, pp. 1123–1124

[McGuinness 2003] McGuinness D. L., Pinheiro da Silva P., (2003) "Infrastructure for web explanations", In Proceedings of the 3rd International Semantic Web Conference (ISWC 2003), pp. 113 – 129.

[Miller 1956] Miller, G. A., (1956) "The magical number seven, plus or minus two: Some limits on our capacity for processing information", Psychological Review, 63, 81–97.

[Miller 2000] Miller R.J., Haas L.M., Hernandez M.A., (2000) 'Schema mapping as query discovery', In Proceedings of the International conference on very Large Data Bases (VLDB), pp. 77 – 88.

[Mitra 2005] Mitra P., Noy N.F., Jaiswal A.R., (2005) 'OMEN: A Probabilistic Ontology Mapping Tool', In Proceedings of International Semantic Web Conference (ISWC 2005), Lecture Notes in Computer Science, Volume 3729, Springer, Heidelberg, pp. 537 – 547.

[Mocan 2006] Mocan A., Cimpian E., Kerrigan M., (2006) "Formal Model for Ontology Mapping Creation", Proceedings of the 5th International Semantic Web Conference (ISWC 2006), Lecture Notes in Computer Science, Volume 4273, Springer, Heidelberg, pp. 459 – 472.

[Morishima 2005] Morishima A., Okawara T., (2005) "SMART: a tool for semantic-driven creation of complex XML mappings", Proceedings of the ACM SIGMOD Conference on the management of Data (SIGMOD 2005), ACM Press, pp. 909 – 911.

[Moulton 1993] Moulton MH., (1993) 'Probabilistic mapping', In Rasch Measurement Transactions, 7(1), pp. 268-269

[Murphy 2002] Murphy G.L., (2002) "The Big Book of Concepts", The MIT Press

[Musciano 2000] Musciano C., Kennedy B., (2000) "HTML & XHTML: The Definitive Guide", 4th edition, O'Reilly & Associates

[Nielsen 1993] Nielsen J., (1993) "Usability Engineering", Academic Press, San Diego/New York

[Noy 2000] Noy N.F., Fergerson R.W., Musen M.A., (2000) "The Knowledge Model of Protege-2000: Combining Interoperability and Flexibility", In Proceedings of EKAW (EKAW 2000), Springer-Verlag, pp. 17 – 32.

[Noy 2002] Noy N.F. and Musen M.A., (2002) "Evaluating ontology-mapping tools: requirements and experience." In Proceedings of OntoWeb-SIG3 Workshop at the 13th International Conference on Knowledge Engineering and Knowledge Management, pp. 1–14

[Noy 2003] Noy N.F., Musen M.A., (2003) "The PROMPT suite: Interactive tools for ontology merging and mapping", International Journal of Human computing Studies, Elsevier, 6(59), pp. 983 – 1024.

[Noy 2004] Noy N. (2004) 'Semantic Integration: A Survey of Ontology-Based Approaches', ACM SIGMOD Record 33(4), pp. 65-70.

[Noy 2008] Noy N., Griffith N., Musen M., (2008) "Collecting Community-Based Mappings in an Ontology Repository", In Proceedings of the 7th International Semantic Web Conference (ISWC 2008), Lecture Notes in Computer Science, Springer, Heidelberg, pp.371 – 386.

[OAEI 2011] The Ontology alignment evaluation initiative (OAEI), (2011) Available at http://oaei.ontologymatching.org/2011/ [Accessed September 2011]

[O'Connor 2003] O'Connor J., Robertson J., Edmund F., (2003) "Student t-test", MacTutor History of Mathematics archive, University of St. Andrews

[Osborne 2004] Osborne J. W. and Overbay A., (2004) "The power of outliers (and why researchers should always check for them)", Practical Assessment, Research & Evaluation, 9(6).

[Owens 2006] Owens, M., (2006) "The Definitive Guide to SQLite", Apress, 1st edition, ISBN 1-590-59673-0

[Pinherio 2008] Pinheiro da Silva P., McGuinness D. L., Fikes R., “A proof markup language for semantic web services”, Technical report, KSL, Stanford University, 2004, Available online at http://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-04-01.pdf [Accessed July 2008]

[Preece 2002] Preece J., Rogers Y., Sharp H., (2002) “Interaction Design: Beyond Human-Computer Interaction” John Wiley and Sons, New York

[Quan 2004] Quan T., Hui S.C., Cao T.H., (2004) “FOGA: A Fuzzy Ontology Generation Framework for Scholarly Semantic Web”, In Proceeding of the Knowledge Discovery and Ontologies Workshop

[Raghavan 2005] Raghavan A., Rangarajan D., Shen R., Goncalves, (2005) “Schema Mapper: A Visualization Tool for DL Integration”, Proceedings the Joint Conference on Digital Libraries (JCDL 2005), pp. 414.

[Robertson 2005] Robertson G.G., Czerwinski M.P., Churchill J.E., (2005) “Visualisation of mappings between schemas”, Proceedings of the SIGCHI conference on human factors in computing systems (CHI 2005), pp. 431-439.

[Ruxton 2006] Ruxton GD., (2006) “The unequal variance *t*-test is an underused alternative to Student’s *t*-test and the Mann–Whitney *U* test”, Behavioural Ecology, 2006

[Sabou 2008] Sabou M., d’Aquin M., Motta E., (2008) “Exploring the semantic web as background knowledge for ontology matching”, Journal of Data Semantics XI, pp. 156 – 190.

[Shamdasani 2009] Shamdasani J., Hauer T., Bloodsworth P., Branson A., Odeh M., McClatchey R., (2009) ‘Semantic Matching Using the UMLS’, In Proceedings of the 6th European Semantic Web Conference (ESWC 2009), Lecture Notes in Computer Science, Springer, Heidelberg, pp. 203-217

[Shi 2009] Shi F., Li J., Tang J., Xie G., Li H., (2009) “Actively Learning Ontology Matching via User Interaction”, In Proceedings of the 8th

International Semantic Web Conference (ISWC 2009), Lecture Notes in Computer Science, Springer, Heidelberg, pp. 586 – 600.

[Shneiderman 1984] Shneiderman B., (1984) "Response time and display rate in human performance with computers", ACM Computing Surveys (CSUR), 16(3), pp. 267-283.

