UNIVERSITY OF SOUTHAMPTON
FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

ENRICHING LIFELONG USER MODELLING IN A SOCIAL NETWORKING AND E-COMMERCE ERA

By
Demetris Kyriacou

A thesis submitted for the degree of Doctor of Philosophy

School of Electronics and Computer Science,
University of Southampton,
United Kingdom.

January 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING

SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

ENRICHING LIFELONG USER MODELLING IN A SOCIAL NETWORKING AND E-COMMERCE ERA

by Demetris Kyriacou

Lifelong User Modelling is an area that aims at harvesting the massive volume of web-based user information, so that it can be used for educational long-term services. Although an amount of research has addressed this issue from an educational angle, by finding ways for educational peer systems to exchange data sets among them, little has been done to investigate the potential of the social networking and e-commerce domains in the Lifelong User Modelling vision to enable Lifelong Learning services to be used.

This thesis presents a Scrutable User Modelling Infrastructure (SUMI) that can accommodate current social networking and e-commerce services for allowing the exchange of user data among these services and educational systems, for potentially improved personalisation services. Special focus has been given to three user modelling aspects: interoperability, scrutability and privacy. The suitability of SUMI regarding these aspects is tested by developing and evaluating a prototype service.

The prototype SUMI service focuses on making the infrastructure scrutable, in terms of allowing users to have control of their imported models in SUMI, while also offering various controls for safeguarding the privacy of their information in SUMI.

To assess our interoperability design decisions, four successful simulation tests have been conducted. To evaluate the proposed scrutability and privacy user privileges, two user-based evaluations have been carried out. For each evaluation, the method and results are described. Overall results have revealed that participants have expressed their general acceptance of having available such an infrastructure and service, while favouring a more holistic approach when exposed to such an environment.

Contents

Chapter 1 Introduction 1

1.1 Research Objectives and Overview 1

1.1.1 Theoretical Background 1

1.1.2 Research Questions and Hypotheses 3

1.1.3 Proposed Solution 3

1.1.4 Evaluation of Proposed Solution for Testing Research Hypotheses . 4

1.1.5 Future Research Plans 5

1.2 Thesis Structure 5

Chapter 2 Personalisation and User Modelling Walkthrough 8

2.1 Introduction 8

2.2 Intelligent Tutoring Systems 8

2.3 Adaptive Hypermedia Systems 10

2.4 User Modelling Shell Systems 13

2.5 User Modelling Servers 15

2.6 Summary 18

Chapter 3 Three Important Aspects of User Modelling 19

3.1 Introduction 19

3.2 Interoperability 20

3.2.1 User Modelling Standards 23

3.2.2 Semantic Web Technologies in User Modelling 25

3.3 Scrutability 26

3.4 User Privacy 28

3.5 Summary 30

Chapter 4 Lifelong User Modelling 32

4.1 Introduction 32

4.2 Why Lifelong User Modelling? 33

4.3 Challenges for Achieving the LUM Vision 35

4.4 Related Work 37

4.4.1 GRAPPLE Project 37

4.4.2 Linked Data Initiative 38

4.4.3 Semantic Data Integration 40
4.4.4 DataPortability Project 40
4.4.5 Open Authorization Protocols 41
4.4.6 Web Data Mashup Tools 42
4.5 Summary 43

Chapter 5 Social Networking and e-Commerce in User Modelling 44
5.1 Introduction 44
5.2 Social Networking Domain 44
5.2.1 Facebook 45
5.2.2 OpenSocial 48
5.2.3 Facebook vs OpenSocial 52
5.3 E-Commerce Domain 54
5.3.1 Amazon 54
5.3.2 eBay 57
5.3.3 Amazon vs eBay 59
5.4 Social Networking and e-Commerce in Lifelong User Modelling 61
5.5 Summary 62

Chapter 6 Research Questions and Hypotheses 64
6.1 Introduction 64
6.2 Identified Problem 64
6.3 Research Questions and Hypotheses 67
6.3.1 Interoperability 67
6.3.2 Scrutability 70
6.3.3 Privacy 71
6.4 Research Approach 73
6.5 Summary 74

Chapter 7 A Scrutable User Modelling Infrastructure 75
7.1 Introduction 75
7.2 SUMI: The Proposed Infrastructure 75
7.2.1 Acquiring User Information from Providers of User Models 77
7.2.2 SUMI Models' Architecture 81
7.2.3 Handling Different Data Models from Various Providers 85
7.2.4 Achieving Communication Among SUMI and Educational Systems 91

7.3 SUMI: The Proposed Service 92
7.3.1 Making the Infrastructure Scrutable 93
7.3.2 Safeguarding Privacy of Information 95
7.4 SUMI Use Cases 98
7.4.1 Generating Personalised Examples 99
7.4.2 Forming Student Groups 100
7.4.3 Recommending Fellow Students and Social Activities 101
7.5 Summary 102

Chapter 8 Evaluation Design and Results for Interoperability 103
8.1 Introduction 103
8.2 Evaluating Interoperability in SUMI 103
8.2.1 Retrieving user data via the SUMI Facebook application 104
8.2.2 Retrieving user data via Facebook Connect 109
8.2.3 Exporting data using the REST protocol 111
8.2.4 Using the SUMI Ontology to add value to raw user data 117
8.3 Summary 119

Chapter 9 Evaluation Design and Results for Scrutability and Privacy 120
9.1 Introduction 120
9.2 Evaluating Scrutability and Privacy in SUMI 121
9.3 Pre-Questionnaires 131
9.4 Scrutability Privileges 135
9.4.1 1st Evaluation 135
9.4.2 2nd Evaluation 138
9.4.3 Comparison of Results and Conclusions 141
9.5 Privacy Privileges 149
9.5.1 1st Evaluation 149
9.5.2 2nd Evaluation 151
9.5.3 Comparison of Results and Conclusions 154
9.6 Post-Questionnaires 162
9.7 Summary 167

Chapter 10 Summary of Contributions and Future Work 168
10.1 Introduction 168
10.2 Summary of Contributions 168
10.2.1 Interoperability 169

10.2.2 Scrutability 172
10.2.3 Privacy 174
10.3 Future Work Agenda 178
10.4 Summary 180

References 181

Appendix A Description of Facebook & OpenSocial APIs 201

Appendix B Description of Amazon & eBay APIs 207

Appendix C Properties of SUMI Ontology 226

Appendix D User Evaluation Structures 229

List of Figures

Figure 2-1 Process for Analysis of Student's Solutions (Brusilovsky, 1994) 10
Figure 2-2 The Revised Taxonomy of Adaptive Hypermedia Techniques (Knutov *et al.,* 2009) 12
Figure 2-3 Architecture of the Personis Server (Kay *et al.*, 2002) 18
Figure 3-1 Protocol for Exchanging User Profiles (Alrifai *et al.,* 2006) 22
Figure 3-2 Relationship between IEEE PAPI and IMS LIP (IMS LIP Specification) .. 25
Figure 4-1 Architectural Overview of Personal UM Clouds (Dolog *et al.,* 2009) 34
Figure 5-1 Standard web application - top VS Facebook application architecture – bottom (Facebook Developers Home Page) 47
Figure 5-2 Facebook Connect Authorization screen/message (Facebook Developers Home Page)..... 48
Figure 5-3 The OpenSocial vision (OpenSocial Home Page) 49
Figure 5-4 Incorporating Friend Connect into a third-party website (Friend Connect Home Page)..... 51
Figure 5-5 Social Networking domain – Facebook vs. OpenSocial. 53
Figure 5-6 Dominance of Facebook and OpenSocial around the world (TechCrunch, 2009) 54
Figure 5-7 E-Commerce domain – Amazon vs. eBay 60
Figure 6-1 The multi-'*domain*'sional big picture (Kyriacou *et al.,* 2009a) 66
Figure 7-1 SUMI Ontology – Protégé Screenshot..... 86
Figure 7-2 Graphical representation of a SUMI use case 99
Figure 8-1 Welcome screen for the SUMI Facebook application 105
Figure 8-2 'My Facebook model' presented in the proposed four-category SUMI models' architecture..... 106
Figure 8-3 'My Friend's Facebook model' presented in the four-category SUMI model's architecture..... 107
Figure 8-4 The SUMI Wall..... 108
Figure 8-5 The SUMI Discussion Board 108
Figure 8-6 Inviting friends to add the SUMI application 109
Figure 8-7 Asking users to login before retrieving their Facebook data 110

Figure 8-8 Retrieving user's Facebook data and displaying them in the four-category SUMI model's architecture 110
Figure 8-9 List of current educational systems' subscriptions for receiving user data 111
Figure 8-10 Second step of the export transaction. List of "Must" and "Want" attributes 112
Figure 8-11 SUMI assigns URIs to the exported attributes 112
Figure 8-12 URI exposes the information regarding the Facebook attribute *activities* 113
Figure 8-13 Example of an AHA! adaptive lesson on HTML unordered lists with a generated personalised example 114
Figure 8-14 AHA!'s generated personalised example based on a user's Facebook activities and interests. 115
Figure 8-15 List of subscriptions by educational systems 116
Figure 8-16 "Must" and "Want" attributes – Users inspect and approve transaction details 116
Figure 8-17 Enriching displayed user data with information from the SUMI ontology 118
Figure 9-1 The three categories of visualisation of user models – ordered, absolute and relative (Vernero *et al.,* 2009) 130
Figure 9-2 Competence results (%) for 1st evaluation 141
Figure 9-3 Competence results (%) for 2nd evaluation 142
Figure 9-4 Understanding of consequences results (%) for 1st evaluation 142
Figure 9-5 Understanding of consequences results (%) for 2nd evaluation 143
Figure 9-6 Acceptance results (%) for 1st evaluation 143
Figure 9-7 Acceptance results (%) for 2nd evaluation 144
Figure 9-8 During-task questions' results (%) for 1st evaluation 144
Figure 9-9 During-task questions' results (%) for 2nd evaluation 145
Figure 9-10 Competence results (%) for evaluation 1 154
Figure 9-11 Competence results (%) for evaluation 2 155
Figure 9-12 Understanding of consequence results (%) for evaluation 1 155
Figure 9-13 Understanding of consequence results (%) for evaluation 2 156
Figure 9-14 Acceptance results (%) for evaluation 1 156
Figure 9-15 Acceptance results (%) for evaluation 2 157
Figure 9-16 During-task questions' results (%) for evaluation 1 157
Figure 9-17 During-task questions' results (%) for evaluation 2 158

List of Tables

Table 3-1 Comparison of IEEE PAPI and IMS LIP 25
Table 5-1 Facebook and OpenSocial APIs 52
Table 5-2 Amazon Product Advertising API and eBay API 60
Table 7-1 List of social networking and e-commerce services (from the 2009 top-50 websites) that currently offer APIs 81
Table 7-2 The four categories of user data that can be identified in the social networking and e-commerce domains 83
Table 7-3 Proposed SUMI models' architecture 84
Table 7-4 Categories of the SUMI ontology 88
Table 9-1 Comparison of the two evaluation frameworks 129
Table 9-2 Pre-Questionnaire – Question 1 132
Table 9-3 Pre-Questionnaire – Question 2 132
Table 9-4 Pre-Questionnaire – Question 3 133
Table 9-5 Pre-Questionnaire – Question 4 134
Table 9-6 Results (% of successful responses) for the proposed Scrutability Privileges in the first evaluation 137
Table 9-7 Results (% of successful responses) for the proposed Scrutability Privileges in the second evaluation 140
Table 9-8 Summary of conclusions for proposed scrutability privileges (1st evaluation) 146
Table 9-9 Summary of conclusions for proposed scrutability privileges (2nd evaluation) 148
Table 9-10 Results (% of successful responses) for the proposed User Privacy privileges (1st evaluation) 151
Table 9-11 Results (% of successful responses) for the proposed User Privacy privileges (2nd evaluation) 153
Table 9-12 Summary of conclusions for proposed privacy privileges for the first evaluation 159
Table 9-13 Summary of conclusions for proposed scrutability privileges for the second evaluation 161
Table 9-14 Post-Questionnaire – Question 1 162

Table 9-15 Post-Questionnaire – Question 2 163
Table 9-16 Post-Questionnaire – Question 3 163
Table 9-17 Post-Questionnaire – Question 4 164
Table 9-18 Post-Questionnaire – Question 5 165
Table 9-19 Post-Questionnaire – Question 6 166
Table C-1 DatatypeProperties in SUMI Ontology 226
Table C-2 ObjectProperties in SUMI Ontology 228
Table D-1 Evaluation 1 - Scrutability User Privilege 1 230
Table D-2 Evaluation 1 - Scrutability User Privilege 2 231
Table D-3 Evaluation 1 - Scrutability User Privilege 3 232
Table D-4 Evaluation 1 - Privacy User Privilege 1 234
Table D-5 Evaluation 1 - Privacy User Privilege 2 234
Table D-6 Evaluation 1 - Privacy User Privilege 3 235
Table D-7 Evaluation 1 - Pre-Questionnaire 237
Table D-8 Evaluation 1 - Post-Questionnaire 238
Table D-9 Evaluation 2 - Scrutability User Privilege 1 239
Table D-10 Evaluation 2 - Scrutability User Privilege 2 241
Table D-11 Evaluation 2 - Scrutability User Privilege 3 242
Table D-12 Evaluation 2 - Privacy User Privilege 1 244
Table D-13 Evaluation 2 - Privacy User Privilege 2 245
Table D-14 Evaluation 2 - Privacy User Privilege 3 246
Table D-15 Evaluation 2 - Pre-Questionnaire 247
Table D-16 Evaluation 2 - Post-Questionnaire 249

Declaration of Authorship

I, Demetris Kyriacou, declare that the thesis entitled “Enriching Lifelong User Modelling in a Social Networking and E-Commerce Era” and work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as :
    1. Kyriacou E, D. and Davis, H. (2008) Moving towards life-long User Modeling. In: Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008), pages, 647-648, July 1st July - 5th July, 2008, Santander, Cantabria, Spain.
    2. Kyriacou E, D. (2008) A Scrutable User Modelling Infrastructure for enabling life-long User Modelling. In: Proceedings of the Adaptive Hypermedia and Adaptive Web-Based Systems - Doctoral Consortium, pages 421-425, 29th July - 1st August 2008, Hannover, Germany.
    3. Kyriacou E, D. (2009) Enriching Lifelong User Modelling with the Social e-Networking and e-Commerce "Pieces of the Puzzle". In: Adjunct Proceedings of the First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation

(UMAP2009), *Doctoral Consortium paper*, pages 1-6, 22nd - 26th June 2009, Trento, Italy.

4. Kyriacou, D., Davis, H. and Tiropanis, T. (2009) Evaluating Three Scrutability and Three Privacy User Privileges for a Scrutable User Modelling Infrastructure. In: Proceedings of the First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation (UMAP2009), pages 428-434, 22nd - 26th June 2009, Trento, Italy.
5. Kyriacou E, D., Davis, H. and Tiropanis, T. (2009) A (multi'domain'sional) Scrutable User Modelling Infrastructure for Enriching Lifelong User Modelling. In: Proceedings of the First Lifelong User Modelling Workshop (in conjunction with the First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation - UMAP2009), pages 46-53, 26th June 2009, Trento, Italy.

Signed: Demetris Kyriacou

Date: 04 January 2010

Acknowledgements

Firstly, I would like to thank my main supervisor, Professor Hugh C. Davis, for taking me on board, making me a member of his postgraduate team and supervising my research during three years of hard work. In addition, I would like to thank my second supervisor, Dr Thanassis Tiropanis, for his extremely valuable feedback throughout my degree.

Secondly, I would like to acknowledge the contribution of my research 'family', the Learning Societies Lab at the University of Southampton, for helping me mature as a researcher and becoming a member of the (Lifelong) User Modelling community.

Lastly, I would like to express my deep appreciation to my parents, Evangelos and Maria Kyriacou, for their constant support and to a special person in my life, Maria Rousha, for being there for me in good and bad times.

Definitions and Abbreviations Used

AHA	Adaptive Hypermedia Architecture
AHS	Adaptive Hypermedia Systems
API	Application Programming Interface
AWS	Amazon Web Services
B2B	Business-to-Business
B2C	Business-to-Consumer
CD	Compact Disc
DERI	Digital Enterprise Research Institute
DVD	Digital Versatile Disc
FBML	Facebook Markup Language
FOAF	Friend Of A Friend
FQL	Facebook Query Language
GALE	GRAPPLE Adaptive Learning Environment
GRAPPLE	Generic Responsive Adaptive Personalized Learning Environment
GUMF	Grapple User Modeling Framework
GUMO	General User Modelling Ontology
HCI	Human Computer Interaction
HTTP	Hypertext Transfer Protocol
IT	Information Technology
ITS	Intelligent Tutoring Systems
JSON	JavaScript Object Notation
LIP	Learner Information Package
LMS	Learning Management System
LUM	Lifelong User Modelling
OWL	Web Ontology Language
PAPI	Personal And Private Information
QTI	Question and Test Interoperability
RDCEO	Reusable Definition of Competency or Educational Objective
RDF	Resource Description Framework
REST	Representational state transfer

RSS	Rich Site Summary
SOAP	Simple Object Access Protocol
SPARQL	Simple Protocol and RDF Query Language
SQL	Structured Query Language
SUMI	Scrutable User Modelling Infrastructure
UM	User Modelling
UMAP	User Modelling, Adaptation, and Personalization
UMS	User Modelling Shells
URI	Uniform Resource Identifier
WSDL	Web Service Definition Language
WWW	World Wide Web
XML	Extensible Markup Language

Chapter 1 Introduction

1.1 Research Objectives and Overview

1.1.1 Theoretical Background

The first objective of this thesis is to summarise the relevant theoretical background regarding the areas of personalisation and Lifelong User Modelling (LUM), with a special focus on three main User Modelling components: interoperability, scrutability and privacy. In addition, it is important to review the current state of the social networking and e-commerce domains while acknowledging any potential contribution to the area of Lifelong User Modelling.

Personalisation is an important factor for standalone and web-based systems that aim to offer effective and efficient services to their users. Every person should be considered independently and every application should personalise its content according to users' needs, goals and characteristics. Our research in this area has revolved around the progress from Intelligent Tutoring Systems to Adaptive Hypermedia Systems to understand the various forms of personalisation components and services offered throughout the years.

User Modelling is the core component for the majority of personalisation systems. By maintaining a model, i.e., a record for each user, User Modelling permits the unique adaptation and presentation of the available resources based on these models, thus potentially improving the services that are offered by the personalisation systems to the users. Our review of the literature in this area has examined the progress from User Modelling Shells, that is, systems that introduced a new internal component that was responsible for user modelling and separated it from the component that was generating personalisation services, to User Modelling Servers, that is, systems that

adopted a client-server architecture for storing and maintaining user information in a central repository; the research has focused particularly on three important User Modelling components:

- *Interoperability*: With the recent introduction of the Semantic Web, the issue of interoperability has become a crucial issue in the area of User Modelling. Exchanging user profiles across various sources in distributed e-Learning environments cannot be achieved without the development, adoption and support of explicit and widely accepted protocols, frameworks and architectures.
    - *User Modelling Standards*: User Modelling Standards ensure consistency on how users should be modelled, since they propose universal structures and set the scope of user information that needs to be gathered so adequate User Modelling can be carried out.
    - *Semantic Web Technologies*: Semantic Web Technologies can play a leading role in finding the solutions to several questions in User Modelling, especially on the aspect of interoperability. Expressing user models in RDF and defining their relationships using ontologies is currently a common practice in the community of User Modelling (Dolog and Schaefer, 2005a; Dolog and Schaefer, 2005b; Alrifai *et al.*, 2006).
- *Scrutability*: The term 'scrutability' in User Modelling signifies that every user model can be controlled by its owner in order to determine what has been modelled about them and how that modelling process was conducted. The benefit of adopting scrutability in User Modelling can be indentified when users are allowed to select the stereotypes in which they prefer to be included or from which excluded. Furthermore, scrutable solutions allow users to alter the value of any single inference that is used for drawing conclusions about them (Kay, 2000).
- *Privacy*: As the literature in the area reveals, the privacy of user information is an important issue to consider in User Modelling, and privacy components ought to be included in every User Modelling solution. Users should be given the option to set the privacy status of their

information and be allowed to determine who can access which part of their models.

Lifelong User Modelling is the driving force towards the Lifelong Learning vision. Various challenges are identified by the LUM community and several attempts and contributions, like the Linked Data Initiative and the DataPortability Project, have led the way by providing pragmatic approaches that help applications to share and connect web-based user data (Kay and Kummerfeld, 2009b). Further research is required to tackle important questions, which will help the users to make the most of their long-term, even lifetime, data while enabling them to take ownership of their personal information.

Special consideration has been given to two rich sources of user information: the social networking and e-commerce domains. Although it has been acknowledged by the Lifelong User Modelling community that these two domains can bring significant value to the area, little has been done to harvest the huge potential that can result when gathering user information from several providers of user models from these two domains, and using that information for educational purposes.

1.1.2 Research Questions and Hypotheses

The second objective of this research is to define research questions and hypotheses, which will be backed up with a previously presented extensive literature review. Furthermore, at that point, we aim to identify a problem in the area of Lifelong User Modelling for which we have proposed a solution.

1.1.3 Proposed Solution

The third objective of this work is to propose a solution that will address the various important issues outlined in section 1.1.1, and attempt to combine them in order to test the pre-defined research hypotheses. A Scrutable User Modelling Infrastructure (SUMI) is presented that blends these User Modelling "ingredients", which, as the literature reveals, are essential components in current LUM contributions, and by taking advantage of Semantic Web technologies, an attempt is made to accomplish the research's primary goal, which is to evaluate whether it is possible to have a user modelling infrastructure that is interoperable, scrutable and maintains the privacy of its

users, while being adequate to accommodate current e-commerce and social networking services. To achieve such an infrastructure, several challenges have to be acknowledged:

- *Interoperability*: to achieve interoperability while retrieving user information from providers of the two examined domains and passing that information towards educational (personalisation) systems. For this to occur, smaller steps have to be taken:
    - Recognise realistic methods for retrieving user data from social networking and e-commerce services by respecting their different policies and rules
    - Examine the possibility of a common models' architecture
    - Introduce a solution for handling the various data models from the several providers of the two examined domains
    - Identify a communication protocol that can assist in passing the retrieved user information to educational systems in order to enrich their internal user models, which are currently being used for personalisation services.
- *Scrutability*: Allow users to have control over their models by offering them solid scrutability privileges.
- *Privacy*: Provide information privacy privileges to users, when retrieving, displaying, processing and exchanging their personal data.

1.1.4 Evaluation of Proposed Solution for Testing Research Hypotheses

The next objective of the work presented in this thesis is to present quantitative proof of the successful evaluation of the research hypotheses and to demonstrate realistically the practical implementation of the proposed solution. For this purpose, we conducted simulation testing using hypothetical scenarios with potential users to evaluate the interoperability aspect, where, to evaluate scrutability and privacy, we executed two user evaluations with 107 and 111 participants respectively. We present the evaluations' structure and frameworks while we support our design decisions and the approaches followed by linking them to significant contributions that are presented in the literature of User Modelling. In addition, we describe the obtained results and

conclusions, which we believe demonstrate and prove how our proposed solution can realistically tackle the identified problem.

1.1.5 Future Research Plans

The final objective of this thesis is to provide our own predictions and estimations for the future of Lifelong User Modelling, not only inside, but also beyond the educational domain, while outlining and proposing further research directions in this area.

1.2 Thesis Structure

This thesis is divided into ten chapters as follows:

- Chapter 1: Introduction
- Chapter 2: Personalisation and User Modelling Walkthrough
- Chapter 3: Three Important Aspects of User Modelling
- Chapter 4: Lifelong User Modelling
- Chapter 5: Social Networking and e-Commerce in User Modelling
- Chapter 6: Research Questions and Hypotheses
- Chapter 7: A Scrutable User Modelling Infrastructure
- Chapter 8: Evaluation Design and Results for Interoperability
- Chapter 9: Evaluation Design and Results for Scrutability and Privacy
- Chapter 10: Summary of Contributions and Future Work

Chapter 2 provides a historical walkthrough in the important areas of:

- Personalisation systems, by focusing on Intelligent Tutoring Systems and Adaptive Hypermedia Systems, and
- User Modelling, by concentrating on User Modelling Shell Systems and User Modelling Servers.

Chapter 3 elicits three critical aspects of User Modelling, as identified from the literature in the field:

- Interoperability

    - This section analyses some of the most significant work published on achieving interoperability in the educational domain while exchanging fragments of user models between different systems.
    - Furthermore, well accepted and widely used standards for User Modelling are described in this section.
    - In addition, with the introduction of the Semantic Web, various technologies have been introduced and this section investigates how they have been applied in the field of User Modelling.
- Scrutability
    - This section examines the issue of scrutability in User Modelling, meaning the process of allowing users to inspect and define how they want to be modelled.
- User Privacy
    - Privacy of user information is an important issue to be considered in User Modelling and it is described in this section.

Chapter 4 introduces the focus area of this thesis, namely, Lifelong User Modelling, by explaining its importance and analysing the main challenges for achieving its lifelong learning vision. Furthermore, it presents related state-of-the-art work that contributes to the area of Lifelong User Modelling.

Chapter 5 elaborates on the two examined domains in this thesis, which are the social networking and e-commerce domains. More specifically, it provides four case studies, two from each domain:

- Facebook and OpenSocial for the social networking domain
- Amazon and eBay for the e-commerce domain

As is illustrated in Chapter 5, by choosing these services for the four conducted case studies we test the part of our claim that a solution that applies to the majority of e-commerce and social networking services is possible. In addition, we point out the relevance between these two important domains and the Lifelong User Modelling field.

Chapter 6 describes the identified problem in the area of LUM and presents our research questions and hypotheses. Furthermore, it describes our approach for testing our hypotheses while looking for answers to the research questions.

Chapter 7 presents our proposed solution: a Scrutable User Modelling Infrastructure (SUMI) that focuses on three key elements of Lifelong User Modelling: interoperability, scrutability and user privacy. Furthermore, we describe our

contribution, which consists of an infrastructure and a prototype online service. Finally, the chapter provides three hypothetical realistic scenarios that demonstrate the potential usefulness of our proposed solution.

Chapter 8 focuses on the evaluation of the interoperability aspect of SUMI by describing the design and results of the four tests we conducted for achieving interoperability in our proposed solution.

Chapter 9 complements Chapter 8 by presenting the frameworks and results of the two conducted user evaluations for scrutability and privacy in SUMI. It clearly provides the numbers behind the results and explains the obtained conclusions and lessons by analysing the participants' responses during the two evaluations.

Finally, *Chapter 10* offers an overview of our contribution to the area of Lifelong User Modelling by revisiting the research questions and hypotheses that were stated in Chapter 6. In addition, it describes our short and long term work agendas, while also expressing our informed projections regarding future research directions for this area.

Chapter 2 Personalisation and User Modelling Walkthrough

2.1 Introduction

Chapter 2 reviews the literature in the field of Personalisation Systems and User Modelling by providing a chronological review of:

- the progress from Intelligent Tutoring to Adaptive Hypermedia Systems in order to identify the extensive work on personalisation in these two areas, and
- the evolution from User Modelling Shells to User Modelling Servers in order to understand the evolution of User Modelling over the years.

The reason we reviewed the literature in these areas was to understand how personalisation is performed by two categories of educational personalisation systems, Intelligent Tutoring and Adaptive Hypermedia Systems, regarding the use of user models for generating personalisation services. Furthermore, our research revolves around enriching user modelling sets that are used for generating personalisation services; thus, we wanted to identify current systems in the area of User Modelling that gather user data, in the form of user models, for educational purposes.

2.2 Intelligent Tutoring Systems

Intelligent Tutoring Systems (ITS) are computer-based instructional systems with models of instructional contents that specify *what* to teach and teaching strategies that specify *how* to teach (Murray, 1999). At the beginning, ITS provided little or no

learning material. Their most important duty was to support students in the process of problem solving. It was assumed, then, that the required knowledge about the subject was acquired by attending lectures or by reading books (Brusilovsky, 2000). The combination of an expertise-intelligent module, for knowledge generation and verification, along with a tutoring module for teaching and student interaction support, and a student model module for storing the knowledge state of every student, formed the structure of ITS (Brusilovsky, 1994). The goal of ITS is the use of knowledge about the domain and the student, and of predefined teaching strategies to support flexible and individualized learning and tutoring (Brusilovsky, 1999).

There are four core ITS technologies (Brusilovsky, 1999):

- *Curriculum sequencing*: the student is provided with the most suitable planned sequence of knowledge units to learn in a sequence of learning tasks.
- *Intelligent analysis of student's solutions:* a solution analyser deals with the student's final answers and provides extensive error feedback. Figure 2-1 presents the underlying process followed by ITS for comparing submitted student solutions with correct solutions and generating feedback.
- *Interactive problem solving support:* the student is provided with intelligent help at each step of problem solving.
- *Example-based problem solving support:* extensive help is provided in the form of examples based on previously solved problems.

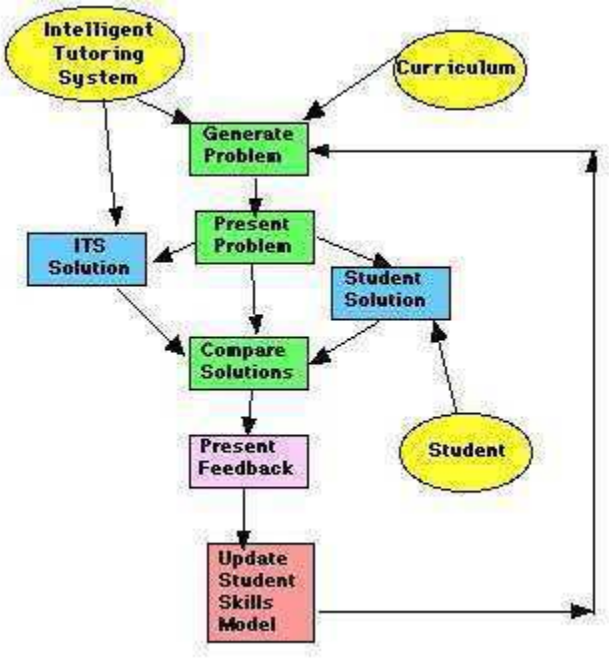

Figure 2-1 Process for Analysis of Student's Solutions (Brusilovsky, 1994)

As computers evolved, it was found reasonable to provide some learning material and for it to be merged into one package along with the supporting features of ITS (Brusilovsky, 2000). The best option for organizing material was found to be hypertext, which later became hypermedia. The combination of ITS' features and an electronic learning environment organized as hypermedia was the starting point of Adaptive Hypermedia Systems (Brusilovsky, 2000).

2.3 Adaptive Hypermedia Systems

Adaptive Hypermedia Systems (AHS) are all the hypertext and hypermedia systems that reflect some features of the user in a model, which can be considered as a record for a user, and that apply this model to adapt various visible aspects of the system to the user (Brusilovsky, 1996a). AHS apply various forms of durable user models to personalise the content and the links of hypermedia pages to every user. Although the intelligent module, which was the main component of the ITS, was omitted from the AHS structure, it was possible to bridge the gap between computer–driven tutoring systems (ITS) and student-driven educational environments (Hypermedia) by taking advantage of the user models to provide effective content adaptation (Brusilovsky, 1995). The goal of AHS is to use the models of their users

throughout the interaction with them, in order to adapt the hypermedia content and links to their needs.

Up until 1996, a number of standalone AHS were built. Education was always the main application area, but all systems developed were essentially laboratory systems for the exploration of new methods and techniques (Brusilovsky, 2000). In 1996, the rapid growth of the WWW and the publication of Brusilovsky's taxonomy (Brusilovsky, 1996a, 1996b, 2001) shifted the focus from standalone systems to the development of web-based systems, since the web platform made them available all around the world and enabled these systems to be viable for much longer. Two of the most important AHS in the educational context, Adaptive Hypermedia Architecture, known as AHA! (AHA! Home Page), and InterBook (InterBook Home Page) provided their students with high-standard personalisation, which demonstrated how usable and useful the addition of the new component, the user model, was:

- AHA! keeps a user model for each user, which is updated based on the user's knowledge about concepts. The knowledge is generated and changed when pages are read and tasks completed during an adaptive course. AHA! supports both adaptive context presentation with fragment variants, and adaptive navigation support with (mainly) link colour annotation (De Bra and Calvi, 1998).
- InterBook allows authors to deliver adaptive electronic textbooks on the WWW. For each user, InterBook creates and maintains a user model, which is also based on the user's knowledge about concepts. It also takes into account the user models to provide adaptive guidance, adaptive navigation support and adaptive help (Brusilovsky *et al.*, 1998).

Two methods for adaptation and the most important techniques for applying these methods are described in Brusilovsky (1996b, 1999), updated in Brusilovsky (2001), revised in Knutov *et al.* (2009) and presented below:

- *Adaptive presentation*: the goal is to adapt the content of a hypermedia page to a user based on the user model. The most popular technique is adaptive text presentation with fragments that can be inserted or removed.
- *Adaptive navigation support*: the goal is to help users find their paths in hyperspace by adapting link presentation and functionality based on the user model. The most popular techniques are direct guidance, adaptive link annotation and adaptive link hiding.

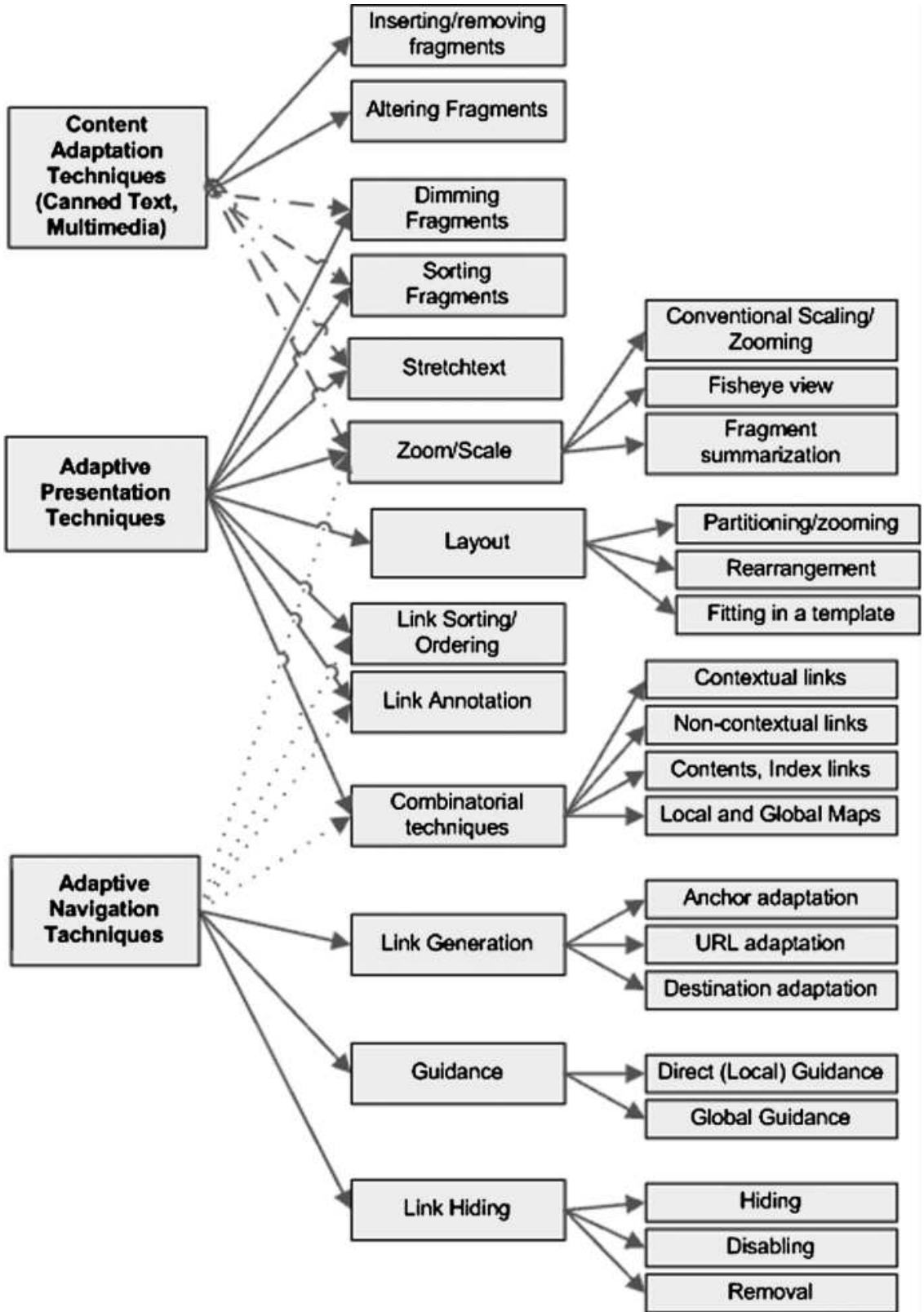

Figure 2-2 The Revised Taxonomy of Adaptive Hypermedia Techniques (Knutov *et al.,* 2009)

2.4 User Modelling Shell Systems

In the initial stages, User Modelling (UM) was performed by the application system; there was no clear distinction of the UM component. From the mid-eighties onwards, the separation from other components was made, but no reusability was supported for future development. Some examples can be found in Kobsa (1985), Sleeman (1985), Kass (1988) and Allgayer *et al.* (1989).

Kobsa (1990) was the first to introduce the term 'User Modelling Shell' (UMS) systems to illustrate software decomposition and abstraction along with support of modifiability and reusability. The common attribute shared by every UMS was the ability to classify users in groups (stereotypes) and make assumptions about their characteristics based on the stereotypes to which they belonged. These assumptions were formed based on the users' interaction history and their behaviour with the system, and they were generalised into stereotypes, such as beginners and experts, once a pattern was identified. The UMS were capable of providing justification for the assumptions made for every user while also providing consistent maintenance of the user models and evaluating each new entry by comparing it with pre-defined standards in order to classify it into a group of users (Kobsa, 1995).

In order for a system to be classified as a UMS it had to meet some essential requirements (Kobsa, 2001):

- Generality, including domain independence
    - Shell systems were required to be usable in as many applications and content domains as possible, and within these domains, for as many User Modelling tasks as possible. Examples can be found in Huang *et al.* (1991), Kono *et al.* (1994) and Paiva and Self (1995).
- Expressiveness
    - Shell systems were expected to be able to express as many types of assumptions about the user as possible at the same time. Examples from this category are found in Kobsa and Wahlster. (1989) and Taylor *et al.* (1996).
- Strong inferential capabilities
    - Shell systems were expected to perform all sorts of reasoning, such as first-order predicate logic, reasoning with uncertainty and conflict resolution when contradiction was detected. Examples are

presented in Kobsa and Wahlster (1989) and Zukerman and Litman (2001).

The most important academic developments, as described in Kobsa (2001), were applied in the educational context:

- UMT
    - This allows the user-model developer to define hierarchically ordered user stereotypes, and rules for user model inferences as well as contradiction detection (Brajnik and Tasso, 1994).
- BGP-MS
    - This allows assumptions about the user and stereotypical assumptions about user groups to be represented in first-order predicate logic (Kobsa and Pohl, 1995).
- DOPPELGANGER
    - This accepts information about the user from hardware and software sensors (Orwant, 1995).
- TAGUS
    - The system represents assumptions about the user in first-order formulae. It allows the stereotype hierarchy to be defined and contains an inference mechanism, a truth maintenance system, and a diagnostic subsystem that includes a library of misconception (Paiva and Self, 1995).
- um toolkit
    - A toolkit for UM that assists in representing assumptions about the user's knowledge, beliefs, preferences and other user characteristics, in attribute-value pairs (Kay, 1995).

A further important contribution is UbisWorld (Heckmann *et al.,* 2009), found at *http://www.ubisworld.org/*, which can be considered as a playground that can be used to represent parts of the real world like an office, a shop, a museum, an airport or a city. It represents persons, objects, locations as well as times, events and their properties and features. UbisWorld could be understood as a virtual coloured blocks world where each colour represents a different category in the ontology. Apart from the representational funtionality, UbisWorld can be used for simulation, inspection and control of the real world. In addition the knowledge about concepts, individuals and relations in UbisWorld is dynamically transferred to the semantic web languages RDF and OWL

and modeled in two ontologies: GUMO, the General User Model Ontology and UbisOntology, the ontology for ubiquitous computing.

In the mid-1990s, personalisation systems shifted towards different domains with less demanding User Modelling requirements, such as user-tailored websites; thus, complex UM and reasoning abilities became redundant. Measuring user behaviour was only used as an information source for making assumptions about a user's attitudes, and was not regarded as a phenomenon that should be analysed and modelled per se. Consequently, UMS did not enjoy a wide distribution, even in the research community (Kobsa, 2001).

2.5 User Modelling Servers

In the late 1990s, the value of web personalisation was increasingly recognized in the area of electronic commerce, and allowed systems to target their product offerings, sales promotions, product news, and so on, to each individual client by taking into account the user's navigation and purchase history along with their interactions with e-business services (Hof *et al.,* 1998; Allen *et al.,* 1998; Hagen *et al.,* 1999).

In order to escape from the structure of the UMS, which had the User Modelling component integrated into the application, a new kind of systems, called User Modelling Servers, was introduced. These systems adopt the client-server architecture and permit information maintenance in a central repository, which allows applications to access that information simultaneously. In this way, data about the users are stored in a non-redundant way and applications can access user information gathered by other systems (Fink and Kobsa, 2000; Billsus and Pazzani, 2000).

The issue of security was a high priority in the development of User Modelling Servers. Systems had to apply security features, such as identification, authentication, access control and encryption of data, in order to protect the users' privacy of information. In every development of a User Modelling Server there were some important requirements that had to be adopted (Kobsa, 2001):

- Comparison of different users' selective actions
    - It was found useful to match users' selective actions (buying items, bookmaking, putting them into the shopping cart) with those of other users in order to predict users' future selective

actions based on those of the most similar-minded users (collaborative filtering).

- Import of external user-related information
    - ODBC interfaces were developed for integrating external customer and marketing data into the User Modelling Servers.
- Privacy support
    - User Modelling Servers must support any company policies, industry policies, privacy norms and any national and international privacy legislations.

According to Fink and Kobsa's review (Fink and Kobsa, 2000), which examined various commercial server systems, three categories of User Modelling Servers can be identified:

- Systems that implement collaborative filtering:
    - GroupLens
        - Employs various collaborative filtering algorithms for predicting users' interests (Net Perceptions Home Page).
    - LikeMinds
        - Similar to Group Lens. Some differences include a more modular architecture, better load distribution, ODBC support, and slightly different input types (Andromedia Home Page).
- Systems that follow a rule-based approach:
    - FrontMind
        - Provides a rule-based development, management, and simulation environment for personalized information and personalized services on the web (Manna Inc. Home Page).
    - Personalisation Server
        - Allows for the definition of rules that assign individual users to one or more user groups based on their demographic data (e.g., gender and age), information about the user's system usage, and information about the user's software, hardware and network environment (ATG Home Page).
- Systems that adopt the method of hierarchical clustering:

- Learn Sesame
  - Allows for the definition of a domain model consisting of objects, object attributes, and event types. It accepts information about the user from an application, categorizes this information based on the domain model, and tries to detect recurrent patterns, correlations and similarities through incremental clustering (Bowne and Co. Home Page).

After the publication of Kobsa's review of the main User Modelling Servers at that time, where the conclusion was that there was considerable room for improvement in various aspects, a significant development, the Personis User Modelling Server (Kay *et al.*, 2002), came along; with its innovative architecture it provided ways to support powerful and flexible UM and at the same time support user scrutiny and control while placing the emphasis on security and user privacy. The design of Personis is based entirely upon the crucial requirement that users should have access to their user model and control over it, thus special emphasis is given to scrutability. MyPlace Locator (Assad *et al*., 2007), a location modelling system, demonstrates how Personis addresses the needs to ensure the user's privacy and control and their ability to scrutinise their user model.

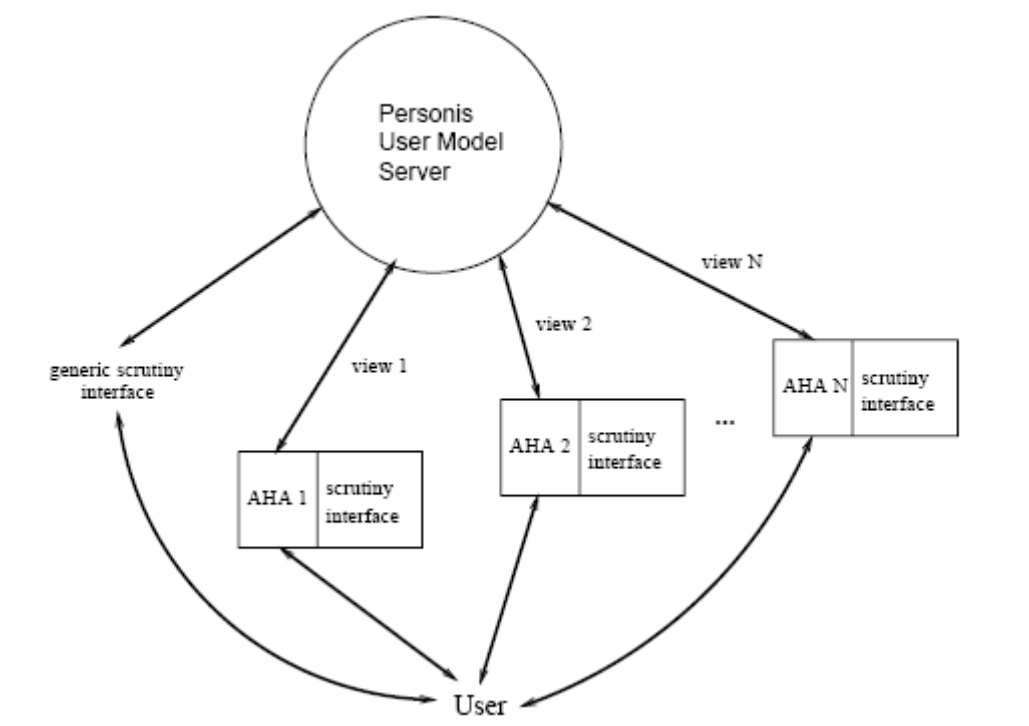

Figure 2-3 Architecture of the Personis Server (Kay *et al.*, 2002)

Figure 2-3 explains how Personis' architecture can support the reuse of a user model over a range of adaptive hypermedia systems. This includes four main components, as shown in the figure:

1. The main server,
2. A list of tools, offered by the um toolkit (Kay, 1995), which allows users to scrutinize and control their own models (marked as scrutiny interfaces)
3. Several adaptive hypermedia systems and
4. The views between the systems that access the user model and the server itself. Note that every system has a different view of a user model, assigned by the server, based on the pre-defined scrutiny settings set by the user/owner.

2.6 Summary

Chapter 2 has summarised a significant theoretical background in the fields of Personalisation Systems and User Modelling by providing a walkthrough of the progress achieved in these two fields throughout the years.

Chapter 3 Three Important Aspects of User Modelling

3.1 Introduction

This chapter introduces and analyses three important aspects of User Modelling, which comprise the central feature of the research presented in this thesis:

- Interoperability, by examining some important UM frameworks and architectures that addressed the issue of overcoming the barriers when exchanging user data across different systems, and the contribution of Standards and Semantic Web Technologies when designing interoperable solutions,
- Scrutability, by analysing the benefit it can bring to UM solutions, and
- User Privacy, by understanding the trade-off between offering personalisation services and implementing privacy policies that respect the preferences of the information owners.

These vital UM components were addressed in the first LUM workshop (Kay and Kummerfeld, 2009b) and in Kay and Kummerfield (2009a) in a discussion regarding the most important challenges that LUM has to address, as will be discussed in more detail in Chapter 4. For the research described in this thesis, although we realise that other important UM aspects like scalability, transparency, trust, standards, reusability, information management, security and representational requirements have to be taken into account as well, these three UM components are identified as crucial when developing LUM solutions and they should be addressed first before tackling other UM

aspects. That is why they are presented and analysed in the following sections of this chapter.

3.2 Interoperability

Interoperability can be considered as "a condition that exists when the distinctions between information systems are not a barrier to accomplishing a task that spans multiple systems" (Aroyo *et al.,* 2006). With the recent introduction of the Semantic Web, the issue of interoperability has become a burning issue in the area of UM. Exchanging user profiles across various sources in a distributed e-learning environment (and others) cannot be achieved unless explicit and widely accepted protocols are developed and adopted that will allow for the description, discovery and exchange of user models that are stored in various systems and implemented in different languages and platforms.

In order to be able to exchange a learner profile between web-based educational systems, we need to provide explicit information about what is going to be exchanged, which values of the specific subject are considered and how the information is bound to the learner (Dolog and Schaefer, 2005b):

- *Learner profile standards and open specifications* provide us with a representation for subjects of exchange, e.g., learner performance, portfolio, preferences, learning style, certificates, evaluations, and assessment.
    - User Modelling Standards most commonly used today, which are elaborated upon in sub-section 3.2.1:
        - IEEE Personal And Private Information (PAPI)
        - IMS learner information package (LIP).
- *Domain ontologies* provide us with exchangeable/sharable models of domains.
    - Learner Ontology:
        - The standards and open specifications guarantee wider acceptance between e-learning systems and, therefore, can be seen as good candidates for the learner-exchange models.
    - Instantiation and Mappings from Internal Models

- The tools, which use a different internal data model and which, it is hoped, could participate in an exchange of learner profiles, have to provide mappings between their internal data model and the exchange model.
- Besides that, evidence about how a learner model was derived should be provided to allow other systems to interpret the model correctly.

Furthermore, in order to exchange fragments of information stored in user models the need for searching and identifying user models' fragments arises. An algorithm has been introduced by Dolog (2004) that allows systems to search and identify relevant fragments of a user's model. The algorithm is briefly described as follows:

- Retrieve all instances of the *identification concept* for the selected user.
- Search instances of the *learner concept* on systems referenced in each identification entry, as retrieved in the previous step.
- If there are further systems referenced in the identification records as remote systems, reapply this algorithm to those systems.
- Retrieve all objects as instances of concepts needed for adaptation as described by the parameters of the call (e.g., learner's interests)

Finally, once discovered, user models' fragments have to be exchanged from one source to another in a consistent and adequate way. One such framework is proposed by Dolog and Schaefer (2005a and 2005b). This framework describes how user profiles can be accessed and imported by various systems:

- Through an extensible Java API that defines classes and properties for retrieving, inserting and updating the user profiles
- Via Web Services where several clients can access one model that is persistent on one server. More information on Web Services is presented at the end of this section.
- Through an RDF querying infrastructure where the user profiles are described in RDF.

The need to introduce a protocol for dealing with consistency while exchanging fragments of information about user profiles between several sources across the Semantic Web had to be addressed at this point and a solution based on Web Services was proposed by Alrifai *et al.* (2006). The proposed protocol, which is presented in

Figure 3-1, defines three operations that User Profile Servers and their interfacing services should use to exchange fragments of information:

- *SendLP* operation: this operation is used to request a Learner Profile (LP) fragment from the server holding this fragment. It is invoked by sending a SOAP message: SendLP message including the LP_ID of the required learner profile and indicating whether the requester wants to register for the notification service about any updates to this learner profile fragment. Upon the receipt of this message, the server sends its local learner information to the requester in a SOAP message.
- *ReceiveLP* operation: this operation is invoked (on the requester's side) by sending a SOAP message: ReceiveLP message including the requested learner fragment. This operation is invoked either as a response to a previous explicit request (through a SendLP message) or upon an update action to the relevant learner information fragment.
- *StopNotification* operation: this operation is used to unsubscribe from the notification service for a particular learner profile and is invoked by a StopNotification message being received.

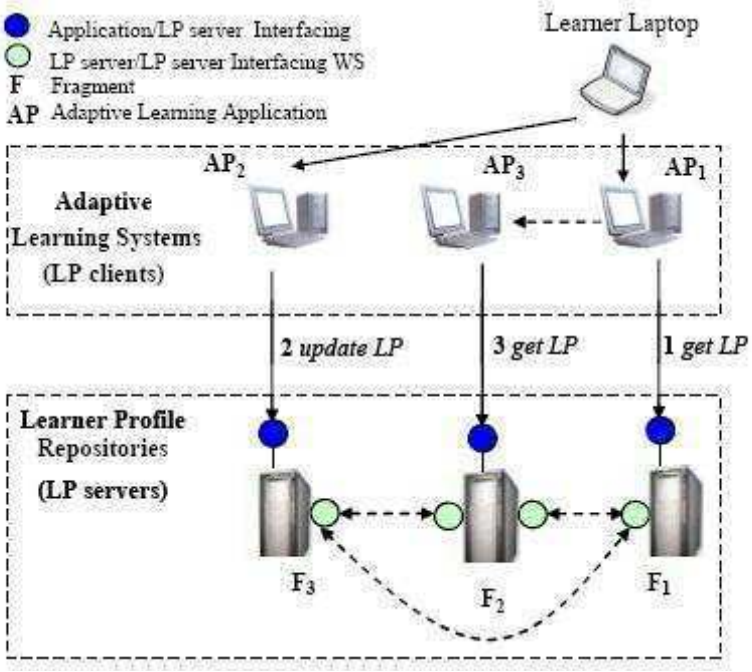

Figure 3-1 Protocol for Exchanging User Profiles (Alrifai *et al.,* 2006)

Finally, two approaches for achieving communication among systems that are favoured by the UM community as explained above, and among the social networking and e-commerce domains that are presented in Chapter 5, are presented below.

Web Services approach: *Web Services*, a set of related standards that enable two computer applications (written in different languages and for different platforms) to communicate and exchange data via the Internet are used to achieve interoperability while exchanging user models' data between one source and another (Deitel *et al*., 2003). In the area of UM, where interoperability is a high priority for enabling the exchange of user model fragments between systems across the web to take place, Web Services, which are structured around an XML-based communication protocol (SOAP) and an XML-based description language (WSDL), appear to solve the problem of platform independency while promoting reusability. Critics of Web services often complain that they are too complex and are based upon large software vendors, rather than on open source implementations. There are also concerns about performance due to Web services' use of XML as a message format, SOAP for enveloping the message and HTTP for transport (W3Schools – Architecture of Web Services).

RESTful approach: Representational State Transfer (REST) refers to a collection of network architecture principles that outline how resources are defined and addressed. Its biggest and most frequently used appliance is the WWW itself. REST assigns a URI to every resource and uses HTTP operations, such as GET, POST, PUT and DELETE, to access and retrieve a value and to update and delete a resource. REST is preferred by many developers today because of its simplicity compared to the alternative Web services, which makes it too attractive to resist. The disadvantage, though, surfaces when sophisticated and/or critical transactions across multiple systems require more advanced solutions than the RESTful approach (IT Professional Blog, 2008).

3.2.1 User Modelling Standards

It is crucial for us to realise that in the UM community, agreement on common structures and the scope of user information modelled is needed (Dolog and Nejdl, 2003). The need for standards was naturally raised and was addressed by two significant organisations, IEEE and IMS, and resulted in two widely accepted UM

standards: PAPI (IEEE PAPI Standard) and LIP (IMS LIP Specification). We present them in Table 3-1 and compare them with Figure 3-2 below:

IEEE PAPI	**IMS LIP**
The *personal category* contains information about the names, contacts and addresses of a learner.	The *identification* category represents demographic and biographic data about a learner.
Relations serve as a category for relationships of a specific learner to other persons.	The *affiliation* category represents information records about membership in professional organizations. The *relationship* category aims to make relationships between core data elements.
Security aims to provide slots for credentials and access rights.	The *security key* is for setting passwords and keys assigned to a learner.
Preference indicates the types of devices and objects that the learner is able to recognize.	The *accessibility* category aims for general accessibility to learner information by means of language capabilities, disabilities, eligibility, and learning preferences.
Performance is for storing information about the measured performance of a learner through learning material (i.e., what does a learner know).	The *transcript* category represents an institutionally-based summary of academic achievements. The *goal* category represents the learning, career and other objectives of a learner. The *QCL* category is used for the identification of qualifications, certifications, and licenses from recognized authorities. The *activity* category can contain any learning related activity in any state of completion.
Portfolio is for accessing previous experience of a user	The *interest* category can be any information describing hobbies and

	recreational activities. The *competency* category serves as a slot for skills, experience and knowledge acquired.

Table 3-1 Comparison of IEEE PAPI and IMS LIP

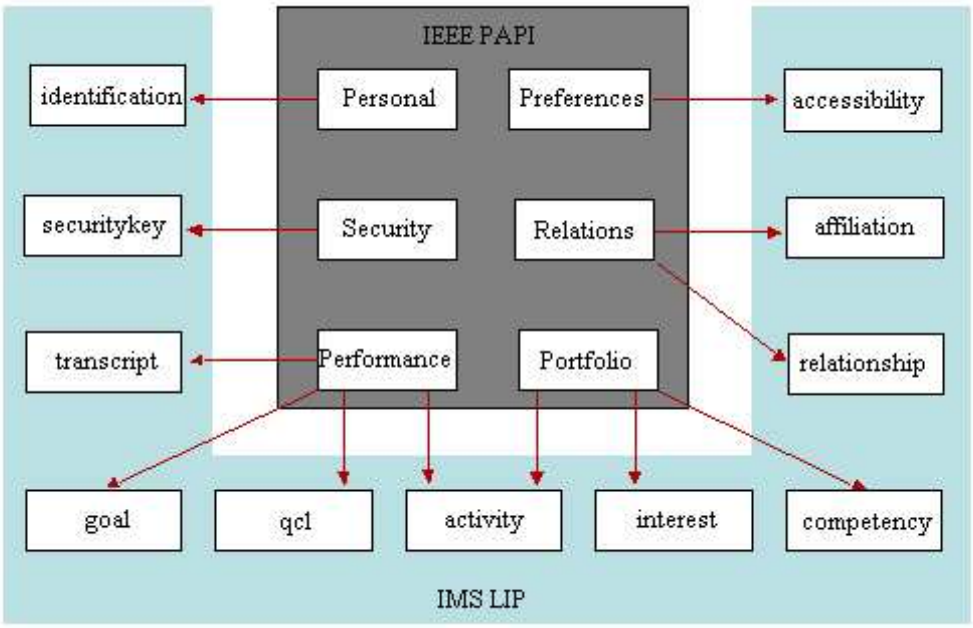

Figure 3-2 Relationship between IEEE PAPI and IMS LIP (IMS LIP Specification)

3.2.2 Semantic Web Technologies in User Modelling

New directions and guidelines for User Modelling have emerged with the introduction of the Semantic Web (Berners-Lee *et al.*, 2001). New technologies have emerged that allow the content of user models to be expressed in a format that can be read and processed by software agents, thus permitting them to find, share and integrate information more easily and efficiently. At the start of section 3.2, we presented some UM solutions that have taken advantage of Semantic Web technologies, in terms of describing user models in RDF and using ontologies written in OWL to describe internal data models and their relationships, for achieving interoperability across several educational systems that are built on different platforms.

The semantic web comprises the standards of XML, XML Schema, RDF, RDF Schema and OWL (Passin, 2004). A brief description of each technology is offered below:

- *XML* (eXtensible Markup Language) provides an elemental syntax for content structure within documents, yet associates no semantics with the meaning of the content contained within (Deitel *et al.*, 2001).
- *XML Schema* is a language for providing and restricting the structure and content of elements contained within XML documents (Deitel *et al.*, 2001).
- *RDF* (Resource Description Framework) is a simple language for expressing data models, which refers to objects (resources) and their relationships. An RDF-based model can be represented in XML syntax (Powers, 2003).
- *RDF Schema* is a vocabulary for describing the properties and classes of RDF-based resources, with semantics for generalized-hierarchies of such properties and classes. RDF Schema is used to represent user models by describing them using its RDF-based vocabulary (Powers, 2003).
- *OWL* (Web Ontology Language) adds more vocabulary for describing properties and classes: among others, relations between classes, cardinality, equality, richer typing of properties, characteristics of properties, and enumerated classes. OWL is often preferred to RDF Schema when describing user models, since its rich vocabulary permits more adequate expressions of the models (Lacy, 2005).

3.3 Scrutability

The term ‘scrutability’ signifies that the model of every user can be scrutinised by the user him/her self to determine what has been modelled about him/her and how that modelling process was conducted (Kay, 2000).

By adopting scrutability in User Modelling, users gain control of their models and this enables them to set their preferences on how the modelling process is applied to them. In addition, users can select in which stereotypes they should be included and which ones not. Furthermore, the users can alter the value of any single inference that is used for drawing conclusions about them (Kay, 2000).

The motivation for implementing scrutable solutions ranges from personal information and its use to user control, self-awareness, approval and acceptability (Kay, 2006):

- There are two motivations relating to personal information and its use:
    - Users have the right to see and appreciate the meaning of the personal information a user model holds about them;
    - The possibility of users correcting errors and/or false assumptions about them in the model.
- There are three rather neglected, but important motivations regarding user control over the personalisation environment and process:
    - To verify the role of the machine (computer) as the servant or aid of the user (human);
    - Accountability of the developers and programmers for the offered personalisation services;
    - Allowing users to feel that they are in control over the adaptation services by controlling the user model, and the way they understand it.
- The following three motivations are associated with self-awareness. In reality, systems, and more specifically user models, have a huge potential to help people reflect their long-term behaviour as they interact with various online applications that store evidence about them:
    - Helping people become more self-aware, since their user models echo their actions and intentions;
    - Educational systems could potentially encourage metacognition and deeper learning;
    - Assisting users to monitor their progress and to plan, especially in a learning environment
- The last two aspects relate to the critical role of making scrutability acceptable to share one's user model for potential benefits. This could even be extended to suggest that user models may have the potential to improve social interaction and collaborative learning by:
    - Motivating people to share their personal user model data because they feel confident about its meaning and use;

- Improving collaboration among team members by enabling them to learn relevant information about each other and help each other more effectively.

3.4 User Privacy

Privacy-Enhanced Personalisation is an area that aims to merge the techniques and goals of User Modelling with privacy considerations and apply the best possible personalisation inside the boundaries set by privacy rules (Kobsa, 2007). As the research in this area shows, there is no ideal solution while attempting to combine these two important elements. Instead, numerous small enhancements must be implemented, depending on the user and application domains in each case, in order to achieve the best possible solution.

One thing can be considered certain regarding user-privacy: users value their privacy privileges extremely highly. Any system, application or service, no matter how successful it may be, that fails to convince the users of its proposed privacy policies will not be successful. Two recent examples, one from the currently biggest social network in the world based on comScore, Facebook (All Headline News, 2008), and one from Netflix, Inc (Ars Technica, 2009b), which is the world's largest online movie rental service, demonstrate exactly how important privacy is to people today:

1. In May 2009, Facebook surprised its users by modifying its privacy policies in secret. The main feature of the change was that any posted information by any user would be considered Facebook's property. The users immediately objected to this change, since it meant that if a user wanted to delete something from Facebook (for example, a picture or a posted video), it would still remain in Facebook's databases and it could be used by Facebook for any reason without any explanation or permission. Some users even filed a lawsuit accusing the social networking giant of violating California's privacy and online privacy laws by circulating private information to third parties for commercial purposes (CNET News, 2009d). In addition, there were significant concerns from the European Commission that Facebook was also violating the European laws for privacy (Gigaom, 2009). Furthermore, Facebook was forced to alter its advertising strategy drastically and shut down its advertising

platform Beacon, due to privacy lawsuits (MediaPost News, 2009; PC Pro, 2009a). Facebook reacted to its users' negative feedback and reversed the change to its privacy rules. In addition, it proposed a completely different policy, which was welcoming users to take part, comment and eventually vote as to which privacy rules should be immediately implemented on the site. Furthermore, Mark Zuckerberg, the creator of Facebook, commented that users would also participate meaningfully in any proposed future privacy policies (InformationWeek, 2009a). In addition, based on a statement released by the Privacy Commissioner Jennifer Stoddart, Facebook has agreed to retrofit its privacy terms to explain better why it collects personal data, make clearer the distinction between *deactivating* and *deleting* an account, and provide consent to have profiles "memorialized" after death (The Register, 2009).

2. In September 2009, Netflix awarded a cheque of 1 million United States dollars to the winners of Netflix's first competition, which lasted for nearly three years. Its objective was to improve the company's recommendation algorithm by 10% (PC World, 2009d). Immediately after the award of the prize to the winning team, Netflix announced a second competition, which caused significant objections to be raised among the Netflix community (PC World, 2009e). Specifically, for the second competition, the company was willing to make available even more user data that would assist participating teams to improve the company's recommendation techniques for predicting movie enjoyment by members who did not rate movies often. This included, among other information, the users' gender, ZIP codes and ages. As experts revealed, research in the area has proven that most American citizens can be identified when ZIP code, gender and birthday are combined (Ars Technica, 2009a). Netflix was releasing both gender and ZIP codes, but not birthdays, although releasing ages would make it very easy to deduce someone's birthday. Consequently, Netflix users were very concerned with the company's approach and the criteria for its second competition. These concerns and objections were expressed in public and experts warned the company that since they had informed the company of the potential privacy implications of its data releases, it could not claim ignorance and this move would be

considered as being liable for damages (Ars Technica, 2009a). Although Netflix assured its users that the released data would be anonymised, it did not convince the public and currently might be facing paying millions of dollars in lawsuit damages.