[Shvaiko 2004] Shvaiko P., Euzenat J. (2004) "A Survey of Schema-based Matching Approaches", Journal of Data Semantics IV, pp. 146 – 171.

[Shvaiko 2005] Shvaiko P., Giunchiglia F., Pinherio da Silva P., McGuiness D., (2005) "Web Explanations for Semantic Heterogeneity Discovery", Proceedings of the 2nd European Semantic Web Conference (ESWC 2005), Lecture Notes in Computer Science, Springer, Heidelberg, pp. 303-317.

[Shvaiko 2008] Shvaiko P. And Euzenat J. (2008) 'Ten Challenges for Ontology Matching' In On the Move to Meaningful Systems: OTM 2008, Volume 5332, pp. 1164 – 1182.

[Siorpaes 2009] Siorpaes K. and Simperl E., (2009) "Human Intelligence in the Process of Semantic Content Creation", World Wide Web Journal (WWW), 13(1-2), pp. 33 – 59.

[Soantag 2008] Soantag D., (2008) "Towards Dialogue-Based Interactive Semantic Meditation in the Medical Domain", 3rd International Workshop on Ontology Matching (OM 2008) at ISWC 2008

[Swrc 2007] "SWRC Ontology", The SWRC (Semantic Web for Research Communities) ontology, available at http://ontoware.org/swrc/, [Accessed January 2007]

[Thakor 2004] Thakor M.V., Borsuk W., Kalamas M., (2004) "Hotlists and web browsing behavior–an empirical investigation", In Journal of Business Research, 57, pp. 776–786.

[Van der Wal 2007] Van der Wal T., "Folksonomy Coinage and Definition", Available at http://vanderwal.net/folksonomy.html, [Accessed 7 March 2011]

[Von Ahn 2006] Von Ahn L., “Games with a purpose”, IEEE Computer Magazine (2006), pp. 96-98

[Walls 2005] Walls C., Breidenbach R., (2005) “Spring in Action”, Manning, ISBN 1932394354

[Welch 1947] Welch, B. L. (1947), "The generalization of "Student's" problem when several different population variances are involved", Biometrika, 34 (1–2), pp. 28–35.

[Widenius 2002] Widenius M., Axmark D., (2002) “MySQL Reference Manual”, O’Reilly & Associates.

[Wright 2010] Wright T., (2010) “The Tragic Cost of Google Pac-Man”, Web blog, Available at http://blog.rescuetime.com/2010/05/24/the-tragic-cost-of-google-pac-man-4-82-million-hours/ [Accessed September 2011]

[Wu 2008] Wu G., Li J., Feng L., and Wang K., (2008) “Identifying Potentially Important Concepts and Relations in an Ontology”, In Proceedings of the 7th International Semantic Web Conference (ISWC 2008) Lecture Notes in Computer Science, Springer, Heidelberg, pp. 33 – 49.

[Yan 2003] Yan B., Frank M., Szekely P., (2003) “WebScripter: Grass-root Ontology Alignment via End-User Report Creation”, In Proceedings of the 2nd International Semantic Web Conference (ISWC 2003), Lecture Notes in Computer Science, Volume 2870, Springer, Heidelberg, pp. 676 – 689.

[Zhdanova 2006] Zhdanova A., Shvaiko P., (2006) “Community-driven ontology matching”, In Proceedings of the 3rd European Semantic web Conference (ESWC 2006), Lecture Notes in Computer Science, Springer, Heidelberg, pp. 34-49

# Appendix A

# Experiment Statistical Tests

This appendix details the results of the statistical tests used in the experiments. Unless otherwise mentioned the statistical test used was an unpaired two-tailed t-Test. In each result table a double dash, --, indicates there is no value in the cell. The P-value is tested against the α-value and represents the probability of significant difference. The statistical significance column indicates whether there was a significant difference in the results between the user groups.

## A.1 Experiment One

**Table A-1:** ***Experiment one*** – The comparison of the time taken to map with COMA++.

| **Test** | **Onto Aware** | **Tech Aware** | **Non Tech** | **P-value** (α = 0.05) | **Statistical Significance** |
|---|---|---|---|---|---|
| Total Time Taken [a] | 489.63 | 545.50 | -- | t = -0.6689, p = 0.52 > α | NONE |
| | 489.63 | -- | 596.00 | t = -1.3315, p = 0.21 > α | NONE |
| | -- | 545.50 | 596.00 | t = -0.4766, p = 0.64 > α | NONE |
| Mapping Task Time [a] | 49.14 | 39.49 | -- | t = 0.6979, p = 0.49 > α | NONE |
| | 49.14 | -- | 35.71 | t = 1.1043, p = 0.30 > α | NONE |
| | -- | 39.49 | 35.71 | t = 0.4525, p = 0.66 > α | NONE |

[a] The time is displayed in seconds

**Table A-2:** ***Experiment one*** – The comparison of the time taken to map with the NL prototype mapping tool.

| **Test** | **Onto Aware** | **Tech Aware** | **Non Tech** | **P-value** (α = 0.05) | **Statistical Significance** |
|---|---|---|---|---|---|
| Total Time Taken [a] | 773.50 | 750.63 | -- | t = 0.1840, p = 0.86 > α | NONE |
| | 773.50 | -- | 791.25 | t = -0.1219, p = 0.90 > α | NONE |
| | -- | 750.63 | 791.25 | t = -0.2725, p = 0.79 > α | NONE |
| Mapping Task Time [a] | 23.44 | 22.75 | -- | t = 0.1840, p = 0.86 > α | NONE |
| | 23.44 | -- | 23.98 | t = -0.1219, p = 0.90 > α | NONE |
| | -- | 22.75 | 23.98 | t = -0.2725, p = 0.79 > α | NONE |

[a] The time is displayed in seconds

**Table A-3:** ***Experiment one*** – The comparison of the mapping precision achieved with COMA++.

| Test | Onto Aware | Tech Aware | Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mapping Precision [a] | 0.9682 | 0.8799 | -- | $t = 1.5524, p = 0.15 > \alpha$ | NONE |
| | 0.9682 | -- | 0.7828 | $t = 2.4211, p = 0.04 < \alpha$ | YES |
| | -- | 0.8799 | 0.7828 | $t = 1.0351, p = 0.31 > \alpha$ | NONE |
| Mapping Recall [a] | 0.5140 | 0.5279 | -- | $t = -0.1151, p = 0.91 > \alpha$ | NONE |
| | 0.5140 | -- | 0.5490 | $t = -0.2626, p = 0.79 > \alpha$ | NONE |
| | -- | 0.5279 | 0.5490 | $t = -0.1708, p = 0.87 > \alpha$ | NONE |
| [a] Represented as a number between 0 and 1 | | | | | |