As has already been stated, we acknowledge the claim that there is no standard trade-off between privacy of information and use of personal information for personalisation purposes (Kobsa, 2007). For the purposes of this study, we consider the following issues to be the most important considerations while attempting to model a wide variety of users:

- Informing the users about the process of gathering their information and explaining to them the potential benefits.
- Allowing users to know how their data are stored and processed so that conclusions about them can be drawn
- Acquiring users' approval when their data are being moved from one system to another.

In addition, we acknowledge the claim, described in Thang *et al*. (2007) and Wang and Vassileva (2009), that central authentication and privacy control mechanisms have to be implemented when mashing up user data; this would mean retrieving user data from several sources and merging them to generate user information from various social networking sites in order to produce UM data sets. Users should be able to log in and express their privacy preferences in one place while the developer handles all the technical challenges.

3.5 Summary

Chapter 3 has presented three important aspects of User Modelling where we lay the basis for the work presented in this thesis:

- Interoperability for exchanging user models' fragments across various online services.
    - Standards to introduce consistency while modelling a large number of users
    - Semantic Web technologies for expressing the content of user models in a machine-readable and process-able format.

- Scrutability in order to pass the control over every user model to its owner.
- User privacy to ensure the protection of personal user information.

Chapter 4 Lifelong User Modelling

4.1 Introduction

Lifelong User Modelling (LUM) can play a critical role in enabling lifelong personalized learning to take place. Although the LUM vision is technically still not possible, due to the fact that we cannot yet identify and harvest every piece of information that we have posted and shared on various services on the WWW, it is the way forward, as Semantic Web technologies and LUM projects, like the Linked Data and Data Portability, have provided us with the tools potentially to achieve this vision. Various challenges have to be addressed, especially in terms of sharing user models effectively with educational systems and between people, while allowing the owners of these models to scrutinize the modelling process (Kay and Kummerfeld, 2009a). This chapter highlights the questions that have just begun to be addressed in the user modelling community while emphasising the most important, according to the latest LUM workshop (Kay and Kummerfeld, 2009b), requirements for making the LUM vision a reality. In addition, it acknowledges related work that establishes a base for the area of LUM and triggers further analysis on some important issues.

The reason we reviewed the work around the area of Lifelong User Modelling was because our research focus addresses modelling social networking and e-commerce user data, which we believe can be considered as life-long user data due to excessive and continuous engagement with people over long periods of time. This will be dealt with in greater detail in Chapter 5.

4.2 Why Lifelong User Modelling?

The importance of lifelong user modelling serving the lifelong learning vision has been acknowledged by many associations and committees, such as the Computing Research Association (CRA) (Computing Research Association Home Page) and the United Kingdom Computing Research Committee (UK Computing Research Committee Home Page).

In addition, various projects have allocated resources and funding to address the various challenges and demands of this area. An excellent example is the Lifelong User Modelling project at the University of Sydney (CHAI LUM Project) which its contribution is elaborated further in this section:

People post and share personal information on the web using several online tools, resulting in the creation of personal user model clouds. Every cloud contains user data, posted by the user or generated based on the user's interaction with the application; it is a subset of a user's complete model, since these clouds may be distributed across multiple machines (Dolog *et al*., 2009). During a symposium held by Gartner Inc., an information technology research and advisory firm, cloud computing was placed at the top of the list for the 10 top technologies for 2010 for which information technology companies need to prepare and plan. Although it is complicated to implement, and poses security risks, cloud computing should be taken seriously and companies should identify those cloud services that suit them best and might give them value when used within each company's specialised context (CNET News, 2009b).

To demonstrate the issues that motivate the need to address personal user model clouds we present the following example, as described in Dolog *et al.* (2009):

"Alice wants to become a software engineer because she has had a long interest in computers. She studied the most challenging computing subjects available at high school and is now nearing completion of a degree in Software Engineering. Several programming subjects used a system call Reflect: this accepts assignment submissions, automatically grading them, provides code reading tasks and it also regularly asks her to assess her own knowledge of concepts. It has an open learner model, used for personalisation and to support reflection. Each subject's Reflect instance maintains its own learner model. This is integrated into the School learner modelling system and to the personal learner models of each student at their preferred machines. The same approach operates for other personalized teaching systems, such as LOGIC-ITA which

teaches logic. Similarly, the School integrates information from the LMS, including subject examination results. The School regularly uses the learner models to review the success of its teaching and to identify areas for improvement to student learning.

Alice happens to meet Bob at the cafeteria and they realise they are doing the same capstone software project. They decide to discuss whether they should be in the same team. Alice uses her smart phone to review the parts of her learner model from the four years of her programming studies. She shares a summary view of this with Bob, to convince him that she could take the programming expert role in the group. He shares part of his user model to show he could be the algorithms expert. They decide to work together and send these user model views to other potential team members."

Figure 4-1 presents a very high-level architectural overview of the illustrated example. Note that there are several sources that hold user models about Alice and Bob, which most probably adopt different privacy policies, and that these models can be accessed via various machines.

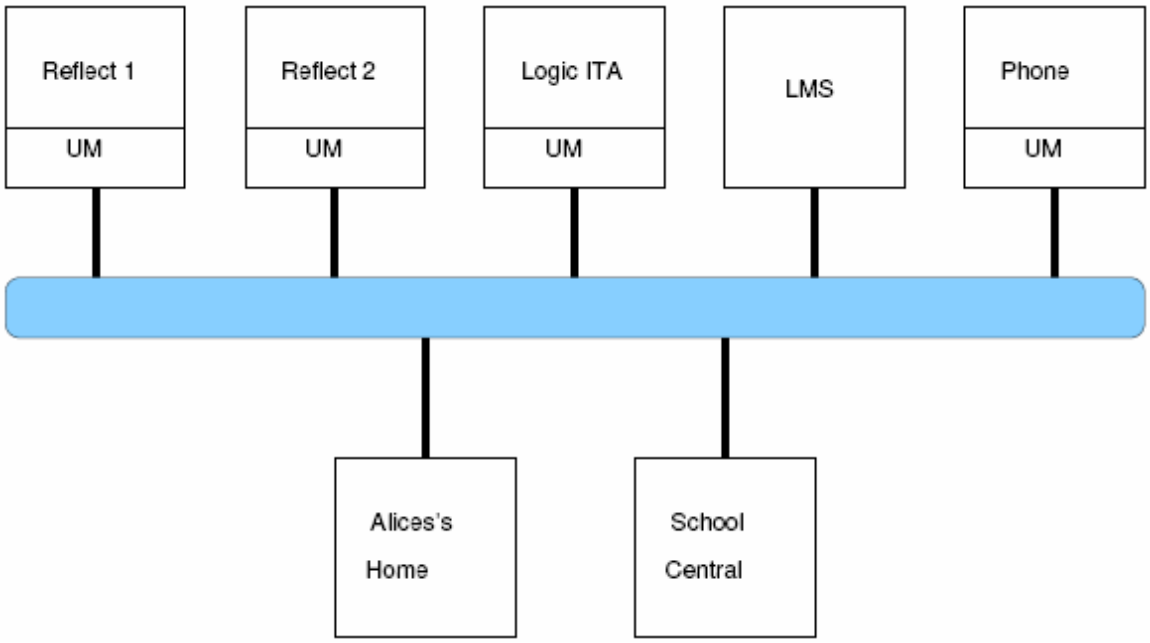

Figure 4-1 Architectural Overview of Personal UM Clouds (Dolog *et al.,* 2009)

While the tools for allowing us to exchange personal data are progressing constantly (for more detail, see Chapter 5) and the storage of large amounts of digital information is no longer an issue, there is a need for the community to find ways to enable users to keep archives of their personal user model clouds. This information is

distributed across various machines, which are built on different platforms, and based on a range of policies and contexts, thus making it extremely difficult for the owners of this information to keep a record of its whereabouts along with its original purpose and value.

On the other hand, these increasingly large personal information collections can create new possibilities for lifelong user modelling and personalisation. Lifelong visions require user models that are not restricted to a single application nor are stored in a single session. Assuming we find a way to harvest this sparse data and create lifelong user models, it will provide the foundation for lifelong personalised services that (ideally) will allow people to make effective use of their lifelong personal digital information (Kay and Kummerfield, 2009a) as follows:

- It will allow people to plan and track long-term goals while allowing the reuse of these models by different applications.
- In addition, it will assist with long-term self-monitoring, planning and reflection while passing full control to the owners of these models.
- Finally, it will support personal information management, which will be based on trust between the models' owners and LUM services.

4.3 Challenges for Achieving the LUM Vision

Although LUM can result in significant benefits, various challenges and demands have to be addressed and (at least partially) resolved, before these potential benefits can be realised. These challenges/demands were discussed during the Lifelong User Modelling workshop conducted in June 2009 at the 1st UMAP conference (Kay and Kummerfeld, 2009b): The participants kept one target-goal in mind while creating this list, which was to capture aspects of a user's life, spanning very long periods of time:

a) What are the requirements that would enable lifelong user models to be useful to a range of applications? Clearly, interoperability is the first aspect that needs to be considered before addressing further issues. Gathering the sparse information, organising it and distributing it to a range of systems is probably the most challenging step. Adequate solutions have been presented in section 3.2, but further and more specialised research is needed to address LUM challenges, such as ensuring consistency when retrieving or updating distributed user models

in Web 2.0 environments, especially from social networking and e-commerce services.

b) How can we ensure users will have control and share their lifelong models effectively? As described in section 3.3, scrutable solutions allow users to inspect and alter any inference that is used for drawing conclusions about them; therefore scrutability can be considered a good method for tackling this issue.

c) What privacy aspects need to be considered when designing technical solutions for LUM? Section 3.4 addressed the important issue of privacy-enhanced personalisation. This, of course, extends to LUM, and the challenge arises when attempting to incorporate models from different systems/sources that are implementing different privacy policies. Although it is hard to agree on and implement a universal solution, various small enhancements can serve both users and providers of user models adequately.

d) What are the relevant existing standards that could be adopted in LUM? Is there a need for additional standards? Section 3.2.1 elaborated on two widely adopted standards: PAPI and LIP. This is clearly an important element in LUM and our research has explored the use of APIs in the form of standards in various constantly progressing Web 2.0 environments, more specifically in the social networking and e-commerce fields.

e) How can semantic web technologies support LUM? Section 3.2.2 provided an overview of the various Semantic Web technologies that could assist LUM in several aspects. In particular, ontologies are a reasonable approach for mapping the relationships between the different data types of the various systems/applications/services.

f) What are the representational requirements for lifelong user modelling? While this is mostly an HCI issue, our research has not focused on this aspect, but we gathered some interesting results during the two evaluations we conducted (for more detail, see Chapters 8 and 9).

Another interesting challenge, which has been addressed by Jameson and Gabrielli (2009), is the study of potential factors that can lead to changes in users' preferences over (long) periods of time. This is, of course, a theoretical overview, based

on experiences concerning the interactions of users and existing adaptive systems, but the accurate conclusion is that users' preferences are most likely to undergo considerable changes over time; hence, LUM has to keep track of these changes to produce valid conclusions regarding users' behaviour.

4.4 Related Work

The related work presented in this section, although it does not explicitly (or fully) relate to the area of LUM, makes a significant contribution to several aspects of LUM and Lifelong Learning.

4.4.1 GRAPPLE Project

The following is taken from GRAPPLE's annual summary: February 2008-February 2009 (GRAPPLE Project Home Page). GRAPPLE is an EU FP7 STREP Project, intended to run from February 2008 until February 2011.

The Generic Responsive Adaptive Personalized Learning Environment (GRAPPLE) project is targeted at taking advantage of users' life-long learning experiences and adapting a technology-enhanced learning environment to their personal preferences, prior knowledge, skills and competences, learning goals and the context in which the learning takes place. The offered environment can be accessed and used anywhere (at home, in school, at work, etc.) and with any device (desktops, laptops, mobile devices, etc.). Furthermore, GRAPPLE provides authoring tools to educators, including adaptive interactive components (visualizations, simulations, virtual reality), to create the learning material, such as designing learning activities, creating or importing teaching resources, and defining adaptation rules for the content and activities.

A distributed user modelling architecture has been adopted that aims at helping users to gain control of the user profiles while at the same time enabling them to obtain personalised access to the various learning applications offered via different LMSs (Van der Sluijs and Hover, 2009).

A first version of the GRAPPLE adaptation engine was built for initial experimentation and it was based on the idea behind the Adaptive Hypermedia Architecture (AHA!). The new adaptation engine, called GRAPPLE Adaptive Learning

Environment (GALE) is a modular and extensive adaptation engine that provides and serves the authored adaptive courses. Its main components are a user modelling service, a domain/adaptation model service and an adaptation engine. These three components communicate with each other through an event bus that is also accessible by other GRAPPLE components (Van der Sluijs *et al.*, 2009).

User information is located mostly in two places inside GRAPPLE: GALE holds the information needed to perform the adaptation and LMSs store personal information, course registration, grades, and other similar information. However, as learners increase their learning experiences, this could result in them having information about themselves and their learning progress stored in several places.

GRAPPLE follows an ontological approach for defining its user model format. A GRAPPLE User Modelling Framework (GUMF), which is based on the Personal Reader Framework (Personal Reader Home Page), permits the retrieval of user information from different sources of which the architecture has been defined (Abel *et al.*, 2009a; Leonardi *et al.*, 2009). GUMF cooperates with GALE and the various LMSs to determine the identification of every source. GUMF has as main enhancements the extensible user modelling ontology format, extensible query interfaces using RESTful, SOAP-based and RSS-based approaches, and a community-based way of sharing and ranking user models. In addition, GUMF can deduct user information by mashing up different (user profile) data streams in RDF using Semantic Web Pipes or Yahoo Pipes. Administrative interfaces exist for managing the GUMF configuration and for manually exploring the resulting user data streams, reasoning mechanisms and ontology extensions. The GRAPPLE User Modelling Ontology, which mainly adopts the approach of General User Model Ontology (GUMO) (Heckmann, 2009), specifies the lingua franca for exchanging user model information in the form of subject-predicate-object statements.

4.4.2 Linked Data Initiative

The following is taken from the official website of the Linked Data Initiative (Linked Data Home Page). The aim of this initiative is to connect distributive data across the Web and it provides the tools for creating and publishing Linked Data in order to produce meaningful information from pieces of distributed data, an approach

that permits (lifelong) reuse by its owners and that has been adopted in our proposed solution and is presented in Chapter 7.

This initiative is concerned with linking data on the Web that were not previously linked, or altering the way that data are currently linked with alternative, better methods. More specifically, Linked Data can be defined as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF" (Berners-Lee, 2006).

Using the Web as a medium for connecting related data is a realistic target if we consider that we currently use the Web to link related documents. The Linked Data initiative has set as a goal the identification of the best practices for publishing, connecting and structuring data on the Web. Key technologies that can help towards achieving this vision are URIs (a generic way for describing and identifying concepts in the Web), HTTP (a simple but effective mechanism for retrieving resources on the Web) and RDF (a standard way of describing and structuring resources on the Web).

In a recent interview, the inventor of the Semantic Web, Sir Tim Berners-Lee, explained that Linked Data actually uses the technologies that make up the Semantic Web, like OWL and SPARQL, and acknowledged the fact that the WorldWideWeb Consortium (W3C Home Page) has been emphasizing the importance of Linked Data to encourage government departments to convert their data into Linked Data (ReadWriteWeb, 2009b).

Several tools are currently available for creating, publishing and discovering Linked Data on the Web. Among them are the following:

- *sqlREST*: sqlREST exposes relational databases as a REST-style Web Service. Using HTTP and XML, user data can be retrieved, removed and altered, while web browsers like Mozilla or MS Internet Explorer can be used to access the data (sqlREST Home Page).
- *D2R Server*: D2R Server is a tool for publishing relational databases on the Semantic Web. It enables RDF and HTML browsers to navigate the content of the database, and allows applications to query the database using the SPARQL query language (D2R Server Home Page).

These are tools for publishing relational databases as Linked Data: Paget (Paget Home Page), which is a framework for building Linked Data applications; Semantic Web Browsers, like Zitgist (Zitgist Home Page); Tabulator (Tabulator Home Page), which

allows you to explore unbounded sets of RDF data sources on the Web; Semantic Search Engines, like Hakia and SenseBot (Pandia Search Engine News), which attempt to make sense of search results based on their context through natural language processing; and Vapour (Vabour Home Page), which is a Linked Data validator.

4.4.3 Semantic Data Integration

Beyond Linked Data there is a need to embed context to web-based user data and relate this context with interoperable well-accepted ontologies across the various communities. In addition, better user interfaces can assist on providing quality meta-data which will potentially improve rule-based assertions performed on the data for generating semantic meaning of the relationships among them. These are issues that extend well beyond the techniques of linked data and form the next set of challenges in gaining broader acceptance for the Semantic Web, which has introduced a vision where computers can understand enough about what the data means in order to process it (Bergman, 2009).

To address this issue the recognition of three types of heterogeneities is required:

- *Syntactic*: being able to handle different data models and formats.
- *Schematic*: being able to understand different data schemata.
- *Semantic*: being able to unify the different data models, formats and schemata by mapping them using specific criteria and constraints.

Located on the higher level of the interoperability scale, semantic data integration allows machines to understand the semantics of the data during exchange across different systems. That can be tackled using various approaches, such as ontology-driven by mapping concepts and entities in a widely-accepted and used ontology, and rule-based by explicitly defining the relationship between concepts and entities using logic rules (Cruz, 2003).

4.4.4 DataPortability Project

The following was taken from the official website of the DataPortability Project (DataPortability Project Home Page). The project's mission is to help people to use and protect the data they create on networked services, and to advocate for compliance with

the values of DataPortability. It was formed in November 2007 and is still in the early stages of development.

Data portability is the term used to describe users' ability to reuse their personal data across different applications, which is the main focus of this work: to assist users to gather their personal data from social networking and e-commerce services and (re)use them for educational purposes. The Data Portability Project works towards achieving the vision of identifying, describing and exchanging resources across interoperable sources. Its goal has been summarised in four words: connect-control-share-remix. More specifically, the vision of this project is to make the data portability experience a reality where people can overcome the barriers of moving between different services across the Web, reusing their data in several contexts while controlling the whole experience and placing special emphasis on safeguarding the privacy of their information.

With Data Portability, users can take with them their identity, list of friends, personal files and browsing histories without having to add them manually to every new service. Then, each service can take advantage of only the user information that is relevant to its context. While users continue to have new experiences, resulting in new or updated data, the information inside the participating services will be updated automatically, with the permission of the user/owner, without the user/owner having to re-enter it or manually update it.

4.4.5 Open Authorization Protocols

Open Authorization protocols are essential when developing central authentication mechanisms in UM solutions that aim to harvest user data mashups from several social networking and/or e-commerce services. Next, we present two widely used protocols in social networking and one well-known in e-learning that allow users to log in to different systems with the same authorization credentials.

OpenID: OpenID is a safe, fast and easy solution to authenticating users in participating websites. It is currently adopted by Google, Facebook, MySpace, Flickr, BBC, PayPal and others (ChannelWeb Network, 2009b). OpenID uses an open decentralized standard to allow users to log on to different services, which trust the authentication mechanism used by OpenID, with the same digital credentials. This revolutionary method replaces the common login process where users were required to

have different usernames and passwords for every service in order to acquire a one-time access to the service. The unique identification that OpenID offers to its users is in the form of a URL that can be used to gain access to the various resources offered by multiple systems (OpenID Home Page).

OAuth: An alternative to OpenID, OAuth was introduced in November 2006; it is an open protocol for permitting secure API user authorization to web applications using a simple but effective mechanism. In contrast with OpenID, which uses a single identity to allow access into many websites, OAuth provides a standard way for websites to offer their services, via an API, without asking users to expose their login credentials. For users, OAuth grants developers secured access to their sensitive personal data. For developers, OAuth gives users access to their services while protecting their account credentials (OAuth Home Page).

Shibboleth: The Shibboleth System is a standard based, open source software package for web single sign-on across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner. The Shibboleth software implements widely used identity standards to provide a federated single sign-on and attribute exchange framework. Shibboleth also provides extended privacy functionality allowing users to control the attributes released to each application. Using Shibboleth-enabled access simplifies management of identity and permissions for organizations supporting users and applications. A user authenticates with his or her organizational credentials. The organization (or identity provider) passes the minimal identity information necessary to the service manager to enable an authorization decision. Shibboleth leverages the organization's identity and access management system, so that the individual's relationship with the institution determines access rights to services that are hosted both on- and off-campus (Shibboleth Home Page).

4.4.6 Web Data Mashup Tools

Web Data Mashup tools and, more specifically, pipes, can assist in producing streams of data that are located on different systems, and can use those streams for presentation purposes. For example, a UM system can take advantage of Yahoo pipes, as presented below, to merge real-time user data from various social networking and e-commerce websites and present the outcome to its owner inside the UM system.

Yahoo Pipes: Pipes is a powerful tool that permits the aggregation, composition and manipulation (also known as mashup) of content from around the web. Adopting the idea of Linux pipes, Yahoo Pipes introduces simple commands that can be combined and so produce outputs that meet any user's needs. The output format can vary, among others, from RSS to JSON. It also provides a visual editor that assists users while creating their own web-data mashups. Furthermore, users can also publish their individual pipes without having any familiarity with programming commands (Yahoo Pipes Home Page).

Semantic Web (DERI) Pipes: Inspired by Yahoo Pipes, Semantic Web Pipes was introduced by the Digital Enterprise Research Institute (DERI Home Page) and its aim is to permit data interoperability and reuse. As with Yahoo Pipes, it also provides a graphical environment where users can perform their own web-data mashups and produce outputs in RDF, XML or JSON formats, which are calculated in real-time when accessed via a pre-assigned URL (DERI Pipes Home Page).

4.5 Summary

In this chapter, we have introduced the area of Lifelong User Modelling, which is a sub-area of User Modelling and is the focus area for this research work. LUM examines the challenges and potential benefits of producing long-term, even lifelong, user models. The arising challenges have been presented and related to the important aspects of UM that were described in Chapter 3. In addition, we have identified potential benefits, which have been discussed thoroughly among the LUM community (Kay and Kummerfeld, 2009b). Finally, we have presented significant related work around the area, which provides a more in-depth and clearer view of the current state of LUM.

Chapter 5 Social Networking and e-Commerce in User Modelling

5.1 Introduction

In this chapter, we will examine two very big and potentially significant contributors to the area of UM: the social networking and e-commerce domains. We have conducted extended research into these two huge sources of user information, identifying which sites have emerged in recent years, how powerful these sites are in terms of user engagement, what value they can bring to the area of LUM when harvesting the large amount of user activity observed on these sites, and especially where they stand and what they offer regarding ownership of the user information that is constantly posted and shared on their platforms.

The reason we reviewed these two domains was because of our research claim that user data from social networking and e-commerce services can bring value to the educational field by enriching user data sets that are currently used by educational systems, like Adaptive Hypermedia Systems, for generating personalisation services.

5.2 Social Networking Domain

In the online world social networking is the term used to describe the way that users build online networks of contacts and interact with these personal or business friends in a common and shared environment. It is probably the most advanced domain in recent years, with tremendous progress being made after the introduction of Web 2.0. In March 2009, a study released by Nielsen Online (Reuters, 2009) revealed that online social networking accounted for 10% of people's time spent on the Internet,

which was more than they spent on email. Furthermore, the same study discovered that one in every eleven minutes spent online globally, was on a social networking site, which means, as supported in the study, in the year 2008 alone, the time spent on these sites increased by 63%. In addition, a recent ethnographical study of more than 800 teenagers and parents in a three-year US project has revealed that participants felt more at home and engaged faster in online social networks compared to the usual public places, like shopping malls, parks and streets (BBC News, 2009c). Moreover based on a report published by Pew Internet & American Life Project, it has been identified that adults are joining social networking sites as well, which supports the case that social networking is a widely-accepted phenomenon (BusinessWeek, 2009). Furthermore, two incidents that have been reported during the last year, specifically, a missing girl's father using Twitter to spread the world of her disappearance and ask for anyone who may have some information to come forward (CNET News, 2008) and, in Australia, two girls who turned to Facebook to ask for help when they were trapped in a storm drain (CNET News, 2009c), revealed that people have turned to social networking sites while seeking help. This proves that we have entered a new era where social networking is playing a major role in our lives.

The list of social network sites is growing every year. Currently, there is not an accurate number of how many such sites exist on the WWW, but a list of the top ones has been debated over the years as presented by a study from comScore (ChannelWeb Network, 2008). For this research, we have focused on the currently biggest and most representative social network site in the world, Facebook, and on the biggest social networking platform in the world, Google's OpenSocial, which has been adopted by, among others, the second biggest social networking site: MySpace. These two representatives have raised the bar extremely high with their distinct progress, and set the standards for future newcomers in this domain; thus, by choosing these services for our case study, we aim to test the part of our claim that a solution that applies to current social networking services is possible.

5.2.1 Facebook

Facebook is the world's largest social network with over 250 million users (Facebook Company Profile). It was founded in 2004 and its initial cause was to serve as an exclusive social network for Harvard students. It did not take long to become a

huge success, and it was expanded within four weeks to serve 30 more colleges. Currently, it holds a 43.84% market share in the top ten social networking websites and forums in the US, which is 29.95% of the global audience (DreamGrow Digital, 2009), and it is the dominant social networking website in many countries (see Figure 5.6). In June 2008 based on comScore, Facebook became the biggest social network site in the world with 132 million users (All Headline News, 2008), dislodging the previous social networking king, MySpace.

The Facebook platform consists of four components of which developers can take advantage for implementing applications that can work inside the Facebook platform only (Facebook Developers Home Page):

- **Facebook (RESTful) API:** The Facebook API is a programming interface, which is based on the REST architecture, for accessing core Facebook services, such as users' profiles, friends lists, photos, joined groups, etc., and performing other Facebook-centric functionalities like logging in, redirecting, updating views, etc. Facebook was the first social e-network service to provide an API for retrieving user information from its platform. In our opinion, it is probably the most significant turning point in the history of social networking because it is the first time users were allowed to have some (limited) control over their own personal information that was posted on a social e-network. It is the main reason for Facebook's huge growth over the past years and it plays a crucial role in this research work.
- **Facebook Markup Language (FBML):** FBML is an HTML-like language used to display pages inside the Facebook canvas.
- **Facebook Query Language (FQL):** FQL is an SQL-based interface into Facebook data. It permits access to many Facebook database tables.
- **Facebook JavaScript:** It permits limited scripting functionality inside a Facebook application.

Figure 5-1 shows the architecture of an application for the Facebook platform (bottom diagram) compared to a standard web application (top diagram):

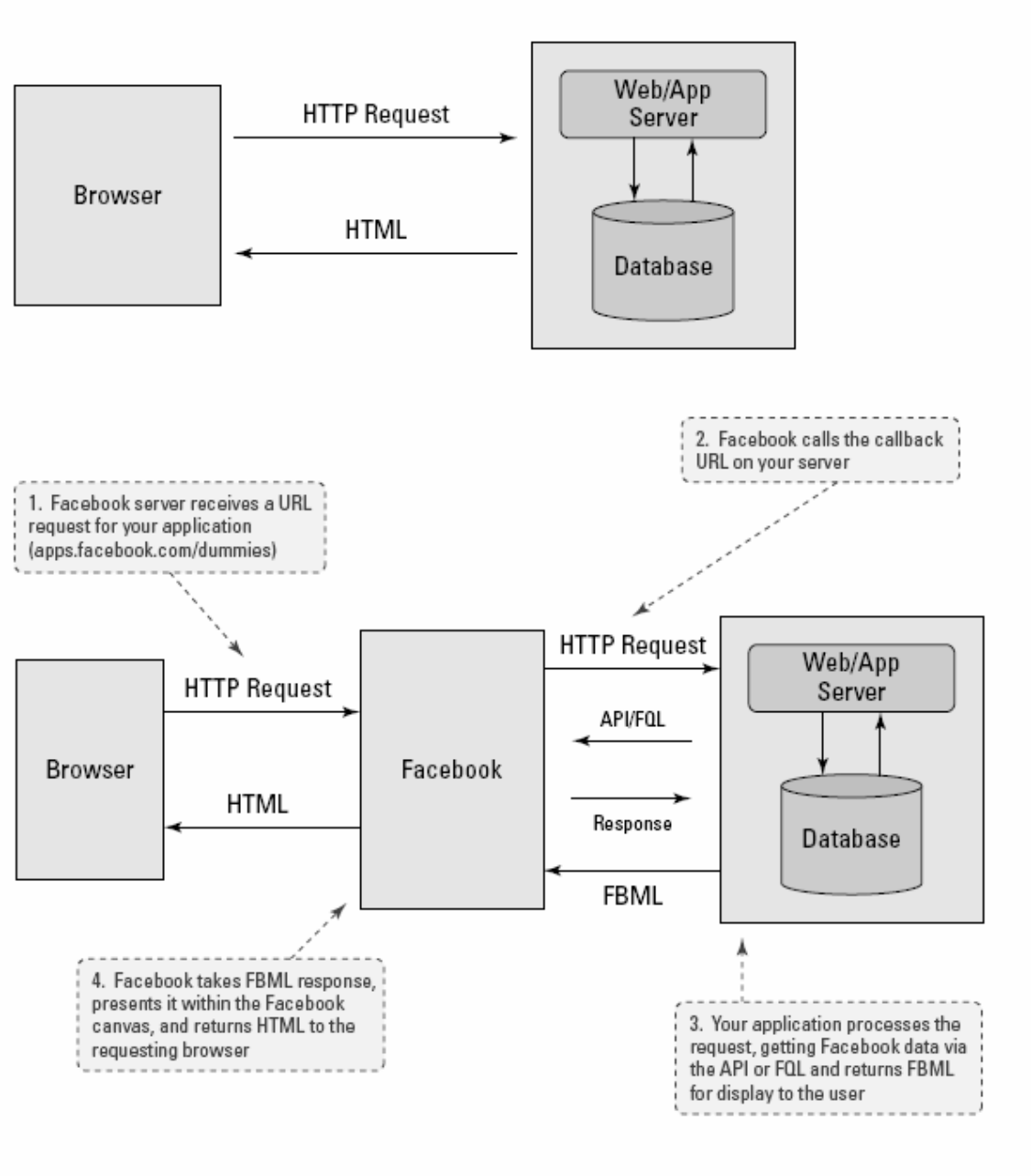

Figure 5-1 Standard web application - top VS Facebook application architecture – bottom (Facebook Developers Home Page)

In addition to the API for building applications inside the Facebook platform, Facebook also announced its data portability initiative in July 2008 and released it in December 2008 (Facebook Developers News, 2008). Facebook Connect is an extension of the Facebook platform that allows Facebook users to "bring their Facebook account information, friends and privacy to any third party website, desktop application or device" (PC Magazine, (2008a). This means that users can sign in to third-party sites with their Facebook ID and bring profile information, profile picture, name, friends,

photos, events, and groups to the new site (Facebook Connect Home Page). Figure 5-2 demonstrates the authorization screen when a Facebook user attempts to connect his Facebook account to an external website. Once the authorisation has been completed successfully, every call from the website to the Facebook platform and vice versa can occur using the regular RESTful API calls.

Figure 5-2 Facebook Connect Authorization screen/message (Facebook Developers Home Page)

To summarize, Facebook has offered:

- an API to developers for implementing applications that only work inside the Facebook platform and
- Facebook Connect, a way for users to bring their personal information that was posted on Facebook to third-party websites.

5.2.2 OpenSocial

Launched in November 2007, Google's OpenSocial is a platform that offers a common API that application developers can use to implement applications that work on any participating social network (TechCrunch, 2007). Instead of introducing another social network site, Google attempted to introduce a "network of social networks" by providing a common tool to develop social applications. The offered API allowed developers and programmers to retrieve user information, along with their friends' information, in a standardized way. Soon, it was adopted by many websites, such as MySpace, Hi5, Orkut, LinkedIn, Plaxo and Friendster, etc. A full list of all websites that currently use the OpenSocial standard can be found at the application's home page

(OpenSocial Home Page) along with the market share each of these websites is holding in the countries where they have the most traffic, according to Alexa.com.

OpenSocial was a reaction from Google to Facebook's move to release an API for developers. It was introduced at a time where developers were complaining about the cost in time and money of learning every single social networking platform in order to adjust their applications accordingly. With this approach, a new stage has been set in social networking where a common platform can be used to obtain distribution across a variety of social networking sites.