**Table A-4:** ***Experiment one*** – The comparison of the mapping precision achieved with the NL prototype mapping tool.

| Test | Onto Aware | Tech Aware | Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mapping Precision [a] | 0.8745 | 0.8333 | -- | $t = 0.6606, p = 0.52 > \alpha$ | NONE |
| | 0.8745 | -- | 0.7235 | $t = 3.6641, p = 0.01 < \alpha$ | YES |
| | -- | 0.8333 | 0.7235 | $t = 1.9319, p = 0.08 > \alpha$ | NONE |
| [a] Represented as a number between 0 and 1 | | | | | |

# A.2 Experiment Two

**Table A-5:** ***Experiment two*** – The comparison of the time taken to map with the tag-based mapping tool.

| Test | Onto Aware | Tech Aware | Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mean Mapping Task Time [a] | 43.23 | 28.3775 | -- | $t = 1.4316, p = 0.2 > \alpha$ | NONE |
| | 43.23 | -- | 41.79 | $t = -0.1273, p = 0.9 > \alpha$ | NONE |
| | -- | 28.3775 | 41.79 | $t = 1.4322, p = 0.2 > \alpha$ | NONE |
| [a] The time is displayed in seconds | | | | | |

**Table A-6:** ***Experiment two*** – The comparison of the mapping precision achieved with the tag-based mapping tool.

| Test | Onto Aware | Tech Aware | Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mean Mapping Precision [a] | 0.6931 | 0.8129 | -- | $t = -0.8619, p = 0.42 > \alpha$ | NONE |
| | 0.6931 | -- | 0.7833 | $t = -0.5598, p = 0.60 > \alpha$ | NONE |
| | -- | 0.8129 | 0.7833 | $t = 0.2065, p = 0.85 > \alpha$ | NONE |
| [a] The precision is represented as a number between 0 and 1 | | | | | |

**Table A-7:** ***Experiment two*** – The comparison of the mapping session prompts entered and finished with the tag-based mapping tool.

| Test | Onto Aware | Tech Aware | Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mapping Session Entered [a] | 0.4339 | 0.5164 | -- | t = 0.4056, p = 0.7 > α | NONE |
| | 0.4339 | -- | 0.3659 | t = -0.4028, p = 0.7 > α | NONE |
| | -- | 0.5164 | 0.3659 | t = -0.8499, p = 0.4 > α | NONE |
| Mapping Session Finished [a] | 0.6387 | 0.7429 | -- | t = 0.8936, p = 0.4 > α | NONE |
| | 0.6387 | -- | 0.4772 | t = -0.5848, p = 0.6 > α | NONE |
| | -- | 0.7429 | 0.4772 | t = -0.9891, p = 0.4 > α | NONE |

[a] The session entered and finished are represented as a number between 0 and 1

## A.3 Experiment Three

**Table A-8:** ***Experiment three*** – The comparison of the mapping results for the Non-Technical user group and the Ontologically Aware user group with the updated tag-based mapping tool.

| Test | Ontologically Aware [a] | Non-Technical [a] | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|
| Mapping Task Time[b] | 14.29 (5) | 24.96 (5) | t = -1.6038, p = 0.17 > α | NONE |
| Mapping Precision[c] | 0.9331 (5) | 0.6381 (5) | t = 2.5578, p = 0.05 = α | YES |
| Mapping Tasks Entered[c] | 0.3072 (6) | 0.2807 (7) | t = 0.1519, p = 0.88 > α | NONE |
| Mapping Tasks Finished[c] | 0.9113 (5) | 0.9032 (5) | t = 0.0906, p = 0.93 > α | NONE |
| Mapping Task mapped per day | 1.3152 (5) | 1.0563 (5) | t = 0.5095, p = 0.62 > α | NONE |
| RSS Stories viewed per day | 0.3976 (6) | 1.3992 (8) | t = -1.1874, p = 0.27 > α | NONE |

[a] The number of participants per user group for each test is displayed in brackets
[b] The time is displayed in seconds
[c] Precision, Tasks Entered & Finished are represented as a value between 0 and 1.

**Table A-9:** ***Experiment three*** – The comparison of the weekly mapping results for the Non-Technical user group using a slope t-Test.

| Test | Wk1[a] | Wk2[a] | Wk3[a] | Wk4[a] | Wk5[a] | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|---|---|
| Mapping Task Time[b] | 23.44 (3) | 11.28 (3) | 28.91 (4) | 32.68 (3) | 12.6 (3) | Slope = -0.02984, t = -0.01, p = 0.49 > α | NONE |
| Mapping Task Precision[c] | 0.5476 (3) | 0.4361 (3) | 0.65 (4) | 0.7667 (3) | 0.7857 (3) | Slope = 0.080675, t = 0.02, p = 0.51 > α | NONE |
| Mapping Task Entered[c] | 0.4785 (4) | 0.3205 (4) | 0.3638 (5) | 0.3669 (5) | 0.5556 (5) | Slope = 0.020047, t = 0.01, p = 0.50 > α | NONE |
| Mapping Task Answered[c] | 0.8556 (3) | 1.0000 (3) | 0.9268 (4) | 0.8056 (3) | 1.0000 (3) | Slope = 0.009444, t = 0.01, p = 0.50 > α | NONE |
| Mapping Task mapped per day | 2.5000 (3) | 1.2444 (3) | 1.1417 (4) | 1.1667 (3) | 1.5000 (3) | Slope = -0.20778, t = -0.07, p = 0.47 > α | NONE |
| RSS Stories viewed per day | 2.4383 (5) | 0.7889 (6) | 1.2200 (5) | 1.3667 (5) | 1.3333 (5) | Slope = , t = -0.06, p = 0.48 > α | NONE |

[a] The number of participants per user group for each test is displayed in brackets
[b] The time is displayed in seconds
[c] Precision, Tasks Entered & Finished are represented as a value between 0 and 1

**Table A-10:** ***Experiment three*** – The comparison of the first half of the experiment results against the second half of the experiment results for the Non-Technical user group using a paired two-tailed t-Test.

| Test (Mean) | 1st half [a] | 2nd half [a] | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|
| Mapping Task Time [b] | 26.34 (5) | 24.76 (5) | t = 0.84, p = 0.45 > α | NONE |
| Mapping Precision [c] | 0.63 (5) | 0.65 | t = -0.25, p = 0.82 > α | NONE |
| Mapping Tasks Entered [c] | 0.32 (7) | 0.25 (7) | t = 0.69, p = 0.51 > α | NONE |
| Mapping Tasks Answered [c] | 0.92 (5) | 0.89 (5) | t = 0.71, p = 0.52 > α | NONE |
| Days Interacted [d] | 10.4 (8) | 10.8 (8) | t = -0.41, p = 0.70 > α | NONE |
| Tasks mapped per day [d] | 1.59 (5) | 1.38 (5) | t = 0.34, p = 0.77 > α | NONE |
| RSS Story viewed at per day [d] | 1.83 (8) | 1.57 (8) | t = 2.45, p = 0.07 > α | NONE |

[a] The number of users for each test is displayed in brackets
[b] The time is displayed in seconds
[c] Precision, Tasks Entered & Finished are represented as a value between 0 and 1
[d] Days, Task per day and RSS per day are represented as an integer

**Table A-11:** ***Experiment three*** – The comparison of Non-Technical user group experiment results with the the standard set in the previous experiment using a one-sample t-test.

| Test | Exp 3 – Non Tech [a] | Exp 2 – Overall | Exp2 – Non Tech | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Mapping Task Time [b] | 25.27 (5) | 37.79 | -- | t = -1.996, p = 0.17 > α | NONE |
| | 25.27 (5) | -- | 41.79 | t = -2.634, p = 0.05 = α | YES |
| Mapping Task Precision [c] | 0.6381 (5) | 0.7613 | -- | t = -1.141, p = 0.32 > α | NONE |
| | 0.6381 (5) | -- | 0.7833 | t = -1.344, p = 0.25 > α | NONE |
| Mapping Task Entered [c] | 0.2772 (7) | 0.4379 | -- | t = -1.439, p = 0.20 > α | NONE |
| | 0.2772 (7) | -- | 0.3659 | t = -0.794, p = 0.46 > α | NONE |
| Mapping Task Answered [c&d] | 0.9004 (5) | 0.7034 | -- | t = 4.477, p = 0.01 < α | YES |
| | 0.9004 (5) | -- | 0.5084 | t = 8.909, p = 0.001 < α | YES |
| RSS Story per Day | 1.3992 (8) | 3.01 | -- | t = -1.921, p = 0.09 > α | NONE |
| | 1.3992 (8) | -- | 2.6722 | t = -1.518, p = 0.17 > α | NONE |
| Mapping Task per Day | 1.0563 (5) | 0.7 | -- | t = 0.991, p = 0.37 > α | NONE |
| | 1.0563 (5) | -- | 0.4667 | t = 1.732, p = 0.16 > α | NONE |

[a] The number of non technical users for each test is displayed in brackets
[b] The time is displayed in seconds
[c] Precision, Tasks Entered & Finished are represented as a value between 0 and 1
[d] Tasks Finished for Experiment 2 include any mapping session in which a mapping task was answered

**Table A-12:** ***Experiment three*** – Comparison of the unintrusive questionnaire results using Mann-Whitney test.

| Questions (1-Very, 2-Mostly, 3-Mildy, 4-Not) | Exp 3 – Non-Tech (9 users) | Exp 3 – Onto Aware (6 users) | Exp 2 – Non-Tech (4 users) | P-value (α = 0.05) | Statistical Significance |
|---|---|---|---|---|---|
| Rate how disruptive SportsFlows was to your daily work? | 3.32 | 3.67 | -- | U = 39, p = 0.18 > α | NONE |
| | 3.32 | -- | 3.00 | U = 23, p = 0.50 > α | NONE |
| Rate how abrupt the timing of the display of the mapping task prompts was? | 3.32 | 3.33 | -- | U = 35.5, p = 0.33 > α | NONE |
| | 3.32 | -- | 3.25 | U = 19.5, p = 0.83 > α | NONE |
| Rate how distractive the sport news alert information was? | 3.22 | 3.00 | -- | U = 31, p = 0.69 > α | NONE |
| | 3.22 | -- | 3.50 | U = 21, p = 0.71 > α | NONE |

**Table A-13:** ***Experiment three*** – Comparison of the appropriate stories questionnaire results using Mann-Whitney test.

| **Questions (1-Not, 2-Mildy, 3-Mostly, 4-Very)** | **Exp 3 – Non-Tech (9 users)** | **Exp 3 – Onto Aware (6 users)** | **Exp 2 – Non-Tech (4 users)** | **P-value ($\alpha$ = 0.05)** | **Statistical Significance** |
|---|---|---|---|---|---|
| Rate how appropriate the sport news stories were to your interests? | 2.56 | 2.00 | -- | U = 9.5, p = 0.69 > $\alpha$ | NONE |
| | 2.56 | -- | 2.75 | U = 19.5, p = 0.83 > $\alpha$ | NONE |
| Did the sport news stories for your interests improve over time? | 2.33 | 1.50 | -- | U = 9.5, p = 0.69 > $\alpha$ | NONE |
| | 2.33 | -- | 2.5 | U = 19, p = 0.94 > $\alpha$ | NONE |

**Table A-14:** ***Experiment three*** – Comparison of the mapping benefits questionnaire results using Mann-Whitney test.

| **Questions (1-No, 2-Mildy, 3-Mostly, 4-Yes)** | **Exp 3 – Non-Tech (9 users)** | **Exp 3 – Onto Aware (6 users)** | **Exp 2 – Non-Tech (4 users)** | **P-value ($\alpha$ = 0.05)** | **Statistical Significance** |
|---|---|---|---|---|---|
| Did answering the tasks change the sport news being displayed? | 2.89 | 1.83 | -- | U = 10, p = 0.69 > $\alpha$ | NONE |
| | 2.89 | -- | 2.00 | U = 27, p = 0.19 > $\alpha$ | NONE |
| Did this help to motivate you to answer more tasks? | 2.56 | 2.17 | -- | U = 10, p = 0.69 > $\alpha$ | NONE |
| | 2.56 | -- | 2.5 | U = 18.5, p = 0.94 > $\alpha$ | NONE |