The advantage this has over Facebook is that OpenSocial is a common API for a number of social e-networks whereas the Facebook API is applied only on the Facebook platform. Developers can use the OpenSocial API to build an application that could work on multiple platforms whereas, with the Facebook API, their applications work on only one platform. Of course, by the time Google announced its own API, developers had already designed and built their applications using the Facebook API. Figure 5-3 presents the vision of Google's OpenSocial, having a common API for building applications for any social network accessed by any standalone or web-based client.

Figure 5-3 The OpenSocial vision (OpenSocial Home Page)

The OpenSocial API is provided in two versions:

- **A JavaScript API:** This is the main and most used version of the OpenSocial API. It allows users to retrieve and update personal information from social e-networks that adopt OpenSocial, via applications that work inside these participating platforms. There are four main feature areas in this API (OpenSocial API Home Page):
    - **People and Relationships:** OpenSocial applications can use the connections between people and their friends.

- **Persistence:** Applications can take advantage of persistence, the ability to store data that can be retrieved when the application runs again at a later time.
- **Activities:** People use social applications to inform others about what they're doing: going to a movie, posting photos, and so on. An activity is an action performed by a user on the social network.
- **Messaging:** Almost all social e-networks offer a way of reading, posting, and deleting messages between users in the network. OpenSocial defines a number of message types including public messages (such as profile comments) and private messages (messages restricted to certain individuals and groups).

- **A RESTful API:** The OpenSocial model also includes a RESTful version of the API that developers can leverage for server-to-server communication. This version allows social applications to run on desktops, mobile devices, and other channels that do not support JavaScript. The REST API provides five main types of data items that can be retrieved via RESTful data calls (OpenSocial REST API Home Page):
  - **Person:** This category returns information about a person. A collection of these items represents a group of people, such as a given user's friends.
  - **Activity:** This category returns information about an activity performed by a user within the social network. A collection of these items represents an activity stream, that is, a list of the activities the user has performed.
  - **Group:** This category supports enquiring for the available groups for a given user.
  - **Album:** Albums support collections of media items posted by a user. This includes any video, image and sound file shared by the user.
  - **AppData:** This is an item of data, such as a preference setting, stored by the social network on a user's behalf for a given application. A collection of these items might be the whole set of user preferences for a Facebook application.

In addition to the OpenSocial API, Google also released a data portability program in December 2008, called Friend Connect (Friend Connect Home Page). Its mission, similar to that of Facebook Connect, was to extend social networking capabilities across the Web. In general, this meant that people could login using their Google credentials and they would have their social networks accompanying them to third-party websites that implemented Friend Connect. For developers of social applications, Friend Connect enables them to integrate social features to external websites directly from participating social e-networks that adopt the OpenSocial specification. In essence, Friend Connect is an extension of the OpenSocial API, since once it validates the requesting user, all calls from the external website to the social e-network and back, are made using the OpenSocial API, either the JavaScript or the RESTful version.

Figure 5-4 provides a graphical demonstration of the advantage of using Friend Connect in a third-party website:

Figure 5-4 Incorporating Friend Connect into a third-party website (Friend Connect Home Page)

To summarize, Google is currently offering:

- A JavaScript and a RESTful version of the OpenSocial API to developers to implement applications that work only inside participating social e-networks
- Friend Connect, a data portability project that allows users to bring social features from social e-networks that adopt the OpenSocial standard to third-party websites.

5.2.3 Facebook vs OpenSocial

Table 5-1 summarizes the collection of available Facebook and OpenSocial APIs that allow users to have (some) control over their personal data posted on several social e-networks:

API	Description
Facebook RESTful API	Method calls are made over the internet by sending HTTP GET or POST requests to the Facebook API REST server.
Facebook Connect API	This API is an extension of the RESTful API. It first validates the user and then permits the integration of the Facebook platform into the third-party website
OpenSocial JavaScript API	The API focuses on people; it lets users share their activities with each other and access information about their friends.
OpenSocial RESTful Data APIs	Also available is a set of RESTful data APIs that developers can leverage for server-to-server communication. These APIs allow applications to run on channels that do not support JavaScript.
Friend Connect	This is an extension of the two offered OpenSocial APIs. First, it validates the user and then it allows the use of either of the two provided APIs for integrating social features into the third-party website.

Table 5-1 Facebook and OpenSocial APIs

Figure 5-5 shows the results after a comparative evaluation of Facebook and OpenSocial in terms of what user information can be retrieved via the provided APIs. The figure illustrates the common attributes/attributes lists between the two platforms

as well as the platform-specific ones. Detailed descriptions of both Facebook and OpenSocial APIs can be found in Appendix A, where we provide full documentation regarding what user data can be retrieved via the offered APIs.

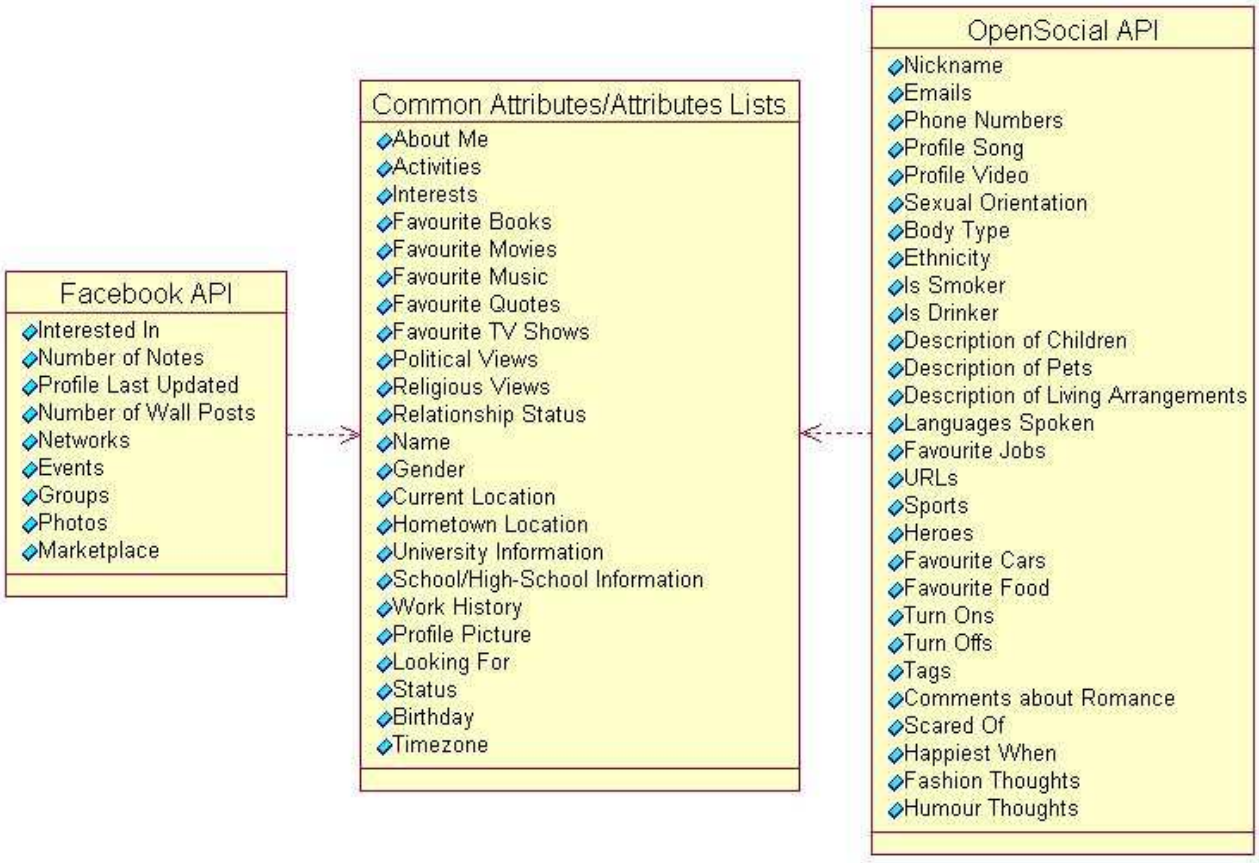

Figure 5-5 Social Networking domain – Facebook vs. OpenSocial.

In addition, Figure 5.6 shows the dominance of Facebook and OpenSocial sites (MySpace, Hi5, Orkut, Cyworld, Friendster, Mixi and Hyves) around the world, according to Alexa.com and Google Trend data; this supports our claim that Facebook and OpenSocial can act as representatives in finding a solution that could assist (almost) every social networking service available today. All of the light green countries are Facebook-dominant. In Russia, the number one social networking site is *V Kontakte*; in China, it is *QQ*; in Brazil and India, it is *Orkut*; in Central America, Peru, Mongolia, and Thailand, it is *hi5*; in South Korea, it is *Cyworld*; in Japan, it is *Mixi*; in the Middle East, it is *Maktoob*; and in the Philippines, it is *Friendster* (TechCrunch, 2009).

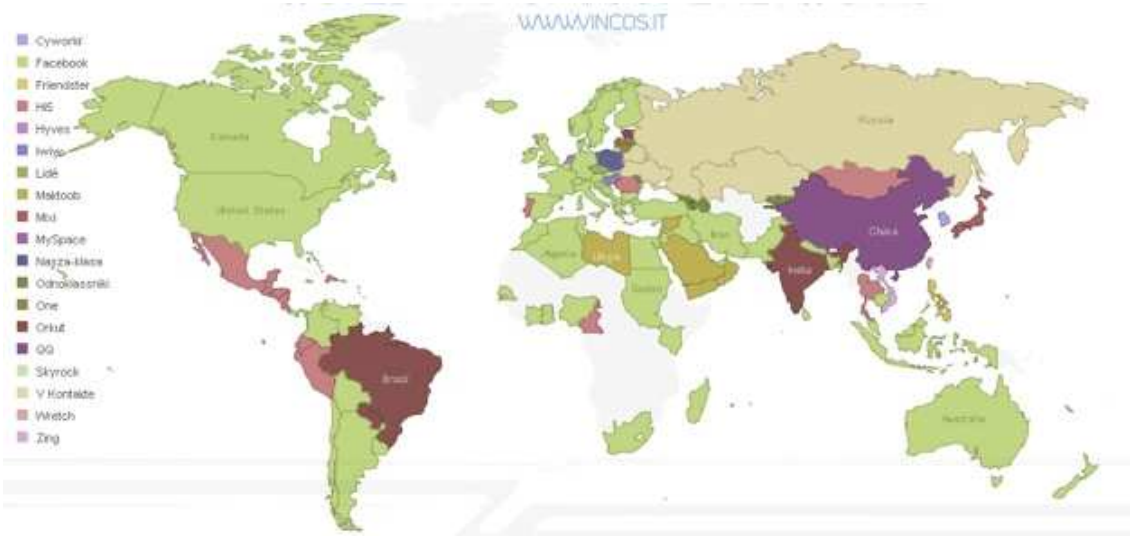

Figure 5-6 Dominance of Facebook and OpenSocial around the world (TechCrunch, 2009)

5.3 E-Commerce Domain

Electronic commerce consists of the buying and selling of products or services over electronic systems and computer networks. With the spread of Internet usage, e-commerce has increased dramatically over the past few decades. E-commerce can be conducted among businesses (B2B) and between businesses and customers (B2C). For our research, we examined two providers who have set the standards in e-commerce, Amazon and eBay, and similarly to the social networking domain, by choosing these services for our case study we aimed to test the part of our claim that a solution that applies to current e-commerce services is possible.

5.3.1 Amazon

Founded in 1994, Amazon began as an online bookstore and has gone from strength to strength. By taking advantage of the bursting of the dot-com bubble, which forced many companies out of business, Amazon survived and continued to evolve. With its innovative technology, it has set the industry standards not only for B2C transactions, but for B2B as well. It was the first commercial website to adopt collaborative filtering methods to analyse customers' behaviour and shopping habits and to generate recommendations based on user data. In addition, it followed a simplistic approach, the one-click ordering, for placing orders and buying products.

According to Internet Retailer (2009) Amazon continues to dominate the e-commerce market because of its sheer size. Last year, sales at Amazon grew more than five times faster than the rest of the B2C e-commerce market. Furthermore, Amazon has been acknowledged as the best overall e-commerce service globally, being awarded the title of "the anchor store for all online shopping" (Time.com and CNN, 2009b). Many new companies attempted to follow Amazon's example, but were unsuccessful; its unique business intelligence, its effective personalisation technologies and its simplified online purchasing has separated Amazon from the other companies and placed it on the front page of the history of e-commerce (ComputerWorld, 2002).

Amazon is currently offering an API that exposes Amazon's product data and e-commerce functionality. This API is called Product Advertising; it can be found at *https://affiliate-program.amazon.com/gp/advertising/api/detail/main.html*, and is part of a collection of remote computing services currently offered by Amazon, known as Amazon Associates Web Service (Amazon Web Services Home Page). Through this API, developers can retrieve product and user information from Amazon, such as items for sale, customer reviews, shopping carts, wish lists, etc. This allows developers and web site publishers to leverage the data that Amazon uses to power its own business, since it makes it extremely easy for developers to build rich, highly effective web sites and applications. Both REST and SOAP versions of the API are provided (Product Advertising API Home Page). In contrast with the social networking domain, developers can take advantage of the Product Advertising API for a fee, since there is a small cost either for every hourly usage of an AWS service or for monthly initiated data transfer and/or storage. Developers pay only for what they use, with no up-front expenses or long-term commitments, making AWS the most cost-effective way to deliver their e-commerce applications to their users.

Amazon Web Services provides a number of already set-up services that developers can incorporate into their applications. From databases to payments, these services help to build great applications cost effectively and with less up-front investment. All AWS services are priced on a pay-as-you-go model, with no up front expenses or long-term commitments. Currently, Amazon is offering seven services which are described below; the following list is a direct quotation from (Amazon Web Services Products):

- **Amazon Elastic Compute Cloud (Amazon EC2):** A web service that provides resizable compute capacity in the cloud. Quickly scale capacity, both up and down, as the computing requirements change.
- **Amazon Simple Storage Service (Amazon S3):** A simple web services interface that can be used to store and retrieve large amounts of data, at any time, from anywhere on the web. It gives any developer access to the same highly scalable, reliable, fast, inexpensive data storage infrastructure that Amazon uses to run its own global network of web sites.
- **Amazon CloudFront:** A web service for content delivery. It integrates with other Amazon Web Services to give developers and businesses an easy way to distribute content to end users with low latency, high data transfer speeds, and no commitments.
- **Amazon SimpleDB:** A web service for running queries on structured data in real time. Amazon SimpleDB is easy to use and provides the core functionality of a database, real-time lookup and simple querying of structured data, without the operational complexity.
- **Amazon Simple Queue Service (Amazon SQS):** A reliable, highly scalable, hosted queue for storing messages as they travel between computers. By using Amazon SQS, developers can simply move data between distributed components of their applications that perform different tasks, without losing messages or requiring each component to be always available.
- **Amazon Elastic MapReduce:** Amazon Elastic MapReduce is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data.
- **AWS Premium Support:** AWS Premium Support is a one-on-one, fast-response support channel to help you build and run applications on AWS Infrastructure Services.

A detailed description of the overall Product Advertising API (Product Advertising API Home Page) can be found at Appendix B, where we present all the user information that can be retrieved from Amazon using this API via any of the seven offered services.

5.3.2 eBay

A second successful e-commerce example is eBay. Its uniqueness is the design on which it was based and with which it has made a huge profit: an online-auction and shopping website where people and businesses can buy and sell their products and services to a worldwide audience. Transactions go through standardized auctions where sellers describe their products and set starting prices before every auction; buyers place their bids hoping to outbid their online opponents. Founded in 1995, initially, eBay was part of a larger website. As the categories of products were constantly increasing, business grew quickly and eBay was expanded worldwide. In 2002, it acquired PayPal, an e-commerce business that allowed payments and money transfers to be made securely through the Internet, and so gained users' trust. Since then, its profits can be estimated in billions of pounds sterling while it stands its ground against rivals and competitors. It has now become the focus of many entrepreneurs hoping to make money by trading online via eBay. During the economic crisis in 2008, many e-commerce providers were forced to change the way they ran their web business and had to use eBay's marketplace as the main source for selling their products (Internet Retailer, 2009). It is not the 1 billion page visits per day nor the 26 billion SQL queries executed every day that emphasize its uniqueness in the e-commerce market; it is the fact that 1.3 billion people make all or part of their living by trading on eBay, which proves its tremendous worldwide success (eWeek.com, 2006). Furthermore, eBay has been acknowledged as one of the top 25 best e-commerce services globally, while earning the title of the best web-based auction service in the world (Time.com and CNN, 2009b).

eBay provides an API that allows developers to communicate directly with the eBay database (eBay API Home Page). By using the API, any application can provide a custom interface, functionality and specialized operations to its users. The API is, in essence, a direct pipeline to eBay and can be used by eBay members, at no financial fee, for 5,000 free API calls per day, which can be initiated by the registered application. The data travels in XML format, making it a consistent and efficient exchange of data, which benefits eBay, the developer, and of course, the end-user.

There are five categories of the complete eBay API; these are all offered by the eBay Developers Program (eBay Developers Program), which allows authenticated access to eBay recourses and data. It is worth mentioning that all categories support

XML and SOAP versions of the eBay API. In addition, almost all categories support REST versions of the API as well.

- **Search**
    - **eBay Finding:** This enables applications to search for eBay items. A REST version of the API is supported.
- **Selling**
    - **eBay Trading:** This offers authenticated access to private eBay data in the areas of listing items, retrieving seller sales status, managing post-transaction fulfilment, and managing private eBay user information, such as *My eBay* details. A REST version of the API is supported.
    - **eBay Large Merchant Services:** This provides a file based, asynchronous execution of large number of transactions in the areas of inventory management and fulfilment. A REST version of the API is supported.
    - **eBay Best Match Item Details:** This category offers authenticated access to private eBay data that enables the sellers understand factors affecting their search ranking. A REST version of the API is supported.
    - **Selling Manager Applications:** This is the next level of platform integration that allows developers to embed their applications where sellers manage their businesses on eBay
    - **eBay Research:** This category lets you retrieve historical eBay data. A REST version of the API is supported.
- **Buying**
    - **eBay Shopping:** This offers access to public read-only data such as searching for items, products, eBay member profiles, popular eBay items and popular searches. A REST version of the API is supported.
    - **eBay Merchandising:** This category offers access to product and item information. A REST version of the API is supported.
- **Users**

- **eBay Feedback:** This provides an easy way to access feedback data, specifically *Detailed Seller Ratings*. A REST version of the API is supported.

- **Alerts**
  - **Server Notifications:** It offers server notifications for subscribed events.
  - **Client Alerts:** It enables retrieval of alert messages for public and private channels.

A detailed description of eBay's Shopping, Trading and Research APIs can be found at Appendix B, where we present all user information that can be retrieved from eBay using these three most popular and most frequently used APIs.

5.3.3 Amazon vs eBay

Table 5-2 provides a summary of the Amazon Product Advertising API and the eBay API:

API	Description
Amazon Product Advertising API. There are SOAP and REST based versions of the API available.	Allows access to much of the data used by Amazon, including the items for sale, customer reviews, and seller reviews, as well as most of the functionality that we see on amazon.com, such as finding items, finding similar items and displaying customer reviews.
eBay API for: • *Selling*	Offers authenticated access to private eBay data to allow automation and innovation in the areas of listing items, retrieving seller sales status, managing post-transaction fulfilment, and managing private eBay user information, such as My eBay and Feedback details.

• *Buying* • *Researching* XML, SOAP and REST based versions of the API exist.	Offers access to public data, such as searching for items, products, eBay member profiles, popular eBay Items and popular searches. Allows the retrieval of historical and statistical eBay data.

Table 5-2 Amazon Product Advertising API and eBay API

Figure 5-7 shows the results after a comparative evaluation of Amazon and eBay, analyzing what user information can be retrieved via the available APIs. Figure 5-8 presents the lists of attributes common to the two providers as well as the provider-specific ones. Detailed descriptions of both Amazon and eBay APIs can be found in Appendix B, where we provide full documentation regarding what user data can be retrieved via the offered APIs.

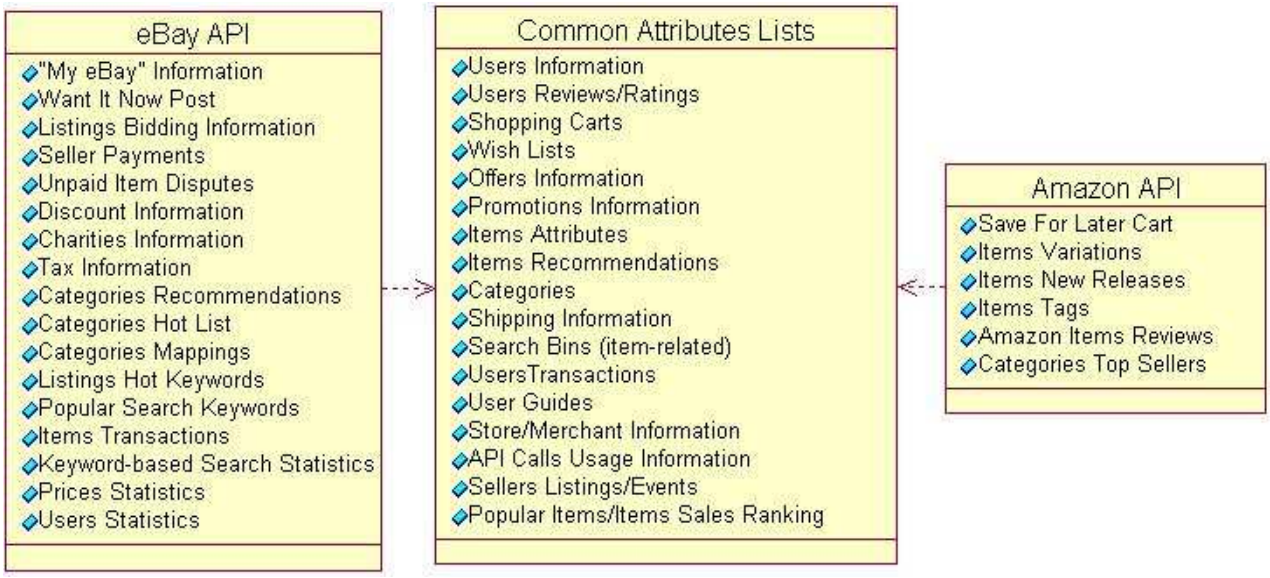

Figure 5-7 E-Commerce domain – Amazon vs. eBay

5.4 Social Networking and e-Commerce in Lifelong User Modelling

After the brief review of the social networking and e-commerce domains, we now examine the relevance of these two domains in the area of LUM, since, as we will describe in Chapter 6, we aim to identify the potential lifelong value these domains can bring to LUM in order to enrich the several user data sets that are currently being used by educational systems for personalisation services.

It seems likely that it would be possible to identify the increasing relevance of social networking in educational areas. People love to socialise online and the influence of social networking services is growing, as has been demonstrated by Facebook and OpenSocial's participatory websites (ChannelWeb Network, 2009c). In addition, people have shown many signs of favouring e-commerce systems, and two very good examples, Amazon and eBay, are described above (IEEE Computer Society, 2002). Furthermore, the extensive research we conducted in this area has led us to the conclusion that while people spend their time online using services from these two examined domains, they do not hesitate to share a significant amount of personal information when visiting services that are built on social networking and e-commerce platforms. Facebook's official statistics clearly demonstrate the degree of user engagement regarding sharing and posting personal data on this social networking platform: more than 3.5 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) are shared each week, more than 2.5 billion photos are uploaded to the site each month, more than 3.5 million events are created each month, and more than 55 million status updates are posted each day (Facebook Official Statistics). Although the majority of users may not understand how these services work and may not comprehend the privacy policies of the several social networking and e-commerce providers, which, as we have shown in section 3.4, have resulted in lawsuits against these providers, they will still trust these services enough to post and share personal information for social and business purposes.

LUM, on the other hand, as we saw in Chapter 4, requires long-term data for the sufficient extraction of conclusions and results. What could be a better source of information than these two domains? If we could find a way to gather all the data located in several services from the two examined domains and manage to make sense of them, since different systems may assign different meanings to the same piece of information, and somehow find a way to pass it to educational systems, we would be

able to enrich with long-term (lifelong) user data those internal user models that are currently being used for personalisation purposes.

Although e-commerce is emphatically neglected when experts consider the potential contribution that Web 2.0 technologies can bring to higher education, it is acknowledged that the "social web indeed matters" today (Armstrong and Franklin, 2008). The benefits and implications have been examined along with the advantages and disadvantages resulting in realistic examples where the social networking domain can bring value to the (educational) table. This has been summarised in a review of current and developing international practice in the use of social networking (Web 2.0) in higher education, which spans the United Kingdom, the United States, Australia, the Netherlands and South Africa (Armstrong and Franklin, 2008). Although this proves the case for the social networking domain, as we will claim in Chapter 6, we believe that e-commerce can offer a similar significant value to personalisation when user information posted and/or generated on e-commerce websites is harvested. In Chapter 7, we will present tangible examples to demonstrate the potential value that both examined domains can add to the current educational value offered today by personalisation services.

Finally, social networking and e-commerce providers have engaged in data portability 'movements', which have been demonstrated by the providers 'opening their doors' to their users. They have released APIs that allow their users to access their own information stored inside the providers' databases. Clearly, different policies and terms and conditions exist across these initiatives, but the main point is that a new era has begun where providers of user models are now passing ownership of information back to its owners.

5.5 Summary

In this chapter, we examined two important domains for UM, namely, the social networking and e-commerce domains. We investigated the two best representatives from each domain and gave a brief report on each one. In addition, we hinted that there is a significant relevance to LUM since the data stored inside applications from the two examined domains are often long-term data; this matches LUM's objectives. Finally, we described how the two representatives from each examined domain have inclined towards data portability initiatives that pass at least some of the control of information

back to its owners. We briefly presented their offered APIs and explained their data portability projects.

Chapter 6 Research Questions and Hypotheses

6.1 Introduction

In Chapter 6, we first describe the identified problem in the area of UM for which we attempted to provide a solution. Next, we present the main research question that motivated the work described in this thesis, followed by the list of sub-questions that helped us to split the problem into three categories based on the three UM components that were analysed in Chapter 3: interoperability, scrutability and privacy. In addition, we present our hypotheses, which will be tested in the following chapters. Finally, we explain the research approach that we adopted to answer our research questions and test our research claims.

6.2 Identified Problem

In Chapter 2, we discussed the progress of UM from User Modelling Shells to User Modelling Servers. In addition, in Chapter 3, we presented the latest state of UM by describing some of the solutions proposed by the community. What has been observed is that while we find UM in a state of transition between the 'old' personalisation approach of autonomous stand-alone and online systems to the 'new' frameworks for achieving interoperable user profiles, and while there is development and progress, it is still being applied single-dimensionally.

- Most adaptive systems developed are using only their own internal models when offering personalisation services to their users. This is understandable to a

certain degree, since the majority of current adaptive systems, for example, AHA and InterBook, were designed and implemented before the recent evolvement of UM, which was triggered by the introduction of the Semantic Web.

- Newly introduced frameworks and architectures, while offering a solution in achieving interoperability across peer systems, do not involve providers of user models beyond the educational domain. As described in Chapter 3, although the proposed solutions are adequate in terms of exchanging user models among educational systems, they do not consider other domains, such as the social networking and e-commerce domains.
- User Modelling Servers, a client-server architecture for allowing central information storing and, simultaneously, data access and retrieval, are mostly designed and developed to meet commercial requirements (Kobsa, 2001). Although, currently, this seems to be the best solution for the future of UM, and this can be justified with Personis (Kay *et al.,* 2002) and its unique scrutable design, further research is required to tackle the identified challenges, especially when considering the different requirements that need to be addressed when interacting with services from the two examined domains we analysed in detail in the previous chapter.

We are losing user information, which is flowing around the WWW, because we are not thinking multi-'*domain*'sionally, i.e., including more than one domain when modelling user data that is located on the WWW. We can enrich UM, and more specifically LUM, if we find a way to model our every day (life-long) interactions with online services from the two popular domains, the social networking and the e-commerce domains, in order to enrich user information sets, which are currently being used by educational systems, like Adaptive Hypermedia Systems, for personalisation purposes, as demonstrated in Chapter 2 (Kyriacou *et al.,* 2009a). However, is this feasible? How can we realistically acknowledge the maturity status of our current knowledge while taking advantage of existing methods and technologies in order to achieve the multi-'*domain*'sional big picture presented in Figure 6-1?

Figure 6-1 The multi-'*domain*'sional big picture (Kyriacou *et al.,* 2009a)

Firstly, we need to understand that there is a vast amount of user information outside the educational domain. Specifically, the social networking domain, e.g., Facebook, MySpace, etc., and the e-commerce domain, e.g., Amazon, eBay, etc., welcome millions of users daily on their respective sites and a huge number of queries are executed every second, storing and retrieving user data held by several systems in these domains. In addition, users post and share personal data on social networking and e-commerce websites for long periods of time, sometimes even lifetime periods. These data are often a better representative of users' interests, activities, goals and knowledge than is any other information posted anywhere else on the WWW.

In addition, recent data portability announcements from the two key players in the social networking domain (PC World, 2008b), Facebook and OpenSocial, have revealed these providers' initiatives to pass user data back to their 'owners'. Various versions of APIs are available from providers of user models, in the social networking and the e-commerce domains, which can be used by developers to 'take a dip' inside the providers' databases, retrieve user information held by these providers, after obtaining the direct users' consent, and enrich their websites with social and/or e-commerce features.

Moreover, Semantic Web Technologies have provided us with effective solutions in viable problems, like the description of resources in a machine-understandable way and standards for communication and exchange among independent providers across various platforms. UM is not an exception; thus, by taking advantage of these

technologies we can develop infrastructures that will allow us to take a step forward, towards multi-'*domain*'sional visions. Expressing user models in RDF, using Web Services for communication between systems and exchanging user information has already become a common practice in the area, as we demonstrated in Chapter 3 (Alrifai *et al.,* 2006; Dolog and Schaefer, 2005a).

6.3 Research Questions and Hypotheses

This main research question that motivated this work is:

What are the requirements for adopting a user modelling infrastructure among users and social networking and e-commerce providers which will:

- *provide an interoperable solution to merge user models from providers of the two examined domains and introduce a communication protocol for enabling the exchange of user data among them and educational personalisation systems,*
- *allow users to have control of their imported data by inspecting and altering the way they are being modelled in such an infrastructure,*
- *offer appropriate privacy options to encourage users to define their privacy preferences in such an infrastructure.*

To answer the main research question as set out above, we created a list of sub-questions that were grouped into three categories, similarly to the examined areas that were presented in Chapter 3. In addition, we formed our hypothesis for each question, which we will address and test in Chapters 7, 8 and 9 in order to verify or dismiss them.

6.3.1 Interoperability

Question 1:

Can we make possible the import of user data from social networking and e-commerce providers to the proposed infrastructure?

Hypothesis 1:

We claim that this will be possible via a web-based service that will allow users to import their information from social networking and e-commerce providers. This service will be responsible for handling the technical requirements when a user initiates

the import of his/her data from a provider, using the available APIs and/or data portability tools, which were presented in Chapter 5.

Question 2:

Where do services from the two examined domains stand regarding passing the control of information back to its owners?

Hypothesis 2:

As shown in Chapter 5, social networking and e-commerce providers offer APIs that developers can leverage to retrieve user data from these providers' platforms, after obtaining the direct users/owners' consent. In addition, the social networking services, Facebook and OpenSocial, have announced their data portability initiatives, Facebook Connect and Friend Connect, which can be used to incorporate the corresponding platforms into third-party websites. It is our hope that this notion of passing the control of information back to its owners will continue to evolve, with newcomers adopting the example of the pioneers, Facebook and OpenSocial, in the social networking domain, and Amazon and eBay in the e-commerce domain.

Question 3:

Is it possible to design a common model architecture that can accommodate the various user models retrieved from social networking and e-commerce providers?

Hypothesis 3:

After the comparative evaluation we conducted for the two representatives from each domain, Facebook and OpenSocial in social networking, and Amazon and eBay in e-commerce, we claim that if we design a common models' architecture that would fit these representatives' data models, based on the market share these providers hold, as demonstrated in Chapter 5, it would be safe to assume that the proposed models' architecture would fit other current social networking and e-commerce services as well.

Question 4:

Can we enable social networking and e-commerce providers of user models to define their data models in the proposed infrastructure, in order to allow the exchange of user data between them and the proposed infrastructure?