**Table A-15:** ***Experiment three*** – Comparison of the simple questionnaire results using Mann-Whitney test.

| **Questions (1-No, 2-Mildy, 3-Mostly, 4-Yes)** | **Exp 3 – Non-Tech (9 users)** | **Exp 3 – Onto Aware (6 users)** | **Exp 2 – Non-Tech (4 users)** | **P-value ($\alpha$ = 0.05)** | **Statistical Significance** |
|---|---|---|---|---|---|
| Were the questions and supplied information easy to understand? | 3.44 | 2.83 | -- | U = 37, $p = 0.27 > \alpha$ | NONE |
| | 3.44 | -- | 2.50 | U = 28.5, $p = 0.11 > \alpha$ | NONE |
| Were there too many suggestions displayed? | 1.89 | 1.83 | -- | U = 28, $p = 0.95 > \alpha$ | NONE |
| | 1.89 | -- | 1.75 | U = 18.5, $p = 0.94 > \alpha$ | NONE |
| The task allowed you to enter the relation in your own words, is this a good way of answering the questions? | 3.33 | 3.50 | -- | U = 35, $p = 0.39 > \alpha$ | NONE |
| | 3.33 | -- | 3.00 | U = 19, $p = 0.94 > \alpha$ | NONE |
| Are the meanings of the suggestions clear? | 2.78 | 2.67 | -- | U = 28.5, $p = 0.86 > \alpha$ | NONE |
| | 2.78 | -- | 3.25 | U = 23, $p = 0.50 > \alpha$ | NONE |

# Appendix B

# Experiment One: Interview Questionnaire

## B.1 User Profile

1. Name: ______________________________
2. Group:______________________________
3. User Preference: COMA++ or Prototype:______________________
4. COMA++ Experiment Number:______________________
5. Matching Group:______________________________
6. Time Spent on Coma++:______________________________
7. Mappings Made:______________________________
8. Valid Mappings:______________________________
9. Prototype Experiment Number:______________________
10. Matching Group:______________________________
11. Time Spent on Prototype:______________________________
12. Mappings Made:______________________________
13. Valid Mappings:______________________________

## B.2 Clarity of Questions

1) Can you see, read and understand what is being asked of the question? i.e. Legible (if not why)

______________________________________________

2) Is the change in color a good way of making a distinction between the pair of terms? (if not why)

______________________________________________

3) Were their appropriate choices available for the question? i.e. Yes, Mostly, No, Skip Question (if not why)

______________________________________________

4) Are the questions too repetitious?

______________________________________________

5) Could the questions be represented/expressed in a different/better way? Should the Correspond/Similar part be emphases? % slider for mostly? What do you think?

______________________________________________

6) What do you think of the question based system? is it too linear?

______________________________________________

## B.3 Clarity of Display of Information in Support of Questions

1) Does the display give you all the information needed to make a logical decision? (if not why)

   ______________________________________________________________

2) Was the representation of attributes in italics and classes in bold a good way of distinguishing between the terms?

   ______________________________________________________________

3) Is the display easy to navigate? Can you find all the information required?

   ______________________________________________________________

4) Was the display of the overall interface easy to understand and navigate (natural, familiar)? Did the alternative choices help you in making your decision?

   ______________________________________________________________

5) Can you understand the information being represented (easy to understand and represented well)? Did it help in making a decision? Do you think you could make another mapping between the descriptions, not already given as a question, with the information? (if not why)

   ______________________________________________________________

6) Was the delay in the enabling of the buttons appropriate? Did it slow you down?

   ______________________________________________________________

## B.4 Comparison Questions Between Tools (NL prototype mapping tool & COMA++)

1) Which Display was easier to understand Tree or Natural Language? Why?

______________________________________________________________

2) Which tool had the better interface for representing the match, all matches and a single match? (why)

______________________________________________________________

3) Which tool was the most intuitive? (why)

______________________________________________________________

4) How difficult was the task of making the mappings with each tool? (why)

______________________________________________________________

5) Any Comments, Suggestions, Queries

______________________________________________________________

# Appendix C

# Added Labels of Information Service Ontology

The information service ontology is provided in the accompanying DVD media under Experiment Datasets – Experiment 2. This section lists the added labels, highlighted in italics, of the concepts in the sport news ontology.

The concept named **Football** had the following added labels:

- *Liverpool*, *Man Utd*, *Manchester United*, *Arsenal*, *Chelsea*, *Tottenham*, *Spurs*, *Man City*, *Manchester City*, *Newcastle*, *Sunderland*, *Barcelona*, *Real Madrid*, *Valencia*, *Inter*, *AC Milan*, *Juventus*, *Roma*, *Celtic*, *Rangers*: These are football teams.
- *Premiership*, *La Liga*: These are football competitions.
- *UEFA*, *FIFA*: These are governing bodies of football.

The concept named **Tennis** had the following added labels:

- *Murray*, *Federer*, *Nadal*, *Djokovic*, *Sampras*: These are tennis players.
- *Masters Title*, *FED Cup*, *Davis Cup*, *French Open*, *Wimbledon*, *Roland Garos*: These are tennis competitions.
- *Straight Sets*, *Clay-Court*, *Grass-Court*, *Hard-Court*: These are tennis terms.

The concept named **Rugby** had the following added labels:

- *European Challenge Cup*, *Six Nations*, *Magners League*, *Heinkein Cup*: These are rugby competitions.
- *Munster*, *Leinster*, *Ulster*, *Toulouse*: These are rugby teams.
- *Declan Kidney*, *Brian O'Driscoll*: These are rugby players.
- *Scrum*, *Lock*, *Coach*, *Cite*: These are rugby terms.

The concept named **MotorSports** had the following added labels:

- *F1*, *Formula One*, *Grand Prix*, *Rallying*, *Motorbikes*, *Superbikes*, *Moto GP*: These are motorsports competitions.
- *Mclaren*, *Ferrari*: These are formula one racing teams.
- *Lewis Hamilton*, *Raikkonen*, *Kovalainen*, *Ecclestone*, *Mosley*, *Rossi*, *Stoner*: These are racing drivers.