Hypothesis 4:

As explained in section 3.2 in the literature review, Semantic Web technologies can help us to answer this question. We claim that ontologies can adequately be used to describe a provider's data model, from both examined domains, assuming we present the necessary interfaces, via the online service for doing so. In addition, the frameworks and architectures we presented in section 3.2 and the GRAPPLE project, which was presented in section 4.4.1, lead us to claim that administrative interactions with the online service will be necessary in order to ensure the proper definition and description of each providers' data model.

Question 5:

How can we map the different data models, from the social networking and e-commerce providers of user models in the proposed infrastructure, in order to overcome the semantic barriers when importing user data from these providers into the proposed infrastructure?

Hypothesis 5:

Similarly to the hypothesis in question 4, and based on section 3.3, ontologies can be used to map the various data models of social networking and e-commerce providers inside the infrastructure, via the necessary interfaces of the web-based service. We also claim that mapping the representatives' data models inside the infrastructure in advance would help other services to relate to them, since we assume, based on the market share that Facebook, OpenSocial, Amazon and eBay hold in the two examined domains, that these representatives' data models are supersets of the other services' data models.

Question 6:

Which communication protocol can best serve the process of exporting user data from the proposed infrastructure to educational personalisation systems?

Hypothesis 6:

As our research in Chapters 3 and 5 revealed, web services and the REST protocol are favoured in the UM community as communication protocols and are currently being adopted by social networking and e-commerce providers as well. We claim that the REST protocol, with its simple but elegant and effective design, can adequately serve us as a communication protocol for exporting user data towards educational personalisation systems, after direct consent from the data owners has been obtained.

Question 7:

How can we enable educational personalisation systems to express interest in receiving specific social networking and/or e-commerce user models?

Hypothesis 7:

As already expressed, we think that the development of an interface to the infrastructure in the form of an online service will help us to answer this question as well. The necessary interfaces can allow educational personalisation systems, and especially the adaptive hypermedia systems that were presented in Chapter 2, to express interest with user modelling providers by subscribing to them, an approach used in the frameworks and architectures that were presented in section 3.2.

In addition, we claim that user information retrieved from social networking and e-commerce services can bring long-term, even lifetime value to personalisation systems, based on the LUM criteria, as described in Chapter 4, in specific use cases, which we aim to identify.

6.3.2 Scrutability

Question 8:

What scrutability privileges should be offered to users? How can we assess whether the proposed scrutability privileges are adequate for such a proposed infrastructure? How can we assess whether potential users will understand and accept the proposed scrutability privileges?

Hypothesis 8:

We claim that the proposed scrutability user privileges, which we offer to potential users and which are presented and evaluated in Chapter 9, will be appropriate to be offered in the proposed infrastructure for the social networking and the e-commerce domains, and accepted by users as well.

Assuming we present to our users the three tasks that will be presented in Chapter 9, where each task will expose one of the proposed scrutability user privileges, we claim that:

a) Users will be able to complete each task successfully, irrespective of the way the tasks are presented to them, (competence).

b) Users will express their acceptance of having the proposed user privilege available for use in the proposed infrastructure, irrespective of how the tasks are presented to them, (acceptance).
c) Users will be familiar with their decisions during and after the completion of each task, and will understand the consequences of each decision they will take, irrespective of how the tasks are presented to them, (consequence).

We claim that we can assess whether a proposed scrutability user privilege is appropriate to be offered in the proposed infrastructure by evaluating users':

a) Competence for completing each task (compare actual outcome after the completion of the tasks with users' answers to the evaluation questions [have they actually done it versus do they think they have done it]
b) Understanding of the consequences of their decisions while interacting with each task.

We claim that we can assess whether a proposed scrutability user privilege is accepted by users by evaluating users' acceptance of the proposed scrutability user privilege by asking them directly what they thought about each task and whether they would like the proposed infrastructure to offer it to its users

Question 9:

How informed are users today regarding the term 'scrutability'? Do they recognise any scrutability options when interacting with social networking and e-commerce providers today?

Hypothesis 9:

Our assumption is that users will not be familiar with the term 'scrutability', as we described in Chapter 4 when the various LUM challenges were presented, although they may have been exposed to some scrutability privileges when interacting with social networking and e-commerce services.

6.3.3 Privacy

Question 10:

What privacy privileges should be offered to users? How can we assess whether the proposed privacy privileges are adequate for such a proposed infrastructure? How

can we assess whether potential users will understand and accept the proposed privacy privileges?

Hypothesis 10:

We claim that the proposed privacy user privileges, which we offer to potential users and which will be presented and evaluated in Chapter 9, will be appropriate to be offered in the proposed infrastructure for the social networking and the e-commerce domains, and accepted by users.

Assuming we present to our users the three tasks that will be presented in Chapter 9, where each task will expose one of the proposed privacy user privileges:

a) Users will be able to complete each task successfully, irrespective of how the tasks are presented to them (competence).
b) Users will express their acceptance of having the proposed user privilege available for use in the proposed infrastructure, irrespective of how the tasks are presented to them (acceptance).
c) Users will be familiar with their decisions during and after the completion of each task, and will understand the consequences of each decision they will take, irrespective of how the tasks are presented to them (consequence).

We claim that we can assess whether a proposed privacy user privilege is appropriate to be offered in the proposed infrastructure by evaluating users':

a) Competence for completing each task (compare actual outcome after the completion of the tasks with users' answers to the evaluation questions [have they actually done it versus do they think they have done it]
b) Understanding of the consequences of their decisions while interacting with each task.

We claim that we assess whether a proposed privacy user privilege is accepted by users by asking them directly what they thought about each task and whether they would like the proposed infrastructure to offer it to its users

Question 11:

What privacy settings will allow users to define their privacy preferences in the proposed context?

Hypothesis 11:

While keeping in mind that there is no perfect trade-off between privacy and user modelling, as discussed in section 3.4, we believe that three categories of privacy settings will satisfy users' needs in such an infrastructure:

a) Category 1: data that the owner is happy to share with everyone and that do not require strong privacy protection
b) Category 2: data that the owner wants to keep completely hidden from other users
c) Category 3: data to which the owner would like to apply some level of privacy that cannot be satisfied with the privacy settings of the first or second categories.

6.4 Research Approach

Following on from the previous section, we now present our approach for testing our hypotheses in order to verify or dismiss them according to the outcome of our research work.

The first task we undertook was building the infrastructure. Using an ontological approach, we created the necessary classes, datatype and object properties in order to define properly the data models, and their relationships, of Facebook and OpenSocial for the social networking domains and the data models, and their relationships, of Amazon and eBay for the e-commerce domains. In addition, we investigated whether we could import and extend other existing ontologies that have been adopted by the UM community. Furthermore, we designed a models' architecture that can fit and handle any imported user information from the four representatives. Moreover we used the Facebook API and Facebook Connect to test whether we could actually copy our own and our friends' data from the Facebook platform to a third-party website. Finally, we adopted the REST approach by trying out the sqlREST tool for exporting the attributes' values in URIs. This work is presented in section 7.2.

Secondly, implementing the interface to the infrastructure was essential for demonstration and evaluation purposes. Therefore, we designed and implemented a web-based service for acting as the interface to the infrastructure and for demonstration and evaluation purposes. We also designed and implemented three scrutability and three privacy user privileges with the aim, following the LUM guidelines presented in

Chapter 4, of allowing potential users to scrutinise how they are being modelled while having the option to express their privacy preferences. This work is described in section 7.3.

Thirdly, we identified some use-cases where the proposed solution could be of benefit to educational personalisation systems, in terms of enriching the sets of user data that are used for generating personalisation services. This work is detailed in section 7.4

Fourthly, we test our hypotheses on interoperability by simulating realistic hypothetical scenarios with one of the four representatives, the social networking king Facebook, and an educational personalisation system, the adaptive hypermedia system AHA!. We tested both retrieving user data via the Facebook API inside the Facebook platform and via Facebook Connect towards the online service that was created. We describe the evaluation design and results for testing our interoperability hypotheses in Chapter 8.

Next, the aim was to evaluate the proposed scrutability and privacy privileges by conducting two user-based evaluations. We provided the online service to potential users and asked them to complete tasks that revealed the proposed scrutability and privacy privileges. In addition, we tested whether different approaches for presenting the tasks to the participants may affect the resulting outcome of the evaluations. The evaluation design and results for testing our scrutability and privacy hypotheses are presented in Chapter 9.

Lastly, we identified our future work agenda while proposing further research in the area of LUM. In addition, we revisited our hypotheses and, based on the evaluation results, we dismissed them or verified them accordingly. More detail of this is given in Chapter 10.

6.5 Summary

Chapter 6 has described the problem in UM that we addressed with our research work, followed by our research questions and testable hypotheses. Moreover, we explained the research approach that we followed to propose a solution to the identified problem and assess whether it satisfies the LUM criteria that we set out to prove: interoperability, scrutability and privacy.

Chapter 7 A Scrutable User Modelling Infrastructure

7.1 Introduction

Chapter 7 is designed to present our proposed solution to the identified problem in the area of UM and, more specifically, in LUM, with the purpose of testing and evaluating our hypotheses and claims, as described in Chapter 6. We start by suggesting a scrutable user modelling infrastructure (SUMI), which we believe brings a realistic solution for harvesting user information that is posted and shared on social networking and e-commerce services, in order to enrich the user models that are currently being used by educational systems for personalisation purposes. Furthermore, a prototype SUMI service is presented, which was built for demonstration and evaluation purposes. First, we explain the approach we adopted to solve the interoperability problem. Next, we address the scrutability and privacy aspects/challenges of our work. Finally, this chapter provides tangible real-life examples that demonstrate the usefulness of our proposed solution.

7.2 SUMI: The Proposed Infrastructure

Following up from the previous chapter, we have identified a gap in the area of UM that falls into the LUM sub-area. Further research is needed to examine the potential of achieving interoperability across the social networking and the e-commerce domains in order to harvest the various user models that are stored in services from the two examined domains in a way that will enrich the current user models of educational systems for potentially improved personalisation services. In addition, we feel that

there is a need to identify how scrutability and privacy can assist in this attempt, especially when seeking users' agreement and cooperation in using their personal information for educational purposes in a way that will be understandable and acceptable to them.

Special consideration has been given to collecting the requirements for employing a Scrutable User Modelling Infrastructure (SUMI), which will also introduce a communication protocol, in an attempt to support both users and providers of user models, in the social networking and e-commerce domains, for exchanging the various user models that these providers currently hold, with educational (personalisation) services (Kyriacou, 2008; Kyriacou, 2009; Kyriacou and Davis, 2008; Kyriacou *et al.,* 2009a). In this section, we address the research questions and hypotheses for interoperability, as presented in section 6.3.1.

After an extensive literature review, which was presented in Chapters 2, 3 and 4, and an analysis of the current situation in social networking and e-commerce domains, which was described in Chapter 5, we produced a four-step approach that could help us answer our research questions by testing our hypotheses for interoperability. These steps, which follow our overall research approach described in section 6.4, are described next and are addressed in detail in the following sub-sections.

1. The first step, which is tackled in section 7.2.1, requires the acquisition of various user data, which are posted and shared in several social networking and e-commerce websites. A standard way of retrieving these data from the various providers of user models, towards a central infrastructure that will allow users/owners to gain control of their own personal information has to be identified while realistically acknowledging the complexity of making this happen given the current situation in social networking and e-commerce domains.
2. The second step is concerned with organizing the user data once they are imported, after obtaining the direct users' consent, into our proposed infrastructure. A close examination of the literature review may result in an infrastructure that could fit (hopefully) all imported models from both examined domains. This issue is addressed in section 7.2.2.
3. The most challenging step is to address the need to cope with the several internal data models of the various providers of user models from the two examined domains. Different meanings could be assigned to the same

attributes and different relationships may exist between the same user data. While our aim is to find a way to import these data into SUMI, we acknowledge the need also to find a way to import the meaning of these data as well, exactly as expressed by each provider of the user models. Section 7.2.3 deals with this challenging step.

4. Finally, the last step is concerned with exporting the user information from SUMI towards educational personalisation systems. What is the best way to achieve this, in which format should we export user information from SUMI, and how can we allow educational personalisation systems to express their preferences regarding which user data they would like to receive? We address this question in section 7.2.4.

7.2.1 Acquiring User Information from Providers of User Models

In Chapter 3, every proposed solution that we presented has, as a requirement, that every participating educational system would follow the proposed protocol in order to ensure consistency when exchanging user models between peer systems. It is our assumption that this is not the case for the social networking and e-commerce domains. We should not expect Facebook and Google to agree on a common API, nor Amazon and eBay to adopt a common internal structure, especially when the level of competiveness is constantly rising, with these 'big' companies acquiring other smaller services in order to challenge each other (BBC News, 2009a). To date, there has not been any indication even to hint that providers from these two domains are even moving towards that direction, and for this thesis, we think it is valid to assume that such co-operations are currently far from likely to happen.

In Chapter 5 and Appendices A and B, we present the APIs offered by Facebook, OpenSocial, Amazon and eBay. On the one hand, we observe even more and more providers adopting data portability initiatives and offering APIs to allow the end-users to retrieve their data that were posted and shared on their platforms (PC World, 2008b). On the other hand, unfortunately, these APIs differ from each other, since each provider has considered its own requirements and internal structure when developing its own API. Although OpenSocial has convinced 33 social e-networks to adopt its proposed standard (OpenSocial Containers), it is safe to assume that the time when every social e-network adheres to a common standard may never come.

To put the previous two paragraphs into perspective, we think that the social networking and e-commerce domains are very different from the educational domain. Providers of user models will play only by their rules and they will currently not follow common standards and protocols to achieve the greater good. This assumption results in one main conclusion: if we want to acquire user data from social networking and e-commerce providers we must comply with their standards and use their APIs.

The question of how many social networking and e-commerce providers are currently offering APIs to their users has to be addressed as well. After conducting extensive research, we still could not identify the exact number of social networking and e-commerce services that are currently offering an API to allow users to retrieve their data from their platforms. However, according to the list of the top 50 websites of 2009 based on user enjoyment and ease of navigation (Time.com and CNN, 2009a), 22 social networking and 6 e-commerce websites that have made the list, and which are summarized in Table 7-1, do have an API available for use. One important observation that is worth mentioning at this point is the fact that all these 28 services offer REST versions of their APIs.

Social Networking Websites	E-Commerce Websites
Flickr (http://www.flickr.com/) – The first site to introduce collaborative tagging on photos uploaded by users.	*Amazon* (http://www.amazon.com/) – As stated in Chapter 5, Amazon is the leading force in e-commerce today
Delicious (http://delicious.com/) – It is considered to be the Flickr of bookmarking while, recently, it has introduced a search engine add-on.	*Kayak* (http://www.kayak.com/) – A travel site that allows users to book plane tickets while offering a search engine for finding the lowest possible price.
MetaFilter (http://www.metafilter.com/) – A weblog to which anyone can contribute a link or a comment.	*Netflix* (http://www.netflix.com/) – A movie-by-mail service that also offers a streaming-video service.
Twitter (http://www.twitter.com/) – A micro-blogging site that allows users to post 140 character-long twits.	*Etsy* (http://www.etsy.com/) – It is the go-to site for handmade fashion, furniture, toys and housewares.
Skype (http://www.skype.com/) – A widely used Voice Over Internet Protocol	*PropertyShark* (http://www.propertyshark.com/) – A

(VoIP) service that allows users to make video calls over the Internet.	subscription-based service that offers an incredible amount of data for selling houses in big population centres.
YouTube (http://www.youtube.com/) – A service that adopts collaborative filtering and allows users to upload their own personal videos.	*Kiva* (http://www.kiva.org/) – A peer-to-peer micro-lending service. Users can lend money to applicants while Kiva acts as the commission-based mediator until the money is paid back.
Hulu (http://www.hulu.com/) – It offers on-demand TV and movies, streamed in high definition	
Vimeo (http://www.vimeo.com/) – A service similar to YouTube with lightly curated content and high video quality.	
Wikipedia (http://www.wikipedia.org/) – The online encyclopaedia that anyone can edit, update and to which they can contribute an entry.	
Internet Archive (http://www.archive.org/) - This site's mission is to offer archives of various other services on the web. For example, users can find out how Times online looked back in 1998 and how eBay came out on its first day.	
Facebook (http://www.facebook.com/) – The biggest social network on the WWW, discussed in Chapter 5.	
Last.fm (http://www.last.fm/) – A radio-killer service that allows users to create their own playlists while also offering suggestions on other songs that users might like based on their preferences.	
Spotify (http://www.spotify.com/) – A	

service that offers an archive of every music album in the world. Users can stream anything to which they would like to listen with no royalty fees.	
Yelp (http://www.yelp.com/) – Currently considered as the mother of all review sites, the place where users can read and post reviews about restaurants, bars, boutiques, dentists, etc.	
CouchSurfing (http://www.couchsurfing.org/) – A worldwide network for making connections between travellers and the local communities they visit.	
Mint (http://www.mint.com/) – A free web-based service that takes the information, after obtaining the user's consent, from banks, brokerages and credit-card companies and collates it into a single easy-to-use record.	
TripIt (http://www.tripit.com/) – With TripIt, users can simply forward their travel information to the service and it will return weather forecasts and maps so users can be fully prepared before they travel.	
Aardvark (http://www.vark.com/) – A new kind of search engine that lets users ask friends and friends of friends any question that may arise at any time and that cannot be answered by traditional search engines.	
drop.io (http://drop.io/) – A private file-sharing service where users can store their	

personal files in order to access them at any time and from anywhere with a connection to the WWW.	
Issuu (http://www.issuu.com/) – An online newsstand with infinite shelf space, Issuu offers hundreds of interesting published projects that can be accessed via an online reader.	
Photosynth (www.photosynth.net/) – Instead of arranging photos in traditional albums, this site finds relationships among pictures and creates a 3-D photo environment called a "synth".	
Fonolo (http://www.fonolo.com/) – Fonolo helps users skip the phone menus when calling large companies by making the call, pressing the right buttons and staying on hold until a real person answers the phone call. Then it rings the user's phone and lets him continue with the call.	

Table 7-1 List of social networking and e-commerce services (from the 2009 top-50 websites) that currently offer APIs

7.2.2 SUMI Models' Architecture

Looking back at the approach we described at the start of section 7.2, the second step requires the architecture of SUMI models that will "prepare" the infrastructure to accept the various social networking and e-commerce models imported by the SUMI users to be defined. The objective in this sub-section is to identify a common models' architecture that would ideally accommodate any provider's data model in the social networking or e-commerce domains.

A SUMI model can be considered as a model of models, where users can determine which models will be added to their SUMI collection. For that reason, the introduced architecture should be flexible and capable of supporting a variety of

models from the two domains. Looking closely at the services that made the list of 2009 top-50 websites, and after conducting a comparative evaluation of the two representatives in each domain (Figures 5.5 and 5.7), we identified four categories of user data one can find stored inside social networking and e-commerce user models. These four categories are presented and described in Table 7-2:

Category of User Data	**Description**	**Inclusive Criterion in SUMI**	**Example**
Common User Input	User data that is manually entered by the user; it is common to services either in the social networking or in the e-commerce domain	All participating social networking services offer it to their users OR All participating e-commerce services offer it to their users	A user's personal interests, which are common user input in participating services from the social networking domain
Service-Specific User Input	User data that is manually entered by the user; it is service specific in both the social networking and in the e-commerce domain	At least one participating social networking service does not offer it to its users OR At least one participating e-commerce service does not offer it to its users	A user's favourite heroes, which is user input one can find in MySpace but not in Facebook.
Common Generated Information	User data that is generated by a service, based on user input that was previously	All participating social networking services offer it to their users OR	A user's recommendations based on his/her purchase records

	provided; it is common to all services in at least one domain	All participating e-commerce services offer it to their users	
Service-Specific Generated Information	User data that is generated by a service, based on user input that was previously provided; it is service specific, in the social networking or in the e-commerce domain	At least one participating social networking service does not offer it to its users OR At least one participating e-commerce service does not offer it to its users	A user's wish list of items that they would like to own; can be found in Amazon, but not in eBay

Table 7-2 The four categories of user data that can be identified in the social networking and e-commerce domains

In addition, we also noticed that every piece of user information that is stored in these models is assigned to a group of user attributes, which aims at grouping the data based on their meaning. For example, users' favourite interests and activities are grouped under *Personal Information*, and users' list of items that they have not yet purchased, but would like to own, are listed under *Wish List*.

Therefore, a four-category architecture of SUMI models is proposed in this sub-section, which has been found to fit adequately all data models of the providers that were presented in Table 7-1, along with the data models of the four representatives, Facebook, OpenSocial, Amazon and eBay, that were examined in Chapter 5. Using this architecture, we are able to provide a common data model structure for every model imported in SUMI, with no barriers if the model is retrieved from either a social networking or an e-commerce provider. This architecture comprises four categories:

Generic User Data: Any input that is being entered by the user manually in any way AND is common to all participating services either in the social networking or the

e-commerce domain; e.g., The "Interests" field that users can find in Facebook and OpenSocial websites (Hi5, MySpace, etc.)

Service-Specific User Data: Any input that is being entered by the user manually in any way AND is specific to at least one, but not all, participating services in the social networking or the e-commerce domain; e.g., MySpace's "Favourite Heroes", which is not provided in Facebook.

Service-Generic Generated Information: Any user information, represented by an attribute or attribute list, that is generated by the provider based on previous user input AND is common to all participating services either in the social networking or the e-commerce domain, e.g., "Items Recommendations", which users receive from both eBay and Amazon.

Service-Specific Generated Information: Any user information, represented by an attribute or attribute list, that is generated by the provider based on previous user input AND is specific to at least one, but not all, participating services in the social networking or the e-commerce domain; e.g., Amazon's "Items New Releases", which is not found on eBay.

Table 7-3 provides an overview of our proposed SUMI models' architecture:

Generic User Data (contains *common user input* data)	**Service-Specific User Data** (contains *service-specific user input* data)
Service-Generic Generated Information (Contains *common generated information* data)	**Service-Specific Generated Information** (Contains *service-specific generated information* data)

Table 7-3 Proposed SUMI models' architecture

It is worth mentioning that this presentation has provided us with an interesting discovery: social networking providers do not currently offer any generated user information, whereas the e-commerce providers are mostly engaged in generating information using recommendation components and techniques, rather than asking their users for manual data input. In our opinion, this will change in the near future, since services will evolve and eventually add both social networking and e-commerce features to their websites in order to target and potentially attract more users.

7.2.3 Handling Different Data Models from Various Providers

The most challenging step in developing such an infrastructure is finding a way to handle the different data models of the various providers. Two providers may assign different meanings to an attribute and store it in a different data structure. For example, in Facebook, a relationship between two users is called a 'friendship', but in LinkedIn is called a 'connection'. Another example can be found regarding the attribute "status", which is offered by social networking services to users to express their thoughts at any particular moment. Twitter offers this attribute with a character limit of 140, whereas Facebook allows users to express themselves without any character limits.

To address this issue, we reviewed the literature of Semantic Web technologies, which is presented in section 3.2.2, and we found ontologies to be an adequate solution to this problem. An ontology can be considered as a formal representation of a set of concepts within a domain and the relationships between those concepts (W3Schools – Introduction to OWL). Furthermore, there is a significant number of contributions in the area (Abel *et al.,* 2009a; Elliot *et al.,* 2009; Heckmann *et al.,* 2009) that have favoured ontologies when addressing this problem. Therefore, we have concluded that this solution can help us to map the different data models of the several providers from the two examined domains *and* the relationships between them, while we attempt to describe each imported data model in SUMI using the proposed architecture we presented in the previous sub-section.

Furthermore, in section 3.2.1, we referenced two UM standards, PAPI and LIP. Although these standards enjoy a significant distribution across contributions that involve educational systems (Dolog and Schaefer, 2005a; Simon *et al.,* 2004), they are sadly ignored in the social networking and e-commerce domains. Therefore, we followed a common practice in the area of UM (Devedzic *et al.,* 2007), which entails adopting some parts from these two standards while introducing new attribute components when developing domain ontologies. In our case, we developed a SUMI ontology, which is described below and which aims to model services from the two examined domains by adopting the proposed SUMI models' architecture.

SUMI Ontology: A SUMI ontology has been developed, and can be found at *http://mysumi.org/sumiOntology.owl*, while keeping in mind the structure of the four best representatives from the social networking and e-commerce domains and the proposed SUMI models' architecture, in order to enable mapping of the various

providers' data models from the two examined domains to occur, for successful communication between them and educational personalisation systems via SUMI. The ontology introduces five categories while adopting some other categories from two already existing and well-accepted ontologies, which incorporate parts from the PAPI and LIP standards:

- the *General User Modelling Ontology (GUMO),* which was created for the uniform interpretation of distributed models in intelligent semantic web enriched environments (Heckmann *et al.,* 2007).
- the *Friend Of A Friend ontology (FOAF)*, which has been developed to describe people, the links between them and the things they create and do (FOAF Vocabulary Specification).

Although we avoid re-creating several features and categories by adopting these two widely used ontologies, we feel they need to be extended since they lack some important elements for successfully modelling the social networking and e-commerce domains. Figure 7-1 shows the five introduced categories, *Attribute, AttributeCategory, DictionaryConcept, Provider,* and *User* that extend the two imported ontologies and define the infrastructure that we describe in this chapter.

Figure 7-1 SUMI Ontology – Protégé Screenshot

Table 7-4 below provides a description of each introduced category in the SUMI ontology and explains which categories from GUMO and FOAF are adopted. In

addition, Appendix C presents the list of *DatatypeProperties* and *ObjectProperties* that exist in the SUMI ontology.

Category	Description
User	*User* uses the *userIsPerson* property to map this category with the *foaf:Person* category and describes SUMI users.
Provider	The category *Provider* describes all providers of user models. They are divided into two sub-categories: • ***Social Networking*** and • ***E-Commerce***.
AttributeCategory	*AttributeCategory* describes the various groups of attributes that can be retrieved from providers of user models using their APIs. These groups are categorised into the following categories: • ***GenericUserData*** • ***ServiceSpecificUserData*** • ***ServiceGenericGeneratedInfo*** • ***ServiceSpecificGeneratedInfo*** Every category in *AttributeCategory* represents one of the four categories of the proposed SUMI models' architecture, which was explained in 7.3.2.
Attribute	*Attribute* describes all attributes that can be retrieved from providers of user models using their APIs. Each attribute belongs to a group of attributes and every group of attributes is assigned to one instance of *AttributeCategory*. There are four kinds of attributes: • ***CommonUserInput*** attributes that belong to *GenericUserData* groups • ***SpecificUserInput*** attributes that belong to *ServiceSpecificUserData* groups • ***CommonGeneratedInformation*** attributes

	that belong to *ServiceGenericGeneratedInfo* groups • **SpecificGeneratedInformation** attributes that belong to *ServiceSpecificGeneratedInfo* groups *Attribute* uses the GUMO:SituationalElements category to map each attribute in SUMI with a situational element in GUMO using the *hasGUMORelation* property. For example, the attribute *Favourite_Movie* in SUMI is mapped with the attribute *Film* in GUMO.
DictionaryConcept	This category provides meaning to the attributes described in *Attribute* using the property *sameAsDictionaryConcept.* As the property shows, we map each attribute with dictionary concepts in order to provide explicit and well-accepted meanings of the attributes' definitions. Further explanation is provided below.

Table 7-4 Categories of the SUMI ontology

Defining meaning with Dictionary Concepts: Defining attributes' meanings using a dictionary that was written expressly for the purpose of explaining terms to people, means it is always possible to explain any ontological reasoning and relationships by showing the user the relevant dictionary entries (Apted *et al.,* 2004). SUMI follows the same approach and provides meaning to the various attributes, which are gathered from the various providers from the two examined domains using the online dictionary WordNet, found at *http://wordnet.princeton.edu/*, which groups nouns, verbs, adjectives and adverbs into sets of cognitive synonyms (called synsets), each expressing a distinct concept while interlinking with each other with conceptual-semantic and lexical relations. It maps each attribute with a dictionary concept to provide explicit and well-accepted meanings.

To achieve this, we followed a manual recursive three-step procedure with all four representatives that were presented in Chapter 5: Facebook, OpenSocial, Amazon and eBay, as follows:

1. Gathered the description of every attribute that could be retrieved using the offered API. These descriptions are provided by the providers and can be found on their respective websites.
2. Searched on the online Oxford English Dictionary and found all possible meanings that could be mapped with the provided descriptions gathered in step 1.
3. Identified the candidate that best met the provided descriptions and mapped each description with one dictionary entry.

At the end, each attribute that can be retrieved using the offered APIs had been mapped with a dictionary entry that can be used to remove any disambiguation and potentially allow successful machine-processing to take place. This approach helps avoid multiple interpretations by people living in different countries, with different cultures and different cognitive social models.

For example, the attribute "Interests", which is common to providers in the social networking domain, has been mapped with the dictionary meaning: "A thing in which one has an interest or concern". Another example is the attribute "Favourite Heroes", which is offered by OpenSocial websites, and refers to real people that a user may consider as important components in his/her life. By matching this attribute to a dictionary definition, we avoid wrong interpretations by users who may think that this attribute refers to imaginary superheroes, like Spiderman or Superman.

It is worth mentioning at this point, that although SUMI informs what constitute a valid input for each imported attribute, we acknowledge the potential scenario where a user may choose to ignore the description of a valid input and enter something different as value for an attribute.

Mapping different data models in the SUMI ontology: As we have already stated, for our research we mainly considered the two best representatives from each domain, and this is discussed in more detail in Chapter 5. Although the comparative evaluation we conducted revealed which attributes are common to both representatives from each domain, it also revealed some differences regarding the various sets of attributes each provider is offering to its users. However, what our research also revealed is that the four representatives could act as pioneers in our attempt to model

the various providers' internal structures from the two examined domains, since we discovered that a large part, if not all of the data models from other providers are subsets of the data models of their representatives. The different part of the data models is identified as being the assignment of different meanings and/or names to the same attributes.

Therefore, in order to tackle this issue, we mapped the four representatives, namely, Facebook, OpenSocial, Amazon and eBay, in our SUMI ontology, which can allow any new provider to use its representatives' data models as a template and define its own data model inside the SUMI ontology accordingly. For this reason, we offer the necessary tools, via a SUMI website, to allow any provider to map its own data model in SUMI manually in order to allow the successful exchange of user data towards SUMI. The need to encourage manual-administrative inspections, additions and updates of the data models from the providers themselves has been found to be a necessary and appropriate solution that is based on the assumption that a service administrator knows best what the service contains and how its internal structure should be defined (Abel *et al.*, 2009a; Abel *et al.,* 2009b). Furthermore, if a provider does not find its representatives' data models adequate to describe its own internal structure, it has the capability to extend the SUMI ontology and define the service's unique attributes, which will be categorized under the service-specific categories. More on this will be presented in section 7.3 where we will introduce the SUMI online service.

The following is a short hypothetical example of how a new social networking provider can map its data model in the SUMI ontology:

- The hypothetical provider "Watch-a-Movie" is a social networking website that allows users to watch movies online. It keeps a model for every user, and updates it accordingly when a user selects and views a movie.
- By using the SUMI service, which is presented in section 7.3, it can see its representatives' data models, i.e., Facebook's and OpenSocial's data models. It can see that both providers are offering an attribute called "movie" and it can read its dictionary definition and its description as given by each provider. It can then assign that attribute to its data model in order to inform SUMI that it is offering that attribute as well. It then

proceeds and defines the attribute's data type, for example, a string with 100 characters, and provides its own description.

- If the new provider is not satisfied with what its representatives are currently offering regarding the concept of movies, it can define a new attribute in the SUMI ontology and store it as a SpecificUserInput attribute, which will also automatically be assigned to the ServiceSpecificUserData category.

In addition to the SUMI ontology, which is responsible for mapping the various relationships between different data models, we maintain a SUMI relational database as well. The database is responsible for holding the most precious data retrieved from the various providers: the actual user information. In this way, we separate the data models from the data, and it assists us when it comes to exporting the user data from SUMI to educational personalisation systems. This is explained in the following sub-section.