The concept named **Golf** had the following added labels:

- *The Open*, *Masters*, *St Andrews*: These are golf competitions.
- *PGA*, *LPGA*: These are golf organisations.
- *Harrington*, *Tiger Woods*, *Darren Clarke*: These are golfers
- *Birdies*, *Bogey*, *Final Hole*, *Shot*: These are golf terms.

The concept named **Cricket** had the following added labels:

- *IPL*, *Indian Premier League*, *Test*, *Twenty20*, *One-day*: These are cricket competitions.
- *ICC*, *International Cricket Council*, *ECB*, *English Cricket Board*: These are cricket organisations.
- *Essex*, *Aussie*: These are cricket teams.
- *Overs*, *County*: These are cricket terms

The concept named **Athletics** had the following added labels:

- *Olympics*, *100m*, *Marathon*, *World Indoor*: These are athletic competitions
- *Sprinter*, *Usain Bolt*: These are athletes

The concept named **Basketball** had the following added labels:

- *Michael Jordan*, *Kobe Byrant*: These are basketball players.
- *Boston*, *Lakers*, *Orlando Magic*, *Suns*, *Pistons*, *Mavericks*, *Rockets*, *New York*, *Hornets*: These are basketball teams.
- *NBA*: This is a competition that basketball teams compete in.

The concept named **GAA** had the following added labels:

- *Dublin*, *Kerry*, *Cork*, *Kilkenny*, *Waterford*, *Wexford*, *Meath*, *Westmeath*, *Tyrone*, *Armagh*, *Derry*, *Mayo*, *Galway*, *Clare*, *Tipperary*: These are football and hurling counties of the Gaelic Athletic Association (GAA).
- *Sam Maguire*, *NFL*: These are competitions the counties of the Gaelic Athletic Association (GAA) compete in.

# Appendix D

# Experiment Two: Questionnaires

## D.1 Pre-Questionnaire

**Name:** ______________________________________________

**Email:** ______________________________________________

1. What is your level of modelling experience?

Experience of ontologies – I know about data modelling/UML – Casual Web user

2. What is your level of experience using tagging on the web (e.g. flickr, del.icio.us, etc ... )

I use tagging every day – I use tagging infrequently – I don't use tagging

3. How often do you use the web

Every Day – Most Day's – Rarely

4. What is your experience of RSS feeds

I subscribe to RSS feeds – I sometimes subscribe to RSS feeds – I don't subscribe to RSS feeds

5. What is your experience of using browser add-ons

I use add-ons – I know about them but don't use them – I don't know what an add-on is

## D.2 Weekly Questionnaire

1) Rate how disruptive the extension is to your daily work:

Extremely disruptive – Very disruptive – Mildly disruptive – Not disruptive

2) Rate how abrupt the timing of the display of the question alert:

Extremely abrupt – Very abrupt – Mildly abrupt – Not abrupt

3) Rate how distractive the sport news alert information is:

Extremely distractive – Very distractive – Mildly distractive – Not distractive

4) Rate how interfering the extension is

Extremely interfering – Very interfering – Mildly interfering – Not interfering

5) Rate how appropriate the sport news information display is for your interests

Not appropriate – Slightly appropriate – Mostly appropriate – Very appropriate

6) Rate how useful it is to have an alert showing new sport news information?

Not useful – Slightly useful – Mostly useful – Very useful

7) Is the sport news information motivating you to answer more questions?

No – Not really – A little – Yes

8) Are the questions and supplied information easy to understand [screenshot of tagging interface being shown above]

All are confusing – Most are confusing – Some are confusing – None are confusing

9) Is tagging a good way to answer the questions

No – Not really – Kind of – Yes

## D.3 Final Questionnaire

1) Rate how disruptive the extension was to your daily work:

Extremely disruptive – Very disruptive – Mildly disruptive – Not disruptive

2) Rate how abrupt the timing of the display of the question alert:

Extremely abrupt – Very abrupt – Mildly abrupt – Not abrupt

3) Rate how distractive the sport news alert information was:

Extremely distractive – Very distractive – Mildly distractive – Not distractive

4) Rate how interfering the extension was:

Extremely interfering – Very interfering – Mildly interfering – Not interfering

5) Did the extension blend into the background (while undertaking your daily tasks):

Stood out – Generally noticeable – Sometimes Noticeable – Not Noticeable

6) Rate how appropriate the sport news information display is for your interests

Not appropriate – Slightly appropriate – Mostly appropriate – Very appropriate

7) Did you find that the sport news information got more tailored as time went on?

No – A little – Mostly – Yes

8) Rate how useful it is to have an alert showing new sport news information?

Not useful – Slightly useful – Mostly useful – Very useful

9) Did you notice that answering the tagging questions changed the sport news information being displayed:

No it did not – A little bit – Mostly it did – Yes it did

10) Did this help motivate you to answer more questions?

No – Not really – A little – Yes

11) Were the questions and supplied information easy to understand [screenshot of tagging interface being shown above]

All were confusing – Most were confusing – Some were confusing – None were confusing

12) Was tagging a good way to answer the questions

No – Not really – Kind of – Yes

13) Did you feel that the tagging interface is similar to other tagging interfaces you have used? Was it easier or harder?

No – Not really – Kind of – Yes

14) Were there too many tag options displayed?

No – Not really – Kind of – Yes

15) Was the meaning of the displayed tags clear to you?

No – Not really – Mostly – Yes

16) Did you use more than one tag to answer any one question?

No – Yes

17) Did you create your own tag or tags rather than the ones displayed? If you did, why?

No – Yes

18) Did you like the transparency effect of this interface?

No – Not really – Kind of – Yes

19) Rate how useful the extension was to you?

Not useful – Slightly useful – Mostly useful – very useful

20) Rate how usable you found the extension?

Not usable – Not really usable – Mostly usable – Very usable

21) Did you like the extension?

No – Not really – Kind of – Yes

22) Would you like to continue using the extension?

No – Yes

23) What improvements would you like to see to the extension?

## D.4 The Initial Questions asked during the Interviews

1. If you didn't create tags, why not?
2. How simple were the mapping tasks?
3. How engaging were the mapping tasks?
4. How interruptive were the mapping tasks?
5. What aspect would you like to change about the extension?
6. Would you rather have answered all the (mapping) questions at the start or was answering a few here and there better

# Appendix E

# Experiment Three: Recruitment Email

Hi everybody,

I would just like to let everybody know about a sports study I am currently carrying out for my PhD research which offers the participants a chance to win an ***Apple iPod Touch*** in a prize draw. The study involves using a newly made sport add-on, called SportsFlows, in your internet browser. This add-on allows you to personalise your own sporting interests, e.g. I want just want 'Football' stories, and will then let you know of any new updates for your interests to sporting websites. The purpose of the study is to investigate if SportsFlows is usable and beneficial to people. Full details can be found at

http://www.cs.tcd.ie/~coconroy/sportsflows

# Appendix F

# Experiment Three: Online Questionnaire

## F.1 Experience Section

1) Enter your Email Address (the same one as you used in the study)

________________________________________________________

2) What is your current occupation

________________________________________________________

3) Which department or school (e.g. Arts, Science, Statistics, etc...)

________________________________________________________

4) Do you know what an ontology is?

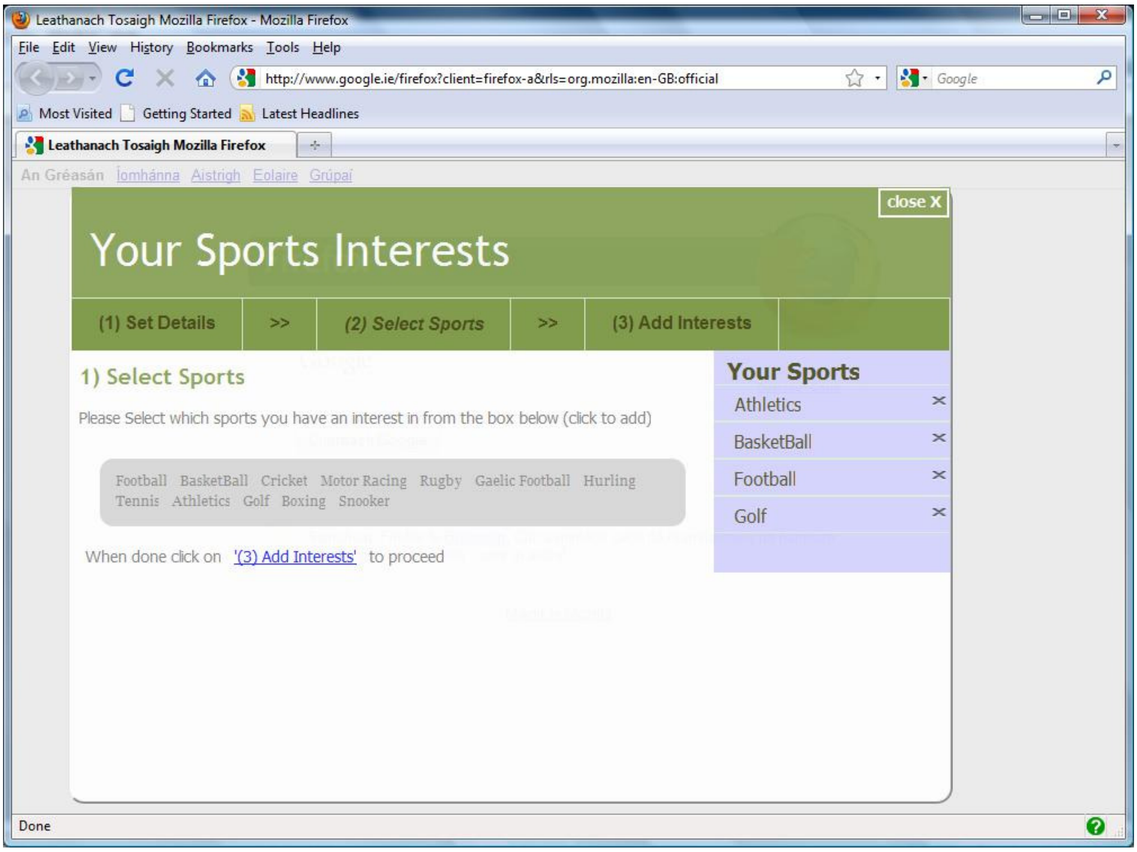
Yes | No

5) If yes, have you worked with ontology tools before?

Yes | No

## F.2 Sports Interest Section (SUS Questionnaire)

This section asks questions about how easy it was for you to add your sporting interests. A sample screenshot is displayed below. In the questions the word "system" refers to the process of adding and removing both sports and interests. If you feel you cannot provide an answer for any particular question select the number 3




1) I think that I would like to use this system frequently to express my sporting interests.

Strongly Disagree 1 2 3 4 5 Strongly Agree

2) I found the system unnecessarily complex.

Strongly Disagree 1 2 3 4 5 Strongly Agree

3) I thought the system was easy to use.

Strongly Disagree 1 2 3 4 5 Strongly Agree

4) I think that I would need the support of a technical person to be able to use this system

Strongly Disagree 1 2 3 4 5 Strongly Agree

5) I found the various functions in this system were well integrated.

Strongly Disagree 1 2 3 4 5 Strongly Agree

6) I thought there was too much inconsistency in this system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

7) I would imagine that most people would learn to use this system very quickly.

Strongly Disagree 1 2 3 4 5 Strongly Agree

8) I found the system very cumbersome to use.

Strongly Disagree 1 2 3 4 5 Strongly Agree

9) I felt very confident using the system.