7.2.4 Achieving Communication Among SUMI and Educational Systems

The final step for SUMI is to decide on the communication protocol that will be used for exchanging user information from SUMI to educational personalisation systems. For this reason, in section 3.2, we analyzed the two popular options for this purpose: Web Services and REST. As we described in Chapter 5, social networking providers offer RESTful APIs whereas e-commerce providers offer both REST and SOAP versions of their APIs. However, the requirement for exchanging user data from SUMI differs from the requirements of the providers' APIs: in contrast with the offered APIs, SUMI allows only the retrieval of user information and does not update or post new information back to the providers. Of course the latter is technically possible and sometimes desirable, especially in the case when the result of an educational activity performed by the user can offer some value to the same user when posted back to the social networking or e-commerce provider, but that yields further research, which is out of the scope of this PhD, to identify potential contribution when such action is allowed and explore whether users would prefer to have such option available for use. By researching the literature and analyzing the API versions that are available from the representatives of the two examined domains, it becomes clear that, although Web services fully satisfy the needs for such an infrastructure, in terms of achieving communication between providers across different domains, it is an advanced solution

to an ‘easy’ problem. A more ‘relaxed’ approach could fit perfectly the requirements enabling communication to take place among providers and the exchange of user information between them and educational personalisation services. Thus, SUMI adopts the REST protocol, an easy, effective and efficient approach, which is also offered in all four representatives of both examined domains. Every attribute inside SUMI is represented with a URI that can be accessed via HTTP methods in order to retrieve the attributes’ values. That URI is generated ‘on the fly’ and only when the user-owner approves the export transaction. For example, a user may post the value “football” as one of his/her Facebook interests. When that user imports his/her Facebook interests in SUMI, using the offered SUMI service, SUMI will assign a URI to that attribute and store its value inside the SUMI database. When, later on, the user exports that attribute towards an educational system, for example, AHA!, then SUMI will expose that URI to AHA! and allow a GET request to be performed on the attribute *interests* for that specific user. In this way, AHA! will retrieve the value “football” and process it accordingly in order to provide a richer personalisation service to that user.

For this service, we have found two adequate tools that can serve SUMI for this purpose, namely, D2R Server and sqlREST, which were presented in section 4.4.2. Both solutions are adequate for achieving communication among SUMI and educational personalisation systems. Since SUMI favours the RESTful approach, we have inclined towards the use of the sqlREST tool for exporting the contents of the SUMI database, which holds the user data, in the form of URIs that can be accessed by the educational systems via HTTP GET requests only.

7.3 SUMI: The Proposed Service

While we presented the proposed infrastructure in the previous section, we have not yet addressed the issue of scrutability and privacy. In a recent article (BBC News, 2009d), the question of who polices Facebook was raised. The number of registered users grows continuously, but what happens to the data that are posted and shared on its platform and how do users feel about how their data are treated? A conclusion that promotes the notion of passing control of information to its owners was supported. Although the infrastructure provides the central feature of our work, it must be scrutable in a way that will allow users to have control and understanding of their

imported user models, while offering them the tools needed to express their privacy preferences, as stated by our research questions and hypotheses for scrutability and privacy. In addition, there is a need to provide a platform users can take advantage of in order to access the proposed infrastructure while enjoying the offered scrutability and privacy privileges.

For that purpose, we have designed and implemented a *prototype* SUMI online service, which can found at *http://www.mysumi.org*, allowing users to customise the settings for their SUMI collections and define their preferred privacy settings. The service exposes the proposed infrastructure to SUMI users and allows them to take control of any social networking and e-commerce model they decide to import into SUMI. In addition, this service allows providers of user models to define their own data models, as explained in the previous section, and allows educational personalisation systems to subscribe to any provider registered in SUMI in order to express interest in receiving user data from them.

This service helped us to evaluate SUMI and demonstrate it in various conferences (Kyriacou, 2008; Kyriacou, 2009; Kyriacou and Davis, 2008; Kyriacou *et al.,* 2009a; Kyriacou *et al.,* 2009b). In this section, we describe the main features of this service; its main purpose is to act as the platform between the infrastructure and the various SUMI stakeholders, namely, users, providers of user models, and educational personalisation systems, while focusing on providing scrutability and privacy privileges.

7.3.1 Making the Infrastructure Scrutable

From the Users' Perspective: Every collection of models in SUMI is a representation of an integrated variety of social networking and e-commerce user models, which have resulted from user interaction with various social networking and/or e-commerce services. The SUMI models' architecture has been divided into four categories, reflecting the two examined domains, as described in section 7.2.2. Every user has absolute control over his/her SUMI collection of models and the first decision they need to make is to determine which models they would prefer to integrate into SUMI.

After users decide to add one of their models to SUMI, they have the ability to import the content of the previously added models in a customised way. SUMI

provides the option between dynamic information, meaning real-time HTTP GET requests and retrieval of real-time data from the user model provider, and static information, meaning the cache copy that was taken when the last dynamic import request was generated by the user and kept inside the SUMI database, and that will be retrieved using SQL queries. Busy network traffics that can be observed when the various providers' APIs are overloaded by millions of simultaneous requests is one of the reasons that users could take advantage of the latter option. Users can instruct SUMI which attributes from every previously added model they would like to import into SUMI, and which attributes they would prefer SUMI to skip when the import procedure is initiated.

The most important feature of the SUMI service is the users' ability to export a part of any of the previously added models towards any subscribed registered educational personalisation service they prefer. This has been achieved by adopting the SUMI export protocol, which adopts the RESTful approach. In case a user decides to export his/her model information to a subscribed service, the 'must' attributes, which are defined by the subscribed educational service, are included by default in the transaction. More on the subscription process is presented below.

From the Providers' Perspective: Providers of user models can manually describe their data model inside SUMI in order to:

- achieve a connection with SUMI, since the providers will "introduce" themselves to SUMI and their data models will be defined and recognised when a data exchange is needed.
- allow users to import the information they post and share on the providers' websites without having to deal with the technical complexity required for initiating an import procedure.
- enable educational personalisation systems to discover them and express interest by subscribing to them. In this way, users can view the subscriptions from the various educational systems and decide toward which of these they will initiate an export procedure from SUMI.

New providers can manually map their data models onto any of the two representatives of their domain, either Facebook and/or OpenSocial for the social networking domain, or Amazon and/or eBay for the e-commerce domain, using the SUMI ontology and the SUMI models' four-category architecture, which is reflected inside the service. If none of the already mapped data models can fully satisfy the needs

of a new provider, then it has the ability to define new attributes and/or attributes lists, which will automatically extend the SUMI ontology.

From the Educational Systems' Perspective: Educational personalisation systems can subscribe to providers of user models, inside SUMI, in order to express their interest, to SUMI users, in receiving models, or parts of models, from them. This approach was presented in section 3.2 where we described various solutions for achieving interoperability in the educational domain, and we think it is adequate to use SUMI for the same purpose. By subscribing to a provider, the subscriber can define explicitly which attributes, from the selected model are 'must' in order for the transaction to take place and which attributes are not so crucially important for the transaction, but which the subscriber would like to receive as well ('want' attributes). In addition, the educational system is required to describe inside SUMI the reason for requesting selected attributes from selected user models. Users have the final word for deciding whether the transaction will take place. Once a user has imported a model into SUMI from a provider to which a service has subscribed, he/she gets to see which attributes were defined as 'must' for the transaction and which ones were listed as 'want' attributes by the subscriber. The user has the final choice to proceed with the export-transaction of his/her information towards the subscribed educational system, or 'pull the plug' and decline to export any data without having to provide any explanations for his/her decision.

7.3.2 Safeguarding Privacy of Information

Safeguarding the privacy of users' and providers' information that is imported or created inside SUMI is a very important requirement in our proposed solution. Using the offered online service, users and providers access the SUMI ontology and the SUMI database where information is updated in various ways:

- Users import their personal information from social networking and e-commerce providers, and until that information is exported towards subscribed educational systems, we understand that it is SUMI's responsibility to safeguard it.
- Providers define their data models inside SUMI and although this issue is not as important as user data, it needs to be addressed with caution to ensure that the applied solution will treat this sensitive information discretely in

order to encourage new providers to share their own data models in SUMI and increase the list of available providers that allow users to gain control of their models via SUMI.

- Educational personalisation systems express their interest in receiving data from various registered social networking and e-commerce providers that can be found in SUMI by subscribing to them. These subscriptions allow users to see which systems are interested in which providers. Admittedly, this kind of private information requires a level of attention for its privacy inside SUMI to be properly safeguarded. Failing to acknowledge its importance may lead to a lack of trust from educational systems towards SUMI, and that would break the chain of events that should happen in order for SUMI to offer the maximum of its potential to all three stakeholders – users, providers of user models and educational personalisation systems.

From the Users' Perspective: In section 3.4, we acknowledged the argument, as stated by (Kobsa, 2007), that there is no perfect trade-off between offering personalisation services and adopting privacy policies. It is assumed that various small enhancements should be introduced to users and the best alternative, based on the requirements at a given situation, should be selected. Furthermore, it is recognised that people become more comfortable and confident in sharing when they gain control of their posted information. It is also important for users to comprehend how their information is treated by the providers whom they trust enough to post on and share in their platforms (InformationWeek, 2009a). For the purposes of this research, three privacy settings were designed to allow SUMI users to control how other SUMI users would access their SUMI collection of models. These three statuses were inspired by two lawsuits that forced Facebook in the area of social networking (The Register, 2009) and Amazon in the area of e-commerce (ChannelWeb Network, 2009a) to review their privacy policies and adopt a different approach that passes the privacy control of the posted data to the users/owners (TG Daily, 2009; PC World, 2009b; DailyTech, 2009; BBC News, 2009b; ChannelWeb Network, 2009d). It is our belief that any posted information located in social networking and e-commerce user models can be assigned to one of the following three privacy categories:

- Category 1: users do not particularly care if these data are accessed by third-party systems and/or other users; therefore, they do not spend time and effort to keep these data private. For example, the Facebook and Twitter

status is a piece of posted data regarding which users do not particularly take action to keep private from third-party applications and/or their peers.

- Category 2: users care which third-party system and/or other user accesses these data, but they do not particularly invest time to safeguard them. For example, wish lists on Amazon do not receive special privacy attention from users, but users do care who accesses them.
- Category 3: users care which third-party system and/or other user accesses these data and care enough to find ways for safeguarding them and keeping them private. For example, posted photos on social networking sites most of the time receive high privacy attention from users, since they care which of their peers may gain viewing rights to their shared photos.

With that in mind, we designed a three-level privacy policy that allows users to obtain control regarding how their imported models are accessed by third-party systems and other SUMI users. More specifically, SUMI users have the choice among three privacy statuses that they can apply to their models as a whole, or to any of the four categories of any of their models:

- *public* – others can see that a public user model exists and anyone can view its content,
- *private* – others can see that a user has a private user model, but they have to place a request with the model's owner to view the model's content, and
- *hidden* – when a user model is set to hidden, others cannot see that the user model exists; therefore, the model's content is accessed only by the model's owner.

In addition, users are allowed to send and receive viewing requests to/from other SUMI users for their *private* models/parts of models. This will be discussed further in the next chapter where we present our evaluation framework.

As we explained in the previous section, all user data are stored inside a SUMI database that is maintained by a professional host, and they are retrieved and processed only when the user-owner initiates an action, via the SUMI service, which generates a database command for accessing his/her data. With this approach, we offer a basic level of security to the user information that is imported into SUMI.

From the Providers' and Educational Systems' Perspective: As we stated in the previous section, we maintain a SUMI ontology and a SUMI database separately from each other. The various data models of the providers of user models are stored

inside the ontology and the various subscriptions to educational systems are stored inside the database. With this approach, we can achieve improvements in performance, speed and security compared to other approaches, which require information storage in one common place or the need to maintain the ontology and database at a common location. Although we acknowledge that this is an important issue for further research, we have not conducted a more in depth investigation since it is not one of the research questions for this thesis.

Chapter 9 describes and presents our evaluation framework and results for scrutability and privacy. We tested the proposed infrastructure and the prototype SUMI service with potential users and evaluated how they would react and behave when given the scrutability and privacy privileges that were presented in this section.

7.4 SUMI Use Cases

Section 7.4 puts everything that was presented in sections 7.1-7.3 into perspective by offering a graphical representation of a use case to show how SUMI can be used, in order to help the reader to comprehend fully the proposed infrastructure and service. Figure 7-2 demonstrates how the SUMI service, marked as *SUMI website*, acts as the platform that users, *Providers of User Models* and *Educational Systems,* can use to access the proposed infrastructure, which is represented by *SUMI Ontology* and *SUMI Database* in the figure. In this example, we use the social networking site Facebook and the educational personalisation system AHA! to demonstrate SUMI's offered functionality and provided options.

- *Facebook,* which is a subclass of the class *Providers of User Models*, communicates with the *SUMI website,* via Facebook Connect or via a Facebook Application, when a SUMI user initiates the import of his/her information from the site.
- SUMI makes sense of the information that is retrieved from Facebook because *Facebook*'s data model is already defined inside SUMI. As we have explained, when a new social e-network registers with SUMI, it has the option to map its own data model to that of any of its representatives and/or extend the *SUMI Ontology* if necessary.
- The SUMI service communicates with the infrastructure, defined using *SUMI Ontology* and the *SUMI Database*, and retrieves the necessary

information in order to serve the SUMI user who has initiated import of his information from Facebook. If necessary, the *SUMI website* can update the database and the ontology when users and/or providers post information back to SUMI accordingly.

- Finally, the educational system *AHA!* uses the *SUMI website* to subscribe to *Providers of User Models* and SUMI uses the REST approach to export user information from the infrastructure towards it, of course after direct consent from the user/owner has been obtained.

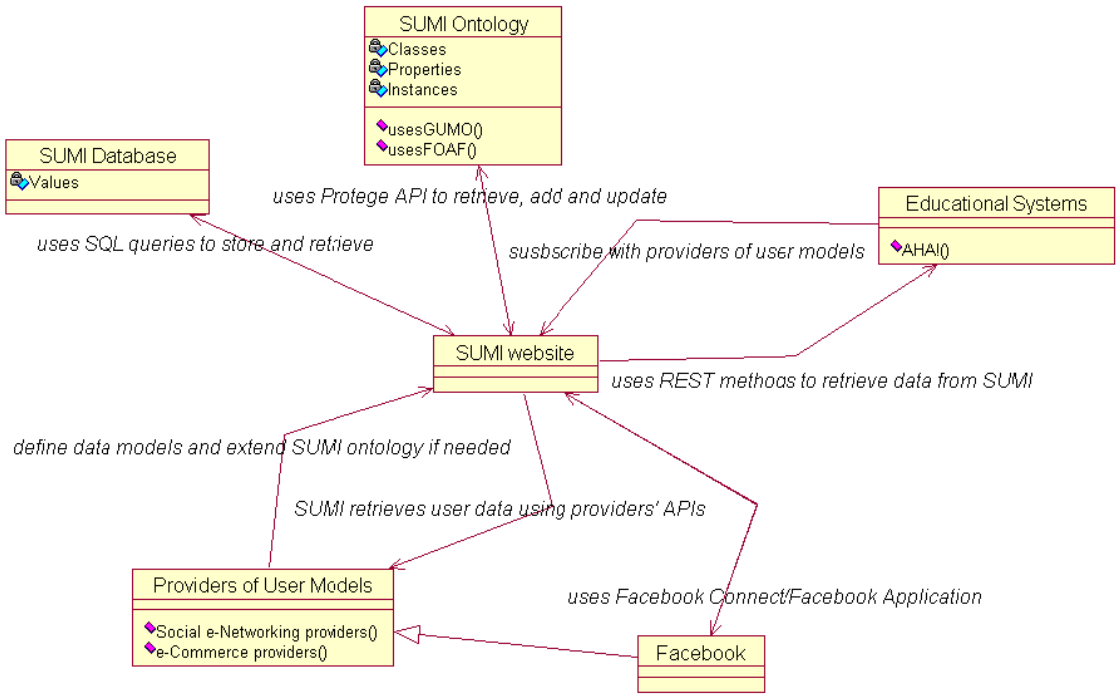

Figure 7-2 Graphical representation of a SUMI use case

Furthermore, in order to assist the reader to comprehend further the potential contribution such an infrastructure can make to the area of Lifelong User Modelling, we offer three more realistic hypothetical use cases that demonstrate SUMI's usefulness.

7.4.1 Generating Personalised Examples

Adaptive courses or lessons can be created using educational personalisation systems like AHA! (AHA! Home Page) and InterBook (InterBook Home Page). The content is adapted based on students' knowledge, which is modelled using the overlay model, and browsing behaviour, which is recorded while the user goes through the

online course, and affects the amount of knowledge assigned for the various concepts that are introduced during the adaptive course. SUMI can enrich personalisation in this case by allowing the generation of personalised examples 'on the fly' that can be offered during the adaptive lessons. These personalised examples can be based on the students' shared information in social networking and/or e-commerce websites, for example, interests posted on Facebook, activities posted on MySpace, the latest movie DVD bought from eBay, the most recent music CD purchase from Amazon, etc. The generation of personalised examples based on students' shared information allows them to comprehend the teaching material better than just going through a default example in every lesson. This can potentially help the students to engage more deeply with the taught subject than before. In addition, by automatically fetching this posted information from social networking and e-commerce websites, the teacher can avoid the time-consuming process of asking each student to fill out (registration) forms beforehand in order to gather information about the students' interests, activities, etc.

7.4.2 Forming Student Groups

In every degree in all universities around the world, there are times when students are required to work in groups. Extensive work has been conducted in academia on identifying the best ways of forming groups in such cases, resulting in the introduction of Group Formation systems (Ounnas *et al.,* 2008). These systems require input, normally from the teacher an administrator, in order to calculate the best way to form groups of students, depending on the circumstances. The teacher has to define the acceptance criteria and any restrictions before any group can be formed. SUMI can assist in this case as well by providing as input for these systems the students' posted information on social networking and e-commerce websites. By automatically retrieving, in real-time, students' hobbies, favourite bands, nationalities, latest book purchases, etc., group formation systems can perform more accurate calculations and produce better output than before. Of course, the instructor must have the final administrative say to approve the resulting groups. We think that this approach would be a preferable solution to the traditional method of asking each student to complete a questionnaire in order to gather personal information that has already been posted and shared on social networking and e-commerce websites.

7.4.3 Recommending Fellow Students and Social Activities

Throughout their academic careers, students participate in social activities by joining societies, sports teams and fellowships. It has been acknowledged that this assists students in relating to and familiarising themselves with their surrounding environment, which can also lead to academic improvement (BBC News, 2009c). SUMI can bring some value to this case as well. By retrieving and analysing the huge amount of information that is constantly posted and shared on social networking and e-commerce websites by millions of students, it can enable social recommender services that can work in the same way as commercial recommender services (Flixster Home Page, Netflix Home Page) to suggest social activities with which every student can engage and participate in if they wish. The algorithm behind such a service can take advantage, via SUMI, of the user data that are currently offered by social networking and e-commerce providers, and recognise similarities in students' interests, activities, origins, etc. Consequently, it can recommend social activities and/or social groups for students to join, while introducing them to like-minded fellow students or those with similar interests. An important issue to consider in this case is the limitation of scrutable UM services, like SUMI, when offering recommendation services. As we have explained above, user consent is required before any data is used by SUMI. That contradicts with the methods used in popular recommender systems, like Amazon and eBay, where the algorithm works silently in the background using data that users share on their platforms without direct consent from the data owners and could result in limited data sets being used in the recommendation algorithm with potentially limited results returned back to the users. Having said that, we argue that scrutable UM solutions could also be considered as the way forward when developing recommender systems. As presented in section 3.4, users do care about their privacy and the hidden use of their data by commercial providers contradicts with the principle of passing ownership of information to its owners. Although acquiring users' consent will require additional steps and extra effort from the developers' side, it will ensure users' control of their own data, used in the recommendation algorithm, and inspire their trust on the recommender system.

7.5 Summary

In Chapter 7, we have presented our proposed solution to an identified problem in the area of LUM. The solution consists mainly of a user modelling infrastructure that helps users to gather their various social networking and e-commerce models in one place while handling the various challenges regarding acquiring the user data, mapping the several different data models of the providers of user models and exporting user information, after obtaining users' direct consent, towards subscribed educational systems. In addition to the infrastructure, we have presented a prototype SUMI service that transforms the infrastructure into a scrutable one by offering scrutability and privacy privileges to its users, providers of user models and educational personalisation systems.

Chapter 8 Evaluation Design and Results for Interoperability

8.1 Introduction

In this chapter, we present the design for the evaluation of the interoperability aspect of SUMI. We test the infrastructure and examine the feasibility of what we proposed in Chapter 7 by testing the hypotheses that were presented in section 6.3.1. To evaluate interoperability, we implemented a real-case scenario, using the Facebook platform and the personalisation educational system AHA! that we designed according to our proposed solution, and we incorporated it into the SUMI prototype service and into a Facebook application. Furthermore, this chapter also presents the results obtained from the four interoperability tests that we conducted, and discusses the various lessons learned and the conclusions drawn.

8.2 Evaluating Interoperability in SUMI

Evaluating interoperability is a technical question and does not necessarily involve users. However, to evaluate interoperability between (a representative from) providers of user models and (a representative from) educational personalisation systems, via SUMI we requested our users' help and co-operation, since users' consent is required for retrieving and processing personal data. Therefore, we incorporated the four interoperability tests into the two user evaluations for scrutability and privacy, which are presented in Chapter 9. The desired outcome, following our research approach for interoperability, which was explained in sections 6.4 and 7.2, was to achieve a connection with a social networking service in order to allow a successful

exchange of user data to occur between that service and an educational personalisation system, via SUMI, which we believe could act as a representative example for testing interoperability among current social networking and e-commerce services, as presented in Table 7-1, and the selected educational personalisation system.

As discussed in Chapter 7 we proposed a four-category SUMI models' architecture and a SUMI ontology to manage the complexity of the providers' data models. In addition, we favoured the use of the REST architecture for exchanging user information between SUMI and the educational systems. We tested this proposed solution by attempting to create a real-life scenario, one that was described in section 7.4, and investigated the possibility of exchanging user data between Facebook and AHA!.

As explained in Chapter 5, two methods can be used to retrieve user information from the Facebook platform. In order to examine both methods, we implemented a Facebook application and used Facebook Connect to bring the Facebook platform into the SUMI service. In addition, once user data were retrieved into SUMI, we tried to pass user information to AHA! using the REST method to make use of it in the context of delivering personalised examples during an adaptive lesson on HTML. Furthermore, we evaluated whether our SUMI ontology, which holds the information about the Facebook data model, could add any value to the raw user data when presented to the data owners. We explain all four tests in detail below while analysing the resulting conclusions from each one.

8.2.1 Retrieving user data via the SUMI Facebook application

As we explained in Chapter 5, users can retrieve their information from Facebook in two ways: via a Facebook application built inside the platform, and via a website implementing Facebook Connect, which brings the Facebook platform into that website. To evaluate the first case, we implemented a Facebook application, found at *http://apps.facebook.com/mysumifirstapp/*, in order to examine the process of obtaining user information this way. Furthermore, we focused not just on evaluating our solution for interoperability when building this application, but also targeted user engagement with the SUMI environment and mentality. For that reason, we first implemented the feature for achieving interoperability and then enriched it with other features that allow users to engage further with SUMI.

The first thing users come across when visiting the application is the welcome page, which explains to them what this application is about (Figure 8-1). This welcome screen introduces the four-category SUMI models' architecture and provides a graphical representation of the SUMI vision. In addition, it provides all navigation options as tabs, which are explained below.

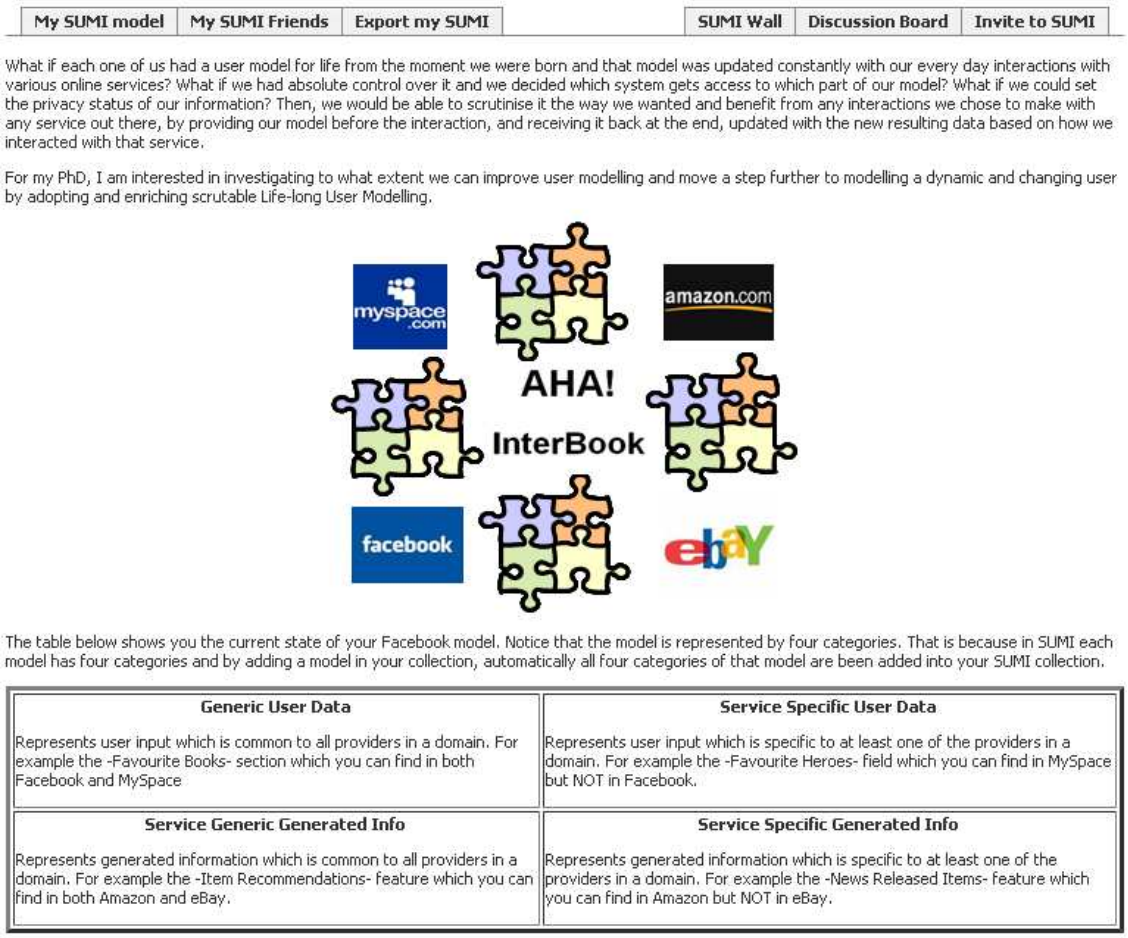

Figure 8-1 Welcome screen for the SUMI Facebook application

The SUMI Facebook application consists of five parts, which are presented as tabs.

1. *My SUMI model*: By selecting this option, SUMI displays the user's current Facebook model in the four-category SUMI models' architecture, as displayed in Figure 8-2. This helps the users to comprehend better the proposed architecture by visualising it with real-time data. In addition, the user data displayed is copied into the SUMI database, unless instructed otherwise by the user/owner. This is discussed in more detail later on.

Generic User Data:	Service Specific User Data:
About Me: Hmmm... Activities: Interests: NBA Basketball, English Premier League, Swimming, User Modelling, Semantic Web Favourite Books: Favourite Movies: Seven, Forrest Gump, Two For The Money, The Dark Knight Favourite Music: 3 Doors Down, Nickelback Favourite Quotes: Favourite TV Shows: Lost, The Office (USA), The Apprentice (UK & USA), Dragons' Den (UK) 24, Desperate Housewives, Boston Legal, Burn Notice	Interested In: Number of Notes: 6 Number of Wall Posts: 242 Profile Last Updated: 0
Service Generic Generated Information: No data	**Service Specific Generated Information:** No data

Figure 8-2 'My Facebook model' presented in the proposed four-category SUMI models' architecture

2. *My SUMI Friends*: In addition to their own model, users have the option of viewing their friends' Facebook model in the proposed SUMI architecture. Figure 8-3 demonstrates how this is presented in the application.

Generic User Data:	Service Specific User Data:
About Me:	Interested In:
Activities:	Number of Notes:
Sailing Walking Camping	1
Interests:	Number of Wall Posts:
Wine Tasting	3
Favourite Books:	Profile Last Updated:
Catch 22 Heller	0
Favourite Movies:	
Local Hero	
Favourite Music:	
See my Last.Fm profile at http://www.last.fm/user/hcd99	
Favourite Quotes:	
Favourite TV Shows:	
I Never Watch Tv	
Service Generic Generated Information:	**Service Specific Generated Information:**
No data	No data

Figure 8-3 'My Friend's Facebook model' presented in the four-category SUMI model's architecture

3. *Export my SUMI:* This option was provided to SUMI users to evaluate our proposed export protocol, which adopts the REST protocol; it is presented in section 8.2.3.
4. *SUMI Wall and Discussion Board:* SUMI also invites users to express their thoughts and socialise with others while using this application by providing a wall where users can state their own suggestions regarding SUMI (Figure 8-4). In addition to the SUMI wall, users have the option of creating a topic about an issue regarding SUMI, where they can participate in a deeper discussion with their peers (Figure 8-5).

Figure 8-4 The SUMI Wall

Figure 8-5 The SUMI Discussion Board

5. *Invite to SUMI:* Finally, SUMI takes advantage of the Facebook invitation feature, which allows a user to ask his/her friends to join him in a social activity, in this case adding the SUMI Facebook application (Figure 8-6)

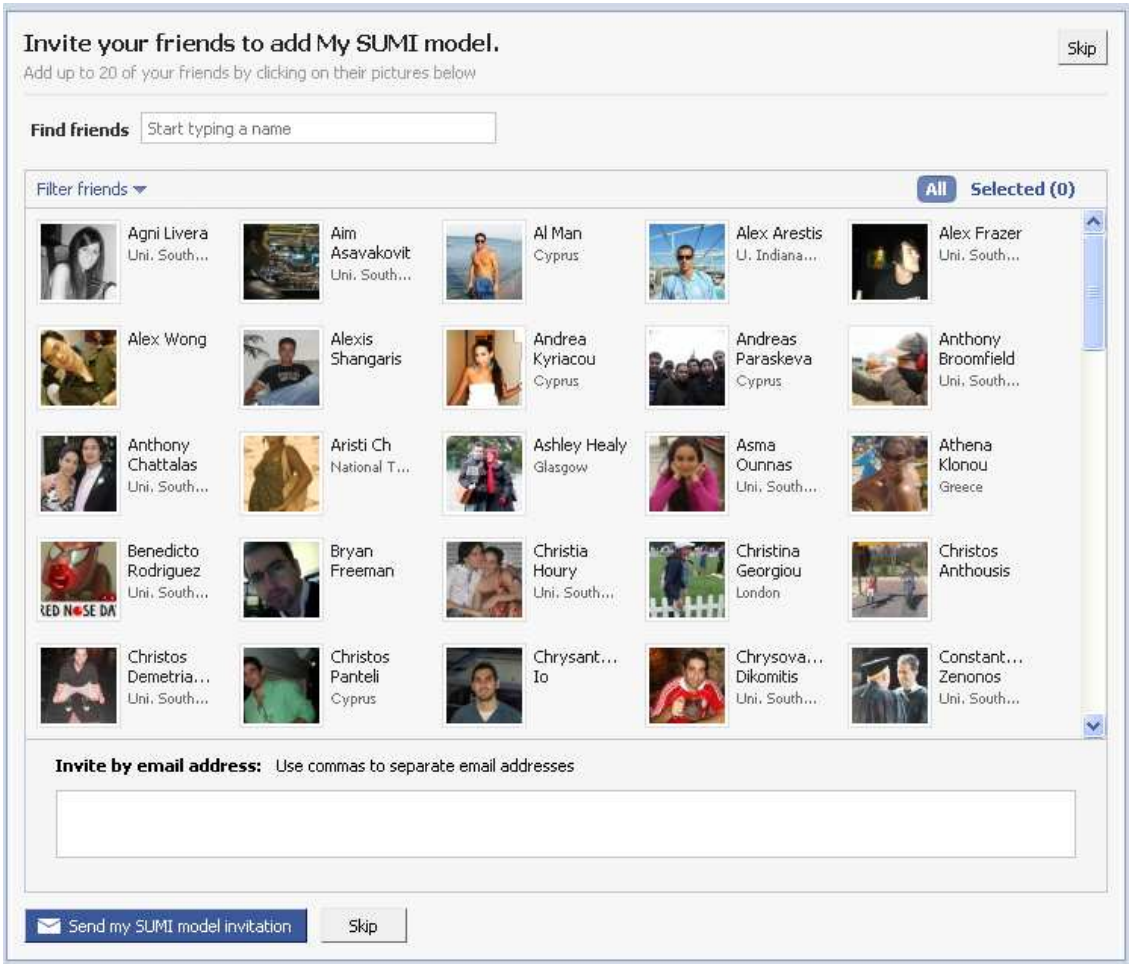

Figure 8-6 Inviting friends to add the SUMI application

Conclusions: Using the available Facebook API to implement the SUMI Facebook application caused us no problems in achieving our objective, which was to retrieve user information from the Facebook platform and move it to SUMI. By clicking the option *My SUMI Model,* the users initiated the transfer of their data from Facebook to SUMI without any problems. The first case for achieving interoperability between Facebook and SUMI passed the test. In addition, it is worth mentioning the several social "bonuses" that can be offered by this application; allowing users to discuss SUMI and to invite their friends can only help in promoting and expanding the SUMI users' network.