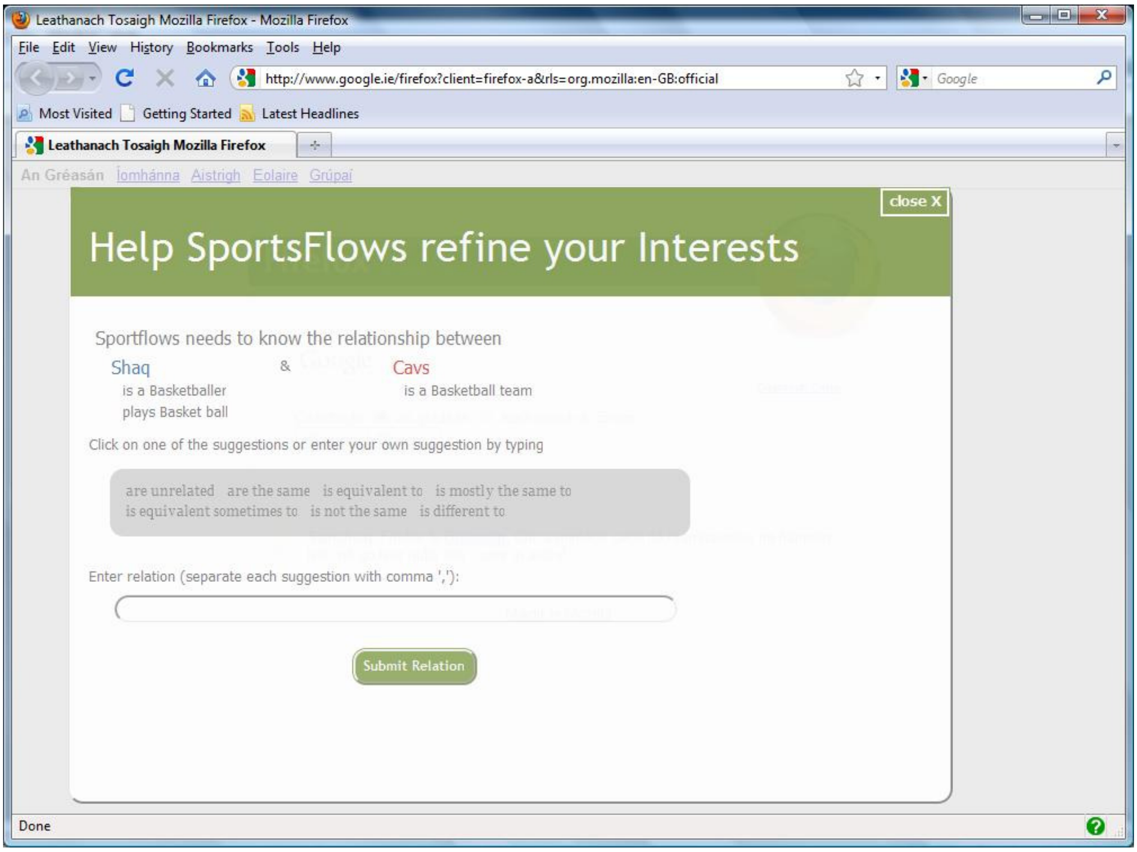
Strongly Disagree 1 2 3 4 5 Strongly Agree

10) I needed to learn a lot of things before I could get going with this system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

## F.3 Interest Refinement Task Section

This section asks questions about how easy it was for you to answer the sporting interests refinement tasks. A sample screenshot is displayed below. In the questions the word "system" refers to the process of answering questions about your sporting interests. If you feel you cannot provide an answer for any particular question select the number 3.




1) I think that I would like to use this system frequently to answer questions about my sporting interests

Strongly Disagree 1 2 3 4 5 Strongly Agree

2) I found the system unnecessarily complex.

Strongly Disagree 1 2 3 4 5 Strongly Agree

3) I thought the system was easy to use.

Strongly Disagree 1 2 3 4 5 Strongly Agree

4) I think that I would need the support of a technical person to be able to use this system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

5) I found the various functions in this system were well integrated

Strongly Disagree 1 2 3 4 5 Strongly Agree

6) I thought there was too much inconsistency in this system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

7) I would imagine that most people would learn to use this system very quickly

Strongly Disagree 1 2 3 4 5 Strongly Agree

8) I found the system very cumbersome to use.

Strongly Disagree 1 2 3 4 5 Strongly Agree

9) I felt very confident using the system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

10) I needed to learn a lot of things before I could get going with this system.

Strongly Disagree 1 2 3 4 5 Strongly Agree

11) Were the questions and supplied information easy to understand?

All were confusing – Most were confusing – Some were confusing – None were confusing

12) Were there too many suggestions displayed?

No – Not really – Kind of – Yes

13) The task allows you to enter the relation in your own words. Is this a good way of answering the questions?

No – Not really – Kind of – Yes

14) Are the meanings of the suggestions clear?

No – Not really – Mostly – Yes

15) Did answering these tasks change the sport news being displayed?

No it did not – A little bit – Mostly it did – Yes it did

16) Did this help to motivate you to answer more tasks?

No – Not really – A little – Yes

## F.4 SportsFlows Section

This section asks questions about your overall experience using SportsFlows.

1) Rate how disruptive SportsFlows was to your daily work

Extremely disruptive – Very disruptive – Mildly disruptive – Not disruptive

2) Rate how abrupt the timing of the display of the Interest Refinement Task Alert

Extremely abrupt – Very abrupt – Mildly abrupt – Not abrupt

3) The majority of participants did not answer all of the tasks prompted, why? (Please click on which are relevant)

I did not feel the need to

I did not have enough time to

I did not notice the prompt

I was busy doing something else

Sometimes when I clicked on the alert nothing happened

I did answer all of the prompted tasks

Other: ___________________________

4) Rate how distractive the sport news alert information was

Extremely distractive – Very distractive – Mildly distractive – Not distractive

5) Rate how interfering SportsFlows was with your browsing experience

Extremely interfering – Very interfering – Mildly interfering – Not interfering

6) Rate how appropriate the sport news stories were to your interests

Not appropriate – Slightly appropriate – Mostly appropriate – Very appropriate

7) Rate how appropriate the stories beginning with * were to your interests

Not appropriate – Slightly appropriate – Mostly appropriate – Very appropriate

8) Did the sport news stories for your interests improve over time?

No it did not – A little bit – Mostly it did – Yes it did

9) Rate how useful SportsFlows was to you

Not usable – Not really usable – Mostly usable – Very usable

10) Did you like SportsFlows?

No – Not really – Kind of – Yes

11) Would you like to continue using SportsFlows?

No | Yes

12) Any Comments, Queries, Problems Encountered, Suggestions

______________________________________________________

# Appendix G

# Accompanying DVD Media: Table of Contents

- Thesis.pdf: a digital copy of the thesis
- Papers: folder containing all digital copy of the academic publications of the
- Experiment Dataset
    - Experiment 1: folder containing the results from experiment 1
    - Experiment 2: folder containing the results from experiment 2
    - Experiment 3: folder containing the results from experiment 3