8.2.2 Retrieving user data via Facebook Connect

Secondly, we incorporated Facebook Connect into the SUMI website, found at *http://www.mysumi.org*, and analysed the complexity of acquiring data from Facebook in this way.

The SUMI website was introduced to the users to evaluate the proposed scrutability and privacy privileges. For that reason, we conducted two user evaluations, the frameworks of which are presented in section 8.3. During the two evaluations, we also took the opportunity to evaluate our proposed solution for interoperability without explicitly stating that intention to our users to avoid any potential confusion and to keep them focused on the proposed scrutability and privacy privileges.

During the two user evaluations, task 2 asked users to import the content of the previously added models. For the users who had previously added Facebook, we used the API for Facebook Connect to retrieve their data from the Facebook platform. Users were requested to login to their Facebook account (Figure 8-7) and initiate the retrieval of their data from Facebook. After the users' direct consent had been obtained, SUMI retrieved the data and displayed them using the proposed four-category SUMI models' architecture (Figure 8-8). In addition, when the user data were retrieved from Facebook, they were copied inside the SUMI database, unless the user/owner instructed otherwise. In this way, we evaluated whether transferring user data from the Facebook platform to an external website would raise any problems. In addition, we took the opportunity to expose our users once again to the proposed SUMI models' architecture.

Figure 8-7 Asking users to login before retrieving their Facebook data

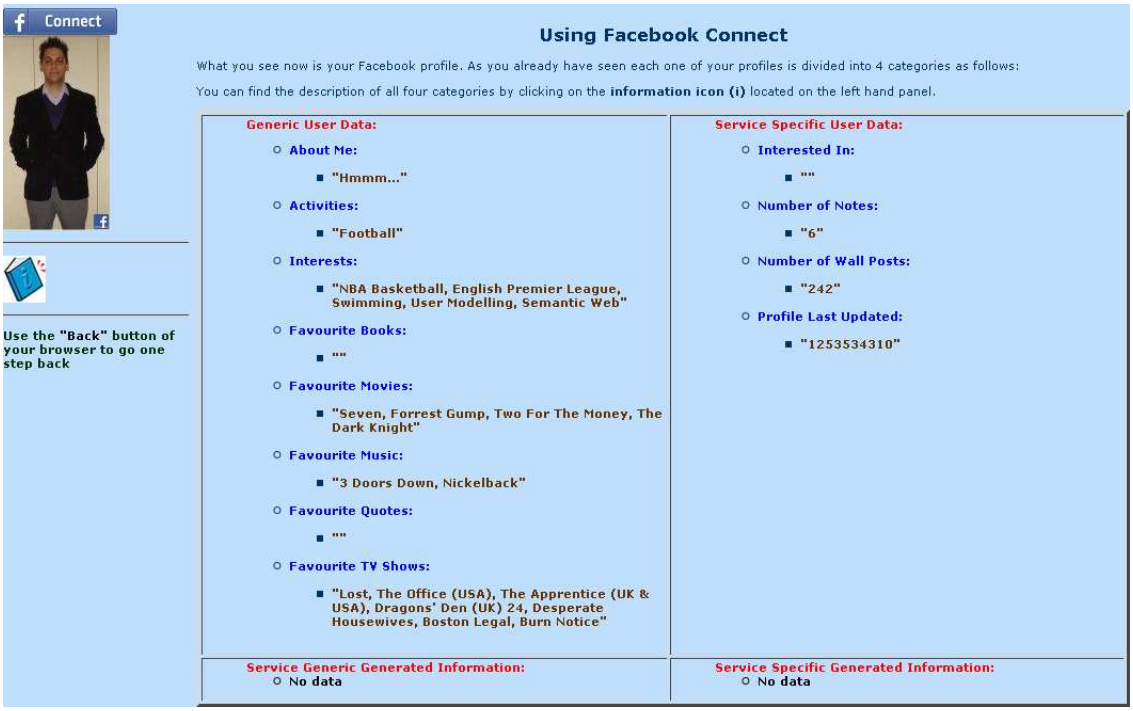

Figure 8-8 Retrieving user's Facebook data and displaying them in the four-category SUMI model's architecture

Conclusions: Similar to the first case, using Facebook Connect to transfer user information from the Facebook platform to the SUMI website, and more specifically to the SUMI database, was achieved without any kind of problems. The second case for

retrieving user information from the social networking site and copying it to SUMI, passed the test as well.

8.2.3 Exporting data using the REST protocol

To evaluate our proposed export protocol, we asked users to use either the SUMI Facebook application or the SUMI website and, taking advantage of the available options, to export their data to a subscribed educational system, in both cases AHA!. Users were required to inspect and approve the transaction details before permitting the exchange of their information. We describe each case below.

For the SUMI Facebook application, we offered an option named *Export my SUMI,* which is probably the most important option offered in the SUMI Facebook application. By visiting this option, users were allowed to export their personal Facebook data towards AHA!. This option is the one used for evaluating our proposed solution for achieving interoperability between the Facebook platform and SUMI, and then from SUMI towards AHA!. As Figure 8-9 shows, AHA! has (hypothetically) subscribed to receive the Generic User Data category of a user's Facebook model.

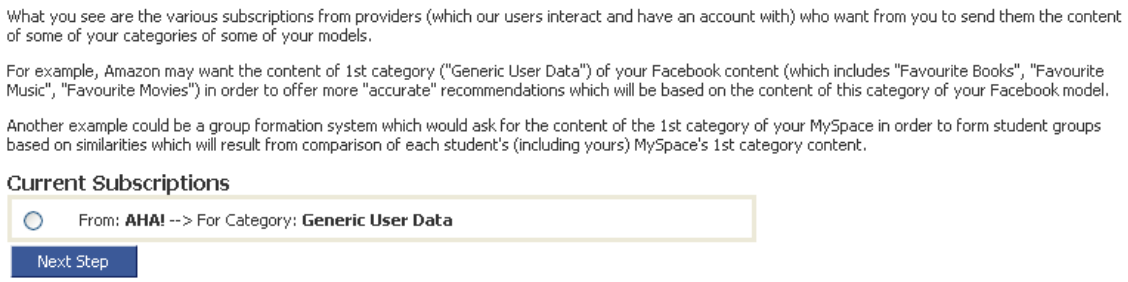

Figure 8-9 List of current educational systems' subscriptions for receiving user data

The second step of this option allows the user to inspect and alter what the educational system will receive, as shown in Figure 8-10. As can be seen in this case, AHA! has marked two attributes as "Must", meaning that the user must accept these two attributes for the export transaction otherwise it would not be interested in receiving any data at all. In addition, the user is also presented with a list of the "Want" attributes that AHA! has stated that it would like to receive only if the user decides to export towards it.

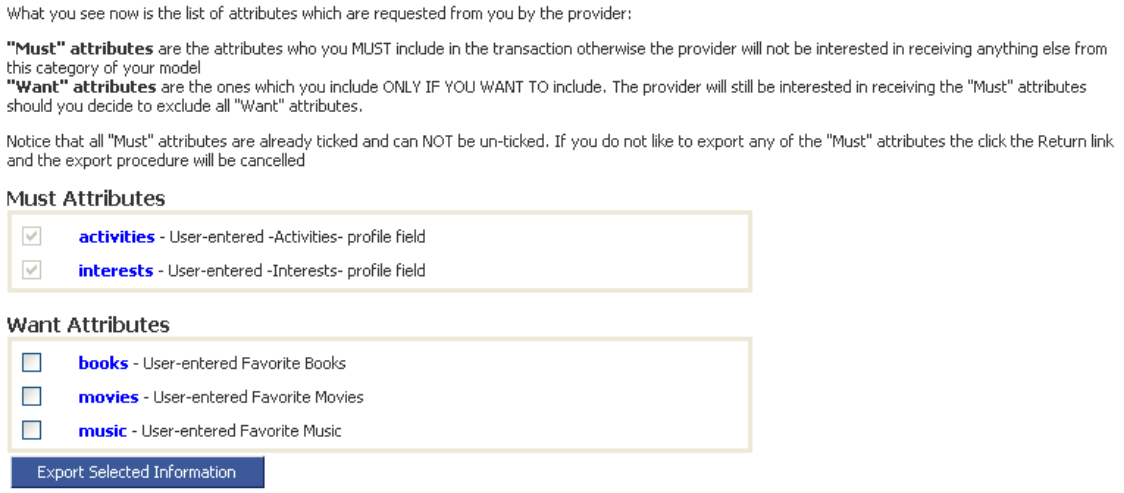

Figure 8-10 Second step of the export transaction. List of "Must" and "Want" attributes

Finally, once the user has decided what to export towards AHA!, the SUMI application assigns a URI to each attribute, via the sqlREST tool, which was presented in section 7.2.4, and notifies the educational system that it can now perform a GET request on these URIs to retrieve those attribute's values. Figure 8-11 shows a scenario where the user has exported the "Must" attributes *activities* and *interests*, and none from the "Want" attributes. In addition, Figure 8-12 demonstrates how the assigned SUMI URI for the attribute *activities* exposes the corresponding attribute to the subscribed educational system.

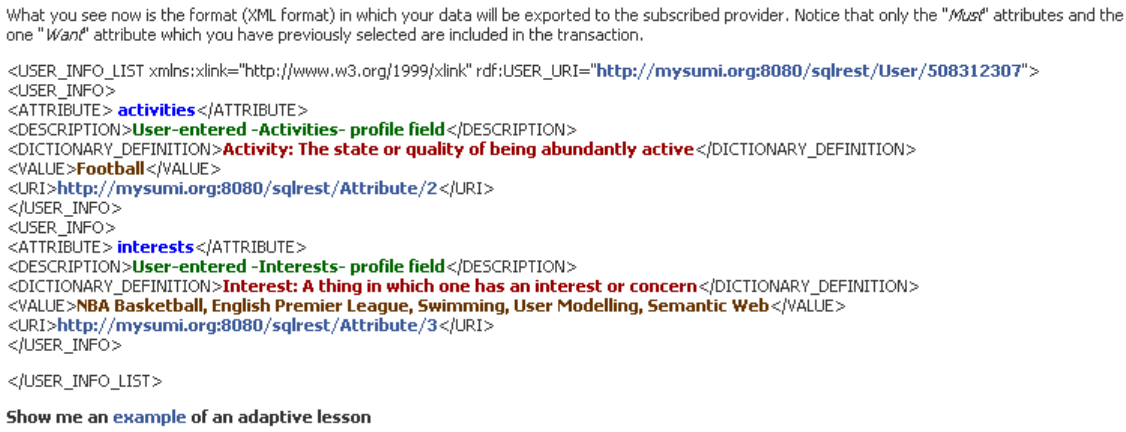

Figure 8-11 SUMI assigns URIs to the exported attributes

```
<Attribute>
   <ID>2</ID>
   <description>User-entered -Activities- profile field</description>
   <dictionaryDefinition>The state or quality of being abundantly active</dictionaryDefinition>
   <name>activities</name>
   <value>Football</value>
   <userID xlink:href="http://mysumi.org:8080/sqlrest/User/508312307/">508312307</userID>
   <flag>1</flag>
   <store>1</store>
</Attribute>
```

Figure 8-12 URI exposes the information regarding the Facebook attribute *activities*

Furthermore, SUMI provides an example to inform the user how his/her exported data would be used by the subscribed educational system AHA!. In this case, an adaptive lesson on HTML unordered lists is presented to the user, where the user goes through the teaching material (Figure 8-13) and then is presented with a personalised example. That example was generated based on the user's Facebook activities and interests; both were attributes that the user has exported towards AHA!. In this case, the personalised example demonstrates how a football team's 4-4-2 formation can be presented using HTML unordered lists (Figure 8-14), since the user has the value "Football" posted as one of his interests.

Adaptive Lesson: HTML Lists

While simple in concept, lists can be very powerful in execution. There are three types of lists: unordered lists, ordered lists, and definition lists. The following lesson attempts to introduce unordered lists and provides a personalised example at the end:

Unordered Lists

The term "unordered list" may be a bit unfamiliar to you, but odds are you've heard of the "bullet list." That's exactly what an unordered list is -- a list of items, each one preceded by a "bullet" (a distinctive character; typically, a small black circle).

The list begins and ends with the tags `<UL>` and `</UL>` respectively. Each item in the list is marked using the `<LI>` tag, which stands for "List Item." `<LI>` has a corresponding `</LI>`, but this closing tag is not required to end the list item (although you could use one if you really wanted to). You can use as many list items as you like, up to your browser's built-in maximum, if any.

Here's the markup for a simple list:

```
<UL>
<LI>Monday
<LI>Tuesday
<LI>Wednesday
<LI>Thursday
<LI>Friday
</UL>
```

If you load an HTML page (webpage) containing the markup above, you would see :

Monday
Tuesday
Wednesday
Thursday
Friday

Personalised Example:

The following personalised example has been generated after a close look of the attributes "Activities" and "Interests" of your Facebook profile which you have just exported to AHA!.

If you like to see a different example, please go and change either your Facebook "Activities" and/or your Facebook "Interests" to one of the following sports: Football, Basketball or Volleyball. Then import the content of this profile again into SUMI.

This is because AHA currently personalises the example you see below based on these three sports. If you do not have any of these three keywords in your Facebook profile, then the default example (To do list) is presented.

Figure 8-13 Example of an AHA! adaptive lesson on HTML unordered lists with a generated personalised example

```
<ul>
     <li>Left Back</li>
     <li>Center Back</li>
     <ul>
          <ul>
               <li>Left Winger</li>
          </ul>
               <li>Defensive Midfielder
                    <ul>
                         <ul>
                              <ul>
                                   <li>Striker</li>
                                   <li>Striker</li>
                              </ul>
                         </ul>
                    </ul>
               </li>
               <li>Defensive Midfielder
                    <ul>
                         <li>Right Winger </li>
                    </ul>
               </li>
          </ul>
      <li>Center Back</li>
      <li>Right Back</li>
</ul>
```

Here's how it looks when a browser displays it:

- Left Back
- Center Back
    - Left Winger
    - Defensive Midfielder
        - Striker
        - Striker
    - Defensive Midfielder
        - Right Winger
- Center Back
- Right Back

Figure 8-14 AHA!'s generated personalised example based on a user's Facebook activities and interests.

The second case involves the SUMI online service; users were asked to export their data using the provided options, which were the same as the ones offered in the Facebook application. We evaluated our proposed solution for exporting user information towards subscribed educational systems using the REST protocol during task six of the two user evaluations:

Similarly to the procedure followed in the Facebook application, users were first given a list of all subscriptions by educational systems (Figure 8-15).

Figure 8-15 List of subscriptions by educational systems

The next step required users to inspect the subscription details and approve the export of all "Must" attributes and any "Want" attributes they wished to send towards AHA! (Figure 8-16).

Figure 8-16 "Must" and "Want" attributes – Users inspect and approve transaction details

Finally SUMI informed users whether the export transaction had been successful and presented an example of how their information could be used by the educational system. We used the same example as the one offered in the SUMI Facebook application: the adaptive lesson for HTML lists where personalised examples were generated based on the users' listed Facebook activities and interests. In order for the transaction to pass the test, SUMI had to assign URIs automatically to the exported

attributes and inform the educational system that by performing GET requests to the attributes' URIs, which were generated by the sqlREST tool, it would retrieve the values of the URIs, i.e., the user information.

Conclusions: Generating URIs to display the attributes' information along with adopting the REST protocol for allowing educational systems to perform GET requests on these URIs worked as expected. The sqlREST tool did not fail the challenge of generating URIs in real-time, and the action of performing a GET request on the generated URI was as easy as anticipated. Although we simulated the AHA! environment inside SUMI instead of using the real AHA system; more specifically we simulated an adaptive lesson on HTML unordered lists that can offered in AHA! in Task 6 of the second user evaluation, it does not invalidate our proposed communication protocol. We manually tested the generated URI for the attributes *activities* and *interests* with AHA v.3 and the value that the URI was exposing was displayed without any problems. The problem arose when we attempted to store the attribute's value inside AHA. The lack of any ability to instruct AHA, in the form of a programming command, to store the value did not allow us to test our proposed solution with the real system. However, the successor of AHA, a system called GALE, will resolve this omission and allow authors to include Java programming commands in conditional statements. This will also allow communication to take place between SUMI and GALE and, therefore, will validate fully our proposed communication protocol.

8.2.4 Using the SUMI Ontology to add value to raw user data

Once we had gathered the data from Facebook, either via the application or the website, we aimed to organise them inside SUMI using the proposed four-category architecture along with the SUMI ontology to map the various relationships between the retrieved information. As we stated in section 7.2.3, we have already mapped Facebook's data model inside the SUMI ontology. In addition, SUMI provides the necessary interfaces for any new provider to map its own data model onto any of its representatives' data models or to extend the ontology by defining its unique attributes. During the second evaluation, where we tested Facebook Connect, we also evaluated the SUMI ontology and whether our users would understand the value added by it.

Once the users initiated data imports and SUMI obtained their Facebook data, we offered them another option, which was showing them what information SUMI was holding about them inside its databases. This option was offered for the requirements of the scrutability evaluations and we used it to test also our SUMI ontology. Besides showing the data to the users, we enriched them with information from the SUMI ontology (Figure 8-17) while displaying all resulting information in the format of the four-category SUMI models' architecture, which we believe could help users understand the metadata better. More specifically, we added to the user data, which is presented in a brown font, the following information from the SUMI ontology:

- The name (and description) of the category where each attribute belongs, presented in a red font
- The name (and description) of each attribute as represented in the SUMI ontology, presented in a blue font
- The dictionary definition for each attribute, presented in a black font

Figure 8-17 Enriching displayed user data with information from the SUMI ontology

Furthermore, describing Facebook's data model in the SUMI ontology and then retrieving user data from the platform using the APIs provided a good test of whether our approach could meet the requirements of separating the data models, which are

stored in the SUMI ontology, and data, which are stored in the SUMI database. The most challenging task was to merge the two components when displaying the data and accompanying metadata to the users, as presented in Figure 8-17, since different programming commands on different platforms and tools must occur simultaneously in order to generate the proper output.

Conclusions: Our approach to separate the various data models from the data worked without any problems when we attempted the pre-described scenario with Facebook Connect and the SUMI website. The Facebook data model was retrieved from the SUMI ontology and displayed to the users with no significant time delay. The data were fetched from the SUMI database, since SUMI kept a copy when users initiated data imports in a previous step, either via the Facebook application or the SUMI website. Both actions were triggered simultaneously and the result, as shown in Figure 8-17, was displayed to all users during the second task of the second evaluation for scrutability and privacy. As the results have shown, and as is presented in the next chapter, 90% of users managed to complete the assigned task while 88% expressed their acceptance of having such an option available to them. This proves that our approach helped the users to understand and to engage better with this presentation, which was a mash up of information from the SUMI ontology and the SUMI database.

8.3 Summary

In this chapter, we have presented the design and the results of the four tests we conducted for interoperability, to test whether our proposed solution on interoperability verifies our hypotheses, which were presented in section 6.3.1. The design for evaluating interoperability was based on setting up a real-time scenario with the Facebook platform, via a Facebook application and via Facebook Connect, and the personalisation system AHA! The resulting outcome of all four executed tests revealed that our proposed approach for achieving interoperability is a solid one that can be used successfully with current social networking and e-commerce services that offer RESTful APIs, which were described in Table 7-1, and the selected educational personalisation system AHA!

Chapter 9 Evaluation Design and Results for Scrutability and Privacy

9.1 Introduction

In this chapter, we present the design and results of the first and second SUMI user evaluations for testing the proposed scrutability and privacy privileges against our hypotheses, which were presented in sections 6.3.2 and 6.3.3. Firstly, we describe the framework for the two user evaluations we conducted in order to gather feedback from potential SUMI users. For the first evaluation, we adopted a sequential approach where participants could complete only one task at a time, whereas for the second evaluation, we followed a more holistic approach and gave to the users all the available options for completing all tasks from the beginning. Secondly, we present the results of these two user evaluations, along with the lessons and conclusions obtained. The following sections clearly provide the statistics behind the results, drawn from the responses in the pre-questionnaires, post-questionnaires and the six tasks that presented the proposed scrutability and privacy privileges to the participants during both evaluations. In addition, we acknowledge which approach is most suitable for exposing the SUMI environment to potential users.

9.2 Evaluating Scrutability and Privacy in SUMI

In contrast with interoperability, the scrutability and privacy of user information are both user-centric requirements. Of course, there is a technical aspect to how SUMI deals with these issues behind the scenes, but for our research, we were mostly interested in finding out what potential SUMI users thought of our proposed scrutability and privacy policies and if they would approve of the proposed privileges we had decided to offer them (Kyriacou *et al.,* 2009b). For this reason, we used the SUMI prototype service, which was accessible via the SUMI website, *http://www.mysumi.org*. In addition, we designed and conducted two user-based evaluations, following a different approach in each one, in order to test the proposed infrastructure along with the proposed communication protocol by giving it to our users and asking them to have a hands-on experience. In this section, we describe the framework behind the two conducted evaluations highlighting the differences between the first and second attempts. The results and conclusions will be presented in the next chapter.

Table 9-1 presents and compares the main components of both evaluation frameworks. In this section, we provide a brief description, while a further detailed description of the structures of the evaluation, regarding what questions were asked and what available answers were provided to the participants, can be found in Appendix D:

	1st Evaluation	**2nd Evaluation**
Purpose	Both scrutability and privacy are crucial UM ingredients, which are mainly user-dependent. The purpose of the evaluations was to obtain the users' perspective on the proposed SUMI's scrutability and privacy user privileges	
Objectives	1. To evaluate whether the proposed scrutability and privacy user privileges are appropriate to be offered in SUMI. 2. To evaluate whether the proposed scrutability and privacy user privileges are accepted by SUMI users	
Hypotheses	We aimed to test our hypotheses for scrutability and privacy, as presented in sections 6.32 and 6.3.3	
Methodology	To expose the evaluation tasks to the participants, we incorporated them into the online SUMI service, which we	

<table>
<tr><td></td><td colspan="2">presented to the users; we then asked users to go through a three-step evaluation process:
First, we asked them to go through a pre-questionnaire (questions with multiple answers) where we identified how much users knew about the following:</td></tr>
<tr><td></td><td>a) Their scrutability and privacy options while interacting with various social networking and e-commerce providers on the WWW.
b) Which social networking and e-commerce providers they use most</td><td>a) The term ‘scrutability’ and if they could identify it when interacting with various social networking and e-commerce providers
b) Whether they would like to have scrutable options when interacting with such providers
c) Whether they were familiar with their privacy options when posting personal information on the various social networking and e-commerce services and whether they were taking advantage of these options
d) Whether they thought they had the option of customising their privacy options to meet their personal</td></tr>
</table>

<table>
<tr><td></td><td></td><td>preferences when interacting with social networking and e-commerce sites</td></tr>
<tr><td></td><td colspan="2">Second, we presented all proposed scrutability user privileges and all proposed privacy user privileges as hands-on tasks and:
a) We asked participants to complete all tasks
b) We introduced multiple-choice questions, during every task and after the completion of all tasks, in order to evaluate whether a proposed user privilege was appropriate to be offered and accepted by the participants as well.</td></tr>
<tr><td></td><td colspan="2">Finally, we asked the participants to go through a post-questionnaire (questions with multiple-choice answers and free text fields) where we examined:</td></tr>
<tr><td></td><td>a) How much users value scrutability and privacy after the completion of the evaluation
b) Users' suggestions for any new scrutability and/or privacy user privileges
c) What users thought about SUMI as a service and if they would use such a service
d) What users thought about the fact that SUMI was keeping a</td><td>a) Whether they would use SUMI and what they thought about the level of control they had during the evaluation
b) How much they valued scrutability after the evaluation and whether they had any suggestions regarding other scrutable privileges they would like SUMI to offer to its users
c) Similar to the previous point, how much they valued privacy after the completion of the</td></tr>
</table>

	copy of their information while interacting with it.	evaluation and whether they would like to suggest other privacy privileges to be offered in SUMI d) The participants' degree of familiarity with SUMI, before they took part in the second evaluation, by asking them if they had taken part in the first evaluation.
Claims	As described in sections 6.3.2 and 6.3.3, we claim that we can assess whether a proposed user privilege is appropriate to be offered in SUMI by evaluating users': a) Competence for completing each task (compare actual outcome after the completion of the tasks with users' answers to the evaluation questions [have they actually done it versus do they think they have done it] b) Understanding of the consequences of their decisions while interacting with each task. We also claim that we can assess if a proposed user privilege is accepted by users by evaluating users' acceptance of the proposed user privilege by asking them directly what they thought about each task and whether they would like SUMI to offer the privilege to its users	
Participants	The target audience for SUMI was undergraduate and postgraduate students, which was also the inclusion criterion.	

	We provided a SUMI prototype service and login information to each of the participants. For the purposes of the first evaluation, we had three pre-setup SUMI accounts that users could use to login to the service and go through the evaluation tasks.	We allowed each user to register and login with his/her own credentials in order to take full advantage of the personalised approach we had implemented in the second evaluation
	Participants were approached via an awarded Facebook game called MouseHunt (*http://apps.facebook.com/mousehunt/*), which attracts between 40k - 50k users every day. For the first evaluation, we kindly asked the users to take part without offering any incentive or reward. For the second one, we offered a small incentive by promising each participant the award of 10000 MouseHunt gold for completing the evaluation. In a period of 30 days, we managed to attract 107 users for the first evaluation and 111 for the second evaluation.	
Tasks exposing the offered *scrutability* privileges	1) Users were required to add at least one social networking and e-commerce profile to their SUMI collection. They were given a list of four providers: Facebook, MySpace, Amazon and eBay 2) Users were asked to import the content of each previously added profile using	1) Users were required to add at least one social networking and e-commerce profile to their SUMI collection. In addition, users had to import the content of these profiles as well. When users initiated data import from Facebook, real time data were imported and stored in SUMI. Users were re-directed to

	two methods: dynamic import, meaning real time data import from the selected provider (which was data that we provided and not actual real-time data), and static import meaning retrieval of user data that SUMI had stored in its database (SUMI was keeping a copy of the data when users initiated dynamic imports). 3) Users were instructed to export their information to a subscribed educational provider, a group formation system. Users had to inspect and approve the transaction before going through.	Facebook in order to login properly before importing their data into SUMI using Facebook Connect. 2) Users were asked to customise SUMI settings for their collection. They had the option of instructing SUMI what information was allowed to be imported and what was not. In addition, users had the option of deleting any value they no longer wanted SUMI to hold about them. 3) Users were instructed to export some of their information to a subscribed educational provider, which was AHA in this case. They had to inspect and approve the transaction before going through it. In addition, users had the option of going through an adaptive lesson about HTML unordered lists, in order to demonstrate how their

		information would be used by the educational system.
Tasks exposing the offered *privacy* privileges	1) The task exposed the users to the four-category architecture and asked them to set the privacy status of each one of the categories of at least one of their previously added profiles. We offered them the three privacy statuses that were introduced in section 7.3.2: public, private and hidden. 2) For the second task, we asked them to respond to a viewing request set by another user and allow that user to see the content of the requested profile 3) Finally the users were asked to place a viewing request to another user's private profile for accessing the content of that requested profile	1) Similarly to task 1 of the first evaluation, users had to set the privacy status for each of the four categories of at least one of the previously added profiles, using the three offered statuses: public, private and hidden. The difference from the first evaluation is that we offered a much improved presentation of each profile, clearly demonstrating the four-category architecture. 2) For the second task, we asked them to respond to a viewing request sent by another user and allow that user to see the private content of the requested profile. After they had completed this part, they were also instructed to take back the privilege they had assigned to the visiting user by taking advantage

		of a new option offered to them. In addition, we gave them the ability to form groups of other SUMI users and assign the same privacy status to the group instead of assigning the same status to every user in that group individually. 3) Similarly to the third task in the first evaluation, we asked users to visit another user's SUMI collection and place a viewing request to a private category of that user's profiles.
Approaches (more on this below)	The approach we followed for the first evaluation was sequential. We guided the users throughout the evaluation by directing them from the first task to the second one and then to the third one and so on. We did not give them the freedom to navigate through the SUMI website. We attempted to keep them focused on the task in hand without worrying about the	For the second evaluation, we followed a completely different approach to the first one. Instead of guiding the user step-by-step, we made all the options available to the users from the beginning and let them find their way around the SUMI website on their own. On the left hand panel we were informing them of their progress throughout the evaluation and on the top we offered links to the start of each

	evaluation flow.	of the six tasks.

Table 9-1 Comparison of the two evaluation frameworks

The feedback from the first evaluation was taken into account when designing the second evaluation; thus, the offered user privileges of the second evaluation can be considered more advanced and complex than the ones of the first evaluation. In addition, the reason we followed a different approach in each evaluation was due to a variety of contributions in the literature that evaluated how better to present user models to users:

- A recent study experimented on which representation of user models is preferred by users. Representations, shown in Figure 8-18, were divided into three categories: ordered (list, medals and podium), absolute (tag cloud, starts and sliders) and relative (pie chart, bricks and coins). Results revealed that the preferred visualisations are those that are commonly used in social websites, such as the list for the ordered representation, the stars and the sliders for the absolute representation, and the pie chart for the relative representation. In addition, users did not express any preference for any of the three categories as a whole (Vernero *et al.,* 2009).

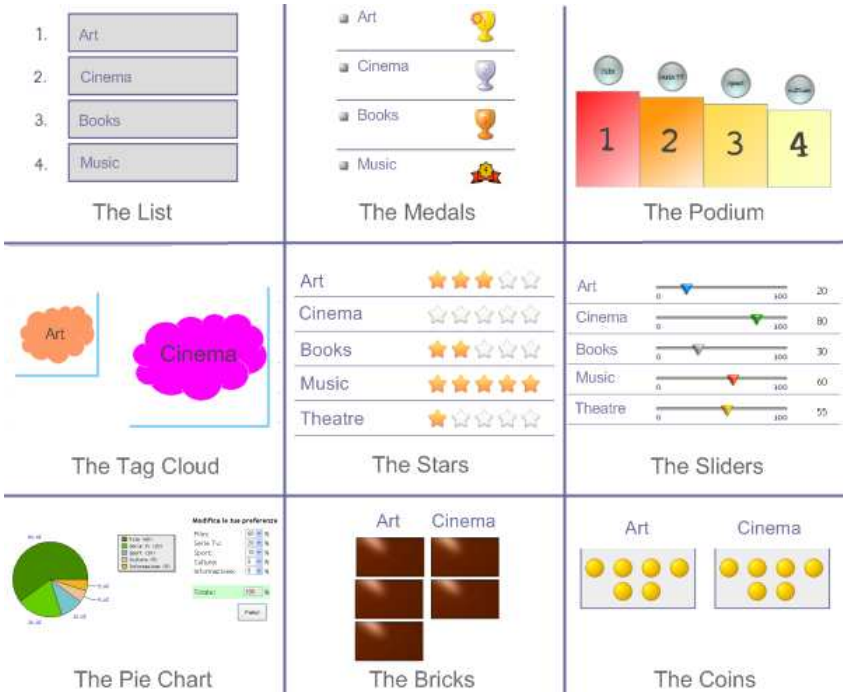

Figure 9-1 The three categories of visualisation of user models – ordered, absolute and relative (Vernero *et al.,* 2009)

- A second study investigated the various phases that products and services go through in terms of user acceptance and engagement. Four categories were identified: enthusiastic phase, where users interact with the product/service as a hobby; professional phase, where users interact with the product/service in the workplace; consumer phase, where the product/service is part of the user's life; and baroque phase, where the product/service is considered to be "too much" for the users to handle and therefore it is rejected and declined. As claimed, the 'sweet spot' of a product or a service should be somewhere between the product and baroque phases in order to achieve the highest level of user acceptance and engagement (Borchers, 2008)
- A third study was concerned with the various challenges that need to be addressed when presenting user models to their owners. To summarise, the most important aspects for assisting the users to comprehend the representation of their models are, first, to ensure that the users understand the modelling ontology behind their representation of their models, i.e., the semantics of the metadata; second, to highlight the most relevant and interesting parts of their models; and third, to establish learning environments that can enable users to focus effectively on the usefulness of their models (Kay, 1997)

In order to identify which method was most suitable for exposing our users to the SUMI environment and to their various social networking and e-commerce models, we decided to adopt and test two approaches.

- *Sequential approach*: This was the approach we followed during the first evaluation. Users were exposed to SUMI and to all six tasks in a sequential way, meaning we obliged the participants to complete one task at a time. After they had completed pre-questionnaire, we presented the first task and asked them to finish it before proceeding to the second one. Similarly, each task was presented only after the completion of the previous one. The list visualisation, incorporated in the proposed four-category models' architecture, was used to represent the various user

models to the participants. In addition, a detailed explanation of every step during the evaluation was displayed at the top of each webpage. At the end of all six tasks, users had to complete the post-questionnaire in order to exit from the evaluation.

- *Holistic approach:* This was the approach we adopted for the second evaluation. Users were exposed to the pre-questionnaire, all six tasks and the post-questionnaire from the beginning of the evaluation, and were asked to find their own way of completing them. At the top of every webpage, we offered links to the start of every task, and on the left hand side we were continually informing participants of their evaluation progress. The list visualisation, incorporated in the proposed four-category models' architecture, was used to represent the various user models to the participants. In addition, metadata information was retrieved from the SUMI ontology and displayed during the second evaluation in order to help the users comprehend the representation of their models. Furthermore, an adaptive lesson was offered to participants to provide an example of how their models could be used by educational personalisation systems. The detailed explanation that was offered in the first evaluation was removed from every webpage and was replaced with buttons that could be used by the participants at any time to receive additional information and explanations regarding each evaluation step.

We tested both the sequential and the holistic approaches under similar evaluation conditions and let the results reveal which one the participants would prefer to be offered in the SUMI context.

9.3 Pre-Questionnaires

The first part in both evaluations was the pre-questionnaires. This part was presented in the form of questions with multiple-choice answers from which the participants had to select one. In Tables 9-2, 9-3, 9-4 and 9-5, we present the percentage of each available response to every question of the pre-questionnaires in both evaluations:

Question 1 (Scrutability) of both evaluations: Are you familiar with the term 'scrutability'? Do you think the various providers (e.g., Facebook, Amazon, etc.) offer you some scrutability user privileges?

Available Response	1st Evaluation	2nd Evaluation
I am familiar with the term and I recognize some scrutability user privileges when I interact with various providers.	10%	3%
I am familiar with the term, but I do NOT recognize any scrutability user privileges when I interact with various providers.	1%	7%
I am NOT familiar with the term, but I do recognize some scrutability user privileges when I interact with various providers.	38%	36%
I am NOT familiar with the term NOR do I recognize any scrutability user privileges when I interact with various providers.	51%	54%

Table 9-2 Pre-Questionnaire – Question 1

Question 2 (Scrutability) of both evaluations: Would you like to have the option of being allowed to inspect and alter, in a variety of ways, the way you are been modelled by various providers (such as MySpace and eBay)?

Available Response	1st Evaluation	2nd Evaluation
Yes, that would be great	80%	77%
Sounds good	17%	16%
I don't really mind	3%	3%
No	0%	0%
I don't understand the question	0%	4%

Table 9-3 Pre-Questionnaire – Question 2

Question 3 (Privacy) of both evaluations: Are you familiar with your privacy options when you interact with various providers (e.g., MySpace, eBay, etc.)? Do you take advantage of the offered privacy options?

Available Response	1st Evaluation	2nd Evaluation
I am familiar with my privacy options and I do take advantage of them.	32%	61%
I am familiar with my privacy options, but I do NOT take advantage of them.	32%	20%
I am NOT familiar with my privacy options, but I do take advantage of them.	2%	0%
I am NOT familiar with my privacy options NOR do I take advantage of them.	34%	19%

Table 9-4 Pre-Questionnaire – Question 3

Question 4 (Privacy) of Evaluation 1: Would you mind if the various providers with which you interact (such as MySpace and eBay) did not allow you to set the privacy status of your information that you had personally entered in a previous stage?

Question 4 (Privacy) of Evaluation 2: Do you think you have the freedom to customise your privacy settings when interacting with various providers (e.g., Facebook, Amazon, etc.) in order to control how the offered privacy privileges are applied to your profiles?

Available Response	1st Evaluation	2nd Evaluation
E1: It would bother me. E2: Yes, I have the option of customising my privacy settings when I interact with some providers.	49%	44%
E1: Yes, I would mind a lot. E2: I recognise some options for customising my privacy privileges, but I would like to be offered more.	42%	49%
E1: I don't really mind. E2: I don't recognise any options for customising my privacy settings.	9%	0%

E1: No. E2: No such options are offered when I interact with various providers.	0%	 7%
E1: I don't understand the question. E2: I don't understand the question.	0%	 0%

Table 9-5 Pre-Questionnaire – Question 4

Conclusions: The pre-questionnaires revealed a number of useful facts regarding the users' familiarity with the evaluated terms 'scrutability' and 'user privacy'.

Regarding scrutability, the results reveal that the majority of students do not know what the term 'scrutability' means, although some can easily identify some scrutability privileges once they have had them explained to them, which was done by presenting a realistic example from the social networking domain; although we acknowledge the possibility that participants may chose their responses in order to express that they understood the offered example and not the term 'scrutability' per se. In addition, considering that the term 'scrutability' is something that was introduced and used in the LUM community, it is interesting to observe that 11% of users stated their familiarity with the term, which may hint that users were expressing their familiarity with the concept of scrutability and not the term per se.

Furthermore, users found the idea of having scrutability privileges available when interacting with various providers a very good idea, which shows the acknowledgment of the importance of scrutability to users, once it has been explained to them.

'User privacy' is a term more familiar to users than is scrutability. It is something they understand and recognize when interacting with several providers, although a significant percentage of participants chose not to take advantage of it. However, at the suggestion of no privacy privileges being available, 91% of users in the first evaluation expressed their concerns. In addition, users said that they identify options for customising their privacy privileges when interacting with various providers and almost half of them clearly stated that they wanted to be offered more. An interesting point regarding this question is that, in fact, users are not given options for customising their privacy privileges. Social networking and e-commerce providers offer a "fixed package

of privacy settings" that users are obliged to follow. Facebook can be considered an exception with its recent announcement that users have the choice of "being as open or as limited in the sharing of their information as they want" (DailyTech, 2009). Several privacy layers are offered to Facebook users, who have the privilege to decide how to apply the various privacy settings in any way they prefer. We wanted to clarify whether users understood the notion of customised privacy privileges, and we believe their responses revealed that they did not fully comprehend the difference between fixed and customisable privacy settings.

9.4 Scrutability Privileges

During both evaluations, three tasks were designed and implemented to reveal the proposed scrutability privileges to the participants. We will briefly describe the tasks in the following sections while providing the percentage of our successful participants' responses. Further on, we will compare the results of the two evaluations and assign "not satisfactory" or "successful" to each task while explaining the reasons for our decisions.

9.4.1 1st Evaluation

Task 1 – Adding models to the SUMI collection: Users were asked to add at least one social networking model and at least one e-commerce provider. They were provided with a list of available providers, comprising Facebook, MySpace, Amazon and eBay. Once the users had added their selected models, they were exposed to the four-category SUMI model's architecture and were also asked a question to evaluate the degree to which users understood the structure of the proposed models' architecture.

Task 2 – Importing the content of previously added models, dynamically and statically: In this task, users were asked to import the content of their previously added models in two ways.

- *Dynamically* – a dynamic import uses the various providers' APIs to import in real time the users' information from the respective providers. For example, a dynamic import of a user requesting his "Favourite Movies" from Facebook, initiates a query that uses Facebook's API in

order to retrieve from Facebook the actual, real-time value for that user's "Favourite Movies" attribute. We would like to note that although such available APIs did exist and were available for use at the period of the first evaluation, they were still immature for use on such an occasion. Therefore, for the purposes of the first evaluation, we used the method of a placebo to let users think that actual requests to the various providers were enabled, when actually we were retrieving the data from our own databases and presenting them as real-time data.

- *Statically* – once a user initiated a dynamic import, SUMI kept a copy of that information inside its own databases. This was implemented in order to offer to users the option of the static import of their data. A static import is when a user initiates a query for data retrieval from the SUMI databases, instead of the providers' databases using their available APIs. . Busy network traffics that can be observed when the providers' APIs are overloaded by millions of simultaneous requests is one of the reasons that could lead users to take advantage of static imports.

Users were first asked to import the content of at least one of their previously added models, both dynamically and statically. In addition, users also had to import the content of another model statically only and notice the difference from the first occasion. Meanwhile, two during-task questions tested users' engagement with the way SUMI was presenting their models' content and tested users' understanding of the difference between dynamic and static import.

Task 6 – Exporting the content of previously added models to (sample) subscribed educational providers: During task 6, users were asked to respond to an export request and export the content of one of their models towards a – hypothetical - subscribed provider, a group formation system. Users were exposed step-by-step to the export process and were allowed to cancel the export process if they felt they wanted to do so. Once again, a during-task question evaluated how users felt about inspecting and approving the process before going through it.

After the completion of all six tasks (including the three tasks for privacy), a series of three-questions-per-task was presented to the users in order to evaluate their degree of the following:

- *competence* on completing each task,

- *understanding of the consequences* of their decisions while interacting with each task,
- *acceptance* of each proposed privilege

Table 9-6 summarizes the results for the three proposed scrutability privileges by presenting the percentage of successful responses to the respective questions of the evaluated categories, namely, competence, consequence and acceptance, along with the percentage of the successful responses to the during-task questions:

Proposed Scrutability Privilege	**Competence(%) Answers/Actual outcome**	**Consequence(%)**	**Acceptance(%)**	**During(%)**
Adding models to SUMI collection	100/100	77	89	73
Importing content of previously added models, both dynamically and statically	79/60	67	100	78, 70 *two during-task questions
Exporting content of previously added models to subscribed providers	91/100	60	94	72

Table 9-6 Results (% of successful responses) for the proposed Scrutability Privileges in the first evaluation

The statistical evidence for the successful responses to the competence questions, the understanding of consequence questions and the during-task questions for the proposed scrutability privileges for the first evaluation is as follows:

- Average of successful responses = 76,70
- Standard Deviation = 11,59
- Confidence Interval for 95% confidence level = 7,18

Similarly, the statistical evidence for the positive responses to the acceptance questions for the proposed scrutability privileges for the first evaluation is as follows:

- Average of positive responses = 94,33
- Standard Deviation = 3,79
- Confidence Interval for 95% confidence level = 4,62

9.4.2 2nd Evaluation

Task 1 - Adding profiles to SUMI collection and importing content of added profiles: We asked users to add their social networking and e-commerce profiles to their SUMI collection. We offered them a list that included Facebook, MySpace, Amazon and eBay and we asked them to select at least one provider from the list. It was up to the participants which profiles and how many they would add to their SUMI collection. Furthermore, once they had added their profiles, they were required to import their profiles' content as well. During this task, the participants were exposed to the four-category SUMI models' architecture, and a during-task question evaluated their understanding of the presentation of their profiles' content as it was incorporated into the proposed models' architecture.

Task 2 - Customising settings for SUMI collection: Using the appropriate options, users could customise the SUMI settings for their previously added profiles according to their preferences. Can SUMI keep a copy of the imported information? If yes, for which attributes of which of their profiles? In addition, which attributes would they like SUMI never to import? These were some of the questions to which the participants were asked to respond. They were required to customise their SUMI collection by taking advantage of the available interfaces. If they did not have any preference, they were asked to follow a hypothetical scenario for which they had to delete the value of their "Favourite Movie" if it existed in SUMI and, in addition,

instruct SUMI never to import that attribute again. As with every task, a during-task question tested whether users comprehended which SUMI settings they had control over and were allowed to customise.

Task 6 – Exporting the content of previously added profiles to an educational personalisation system (AHA!): SUMI enables educational services to subscribe to the various providers of user profiles with which users have an account, and request them to export their information towards them. Users were asked to inspect and approve the details of an export transaction towards the educational system AHA! before allowing the export transaction to be completed. Once they had exported their data, users had the chance of going through an example of an adaptive lesson that demonstrated how their data could be used by AHA!. The example we offered was an adaptive lesson on HTML lists, which provided a personalised example based on the value of the users' Facebook *interests* and *activities* attributes. In this task, the during-task question evaluated whether users felt that they had inspected and approved the transaction before going through the example.

Similarly to the first evaluation, after the completion of all six tasks (including the three tasks for privacy), a series of three-questions-per-task was presented to the users in order to evaluate their degree of the following:

- *competence* on completing each task,
- *understanding of the consequences* of their decisions while interacting with each task,
- *acceptance* of each proposed privilege

Table 9-7 summarizes the results for the three proposed scrutability privileges by presenting the percentage of successful responses to the respective questions of the evaluated categories, namely, competence, consequence and acceptance, along with the percentage of the successful responses to the during-task questions:

Proposed Scrutability Privilege	**Competence(%) Answers/Actual outcome**	**Consequence(%)**	**Acceptance(%)**	**During(%)**
Adding models to SUMI collection	98/100	95	93	93

and importing content				
Customising SUMI settings of previously added profiles for their collection	89/91&90 (users were required to delete the value of "Favourite Movies" and instruct SUMI never to import this attribute again)	89	88	90
Exporting content of previously added models to subscribed providers	98/100	89	93	89

Table 9-7 Results (% of successful responses) for the proposed Scrutability Privileges in the second evaluation

The statistical evidence for the successful responses to the competence questions, and the understanding of the consequence questions and of the during-task questions for the proposed scrutability privileges for the second evaluation is as follows:

- Average of successful responses = 92,22
- Standard Deviation = 3,90
- Confidence Interval for 95% confidence level = 2,55

Similarly, the statistical evidence for the positive responses to the acceptance questions for the proposed scrutability privileges for the second evaluation is as follows:

- Average of positive responses = 91,33
- Standard Deviation = 2,89
- Confidence Interval for 95% confidence level = 3,27

9.4.3 Comparison of Results and Conclusions

Competence: Figures 9-2 and 9-3 present a comparison of the two evaluations regarding the competence results. The red colour (and green colour) signifies the actual successful outcome while the blue colour shows the participants' positive responses to the evaluation question "*Do you think you have managed to complete the task…?*"

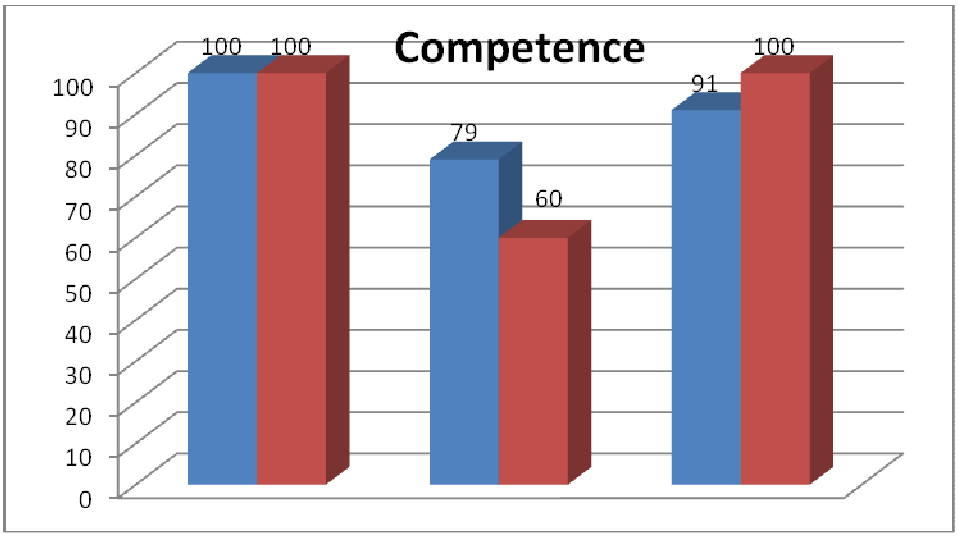

Figure 9-2 Competence results (%) for 1st evaluation

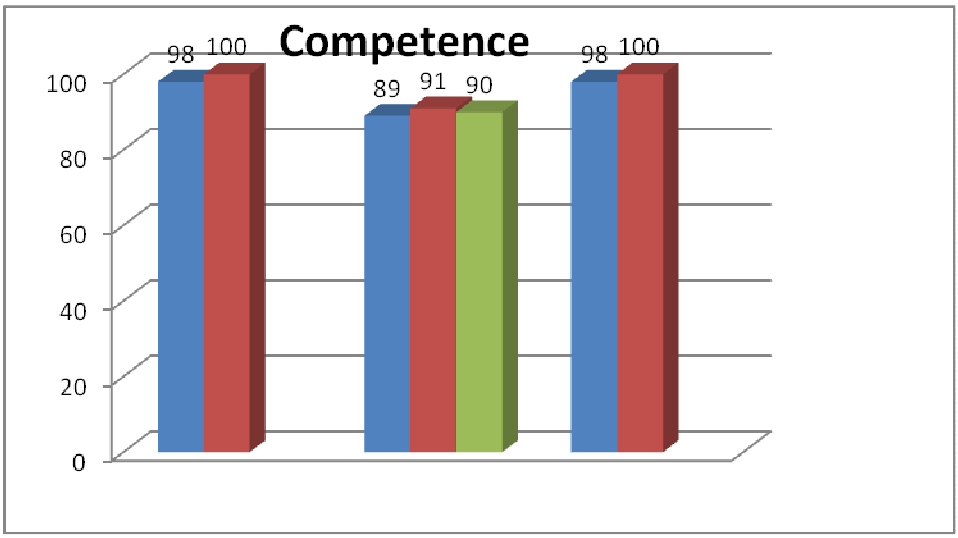

Figure 9-3 Competence results (%) for 2nd evaluation

Understanding of Consequences: Figures 9-4 and 9-5 present a comparison of the two evaluations regarding the successful responses to the questions that evaluated the participants' understanding of consequences while attempting to solve the three scrutability tasks.

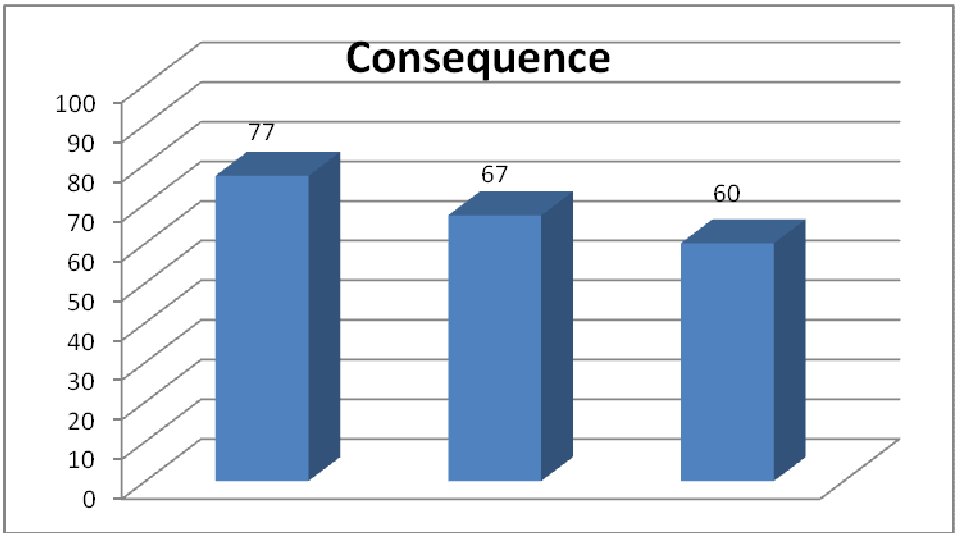

Figure 9-4 Understanding of consequences results (%) for 1st evaluation

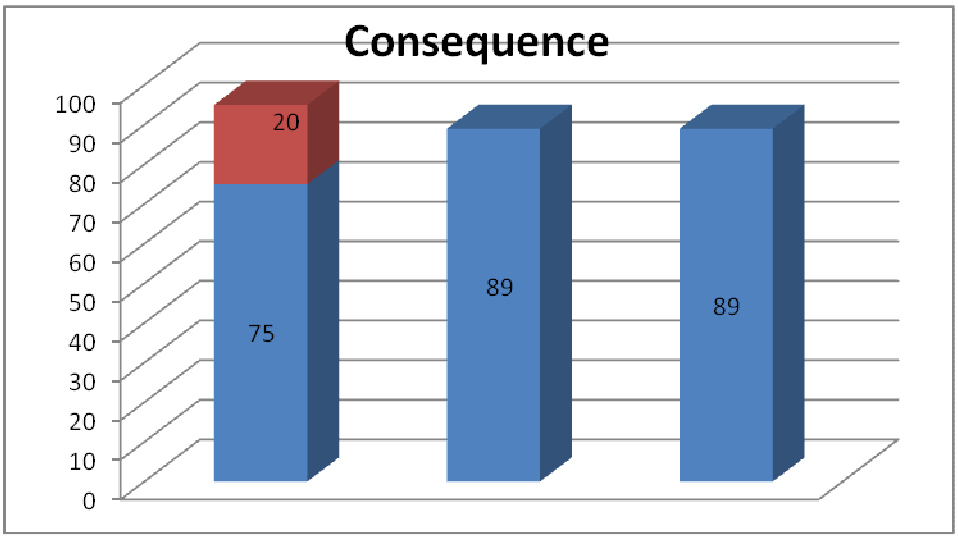

Figure 9-5 Understanding of consequences results (%) for 2nd evaluation

Acceptance: Figures 9-6 and 9-7 present a comparison of the two evaluations regarding the positive responses to the questions that evaluated the participants' acceptance of the proposed scrutability privileges.

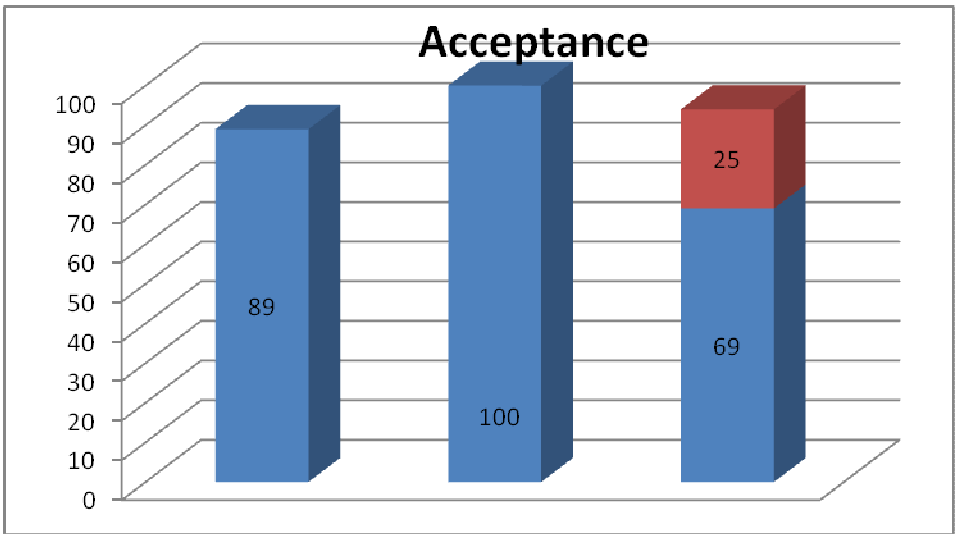

Figure 9-6 Acceptance results (%) for 1st evaluation

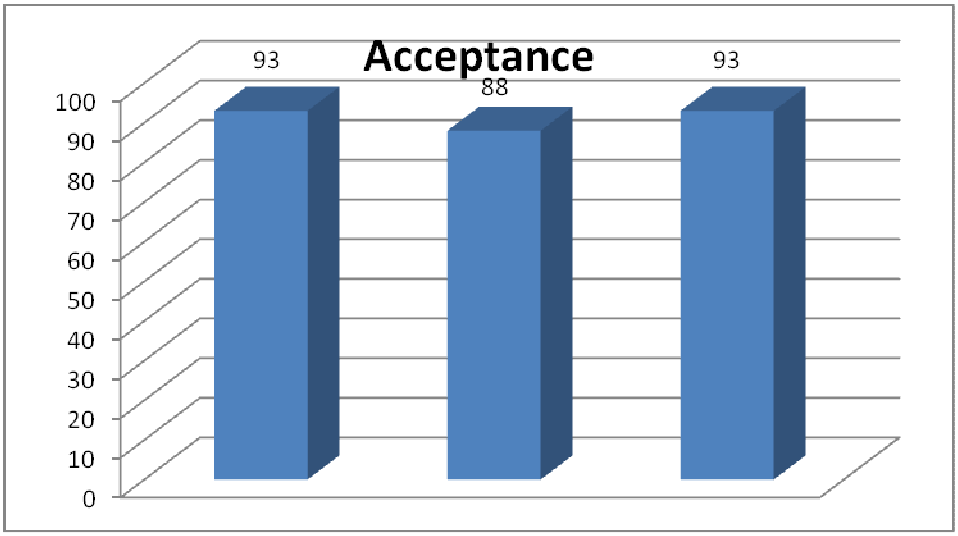

Figure 9-7 Acceptance results (%) for 2nd evaluation

During-Tasks Questions: Figures 9-8 and 9-9 present a comparison of the two evaluations regarding the successful participants' responses to the during-tasks questions.

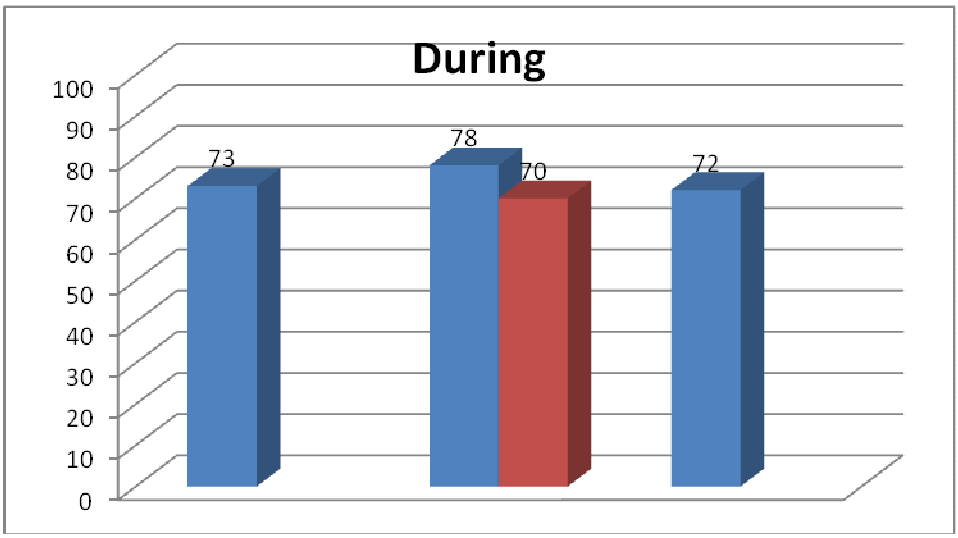

Figure 9-8 During-task questions' results (%) for 1st evaluation

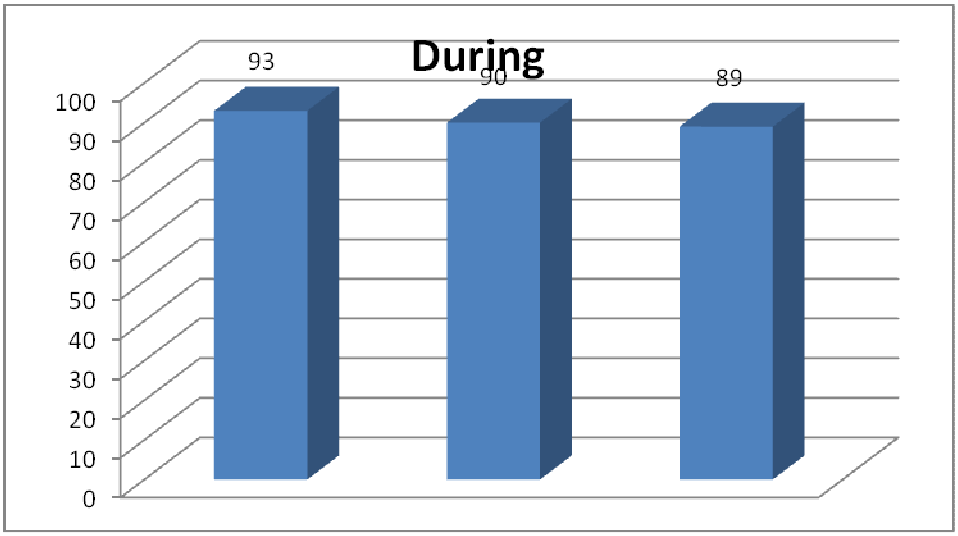

Figure 9-9 During-task questions' results (%) for 2nd evaluation

Evaluation 1 Summary: Table 9-8 summarizes our conclusions regarding which privilege was successful in which category – competence, consequence and acceptance - for the first evaluation. The X symbol means "not satisfactory", whereas V means "successful".

Proposed Scrutability Privilege	**Competence**	**Consequence**	**Acceptance**
Adding models to SUMI collection	V	X	V
	All participants managed to complete the task. In addition, all users answered positively the competence question	77% actually understood the consequences of their decisions while completing the task, whereas 73% answered the during-task question correctly	89% of users expressed their acceptance of this scrutability privilege

Importing content of previously added models, both dynamically and statically	X	X	V
	79% of users thought they had managed to complete the task whereas, in fact, only 60% actually had managed to complete it	67% actually understood the consequences of their decisions while completing the task, whereas 78% and 70% respectively answered the two during-task questions correctly	All users expressed their acceptance of this scrutability privilege
Exporting content of previously added models to subscribed providers	V	X	V
	91% of users thought they had managed to complete the task whereas in fact all users actually had managed to complete it	60% actually understood the consequences of their decisions while completing the task, whereas 72% answered the during-task question correctly	94% of users expressed their acceptance of this scrutability privilege

Table 9-8 Summary of conclusions for proposed scrutability privileges (1st evaluation)

Evaluation 2 Summary: Table 9-9 summarizes our conclusions regarding which privilege was successful in which category – competence, consequence and acceptance - for the second evaluation. The X symbol means "not satisfactory", whereas V means "successful".

Proposed Scrutability Privilege	**Competence**	**Consequence**	**Acceptance**
Adding models to SUMI collection and importing content	V	V	V
	98% of users thought they had managed to complete the task, whereas, in fact, all users actually had managed to complete it	95% actually understood the consequences of their decisions while completing the task, whereas 93% answered correctly to the during-task question	93% of users expressed their acceptance of this scrutability privilege
Customising SUMI settings of previously added profiles for their collection	V	V	V

	89% of users thought they had managed to complete the whole task, whereas, in fact, only 91% and 90% respectively actually had managed to complete the two parts of this task	89% actually understood the consequences of their decisions while completing the task, whereas 90% answered the during-task question correctly	88% of users expressed their acceptance of this scrutability privilege
Exporting content of previously added models to subscribed providers	V	V	V
	98% of users thought they had managed to complete the task whereas, in fact, all users actually had managed to complete it	89% actually understood the consequences of their decisions while completing the task, whereas 89% answered the during-task question correctly	93% of users expressed their acceptance for this scrutability privilege

Table 9-9 Summary of conclusions for proposed scrutability privileges (2nd evaluation)

Conclusions: Although users expressed their approval of all three privileges in both evaluations, the sequential approach followed in the first evaluation did not help the users to comprehend the consequences of their interaction with the SUMI environment. Furthermore, the sequential approach did not help to explain to the users the notions of adding profiles and importing the content of those profiles. Both these obstacles were overcome in the second evaluation, which adopted the holistic approach.

Users were aware of the consequences of their decisions while completing the tasks and managed to fulfil the requirements of the second evaluation successfully. We also noticed that adding metadata information from the SUMI ontology and presenting an example of an adaptive lesson helped the participants to engage more deeply in contrast to the sequential approach of the first evaluation, where the detailed explanation of every evaluation step did not produce the expected results. By analysing the obtained results of the first evaluation, we conclude that users did not read the heavy text fragments at the beginning of every webpage since that would have helped them to complete the first evaluation successfully. In addition, users rarely clicked on the "Help" button that was offered on every webpage during the second evaluation and that provided additional explanation and information for every evaluation step.

9.5 Privacy Privileges

Similarly to the approach for exposing the offered scrutability privileges, during both evaluations, three tasks were designed and implemented to expose the proposed privacy privileges to the participants. We will briefly describe the tasks in the following sections while providing the percentage of our successful participants' responses. Further on, we will compare the results of the two evaluations and assign "not satisfactory" or "successful" to each task while explaining the reasons for our decisions.

9.5.1 1st Evaluation

Task 3 - Setting privacy status for all four categories of previously added models: During this task, users had to set the privacy status of all four categories of at least one of their models using the proposed three privacy settings:

- *public* – others can see that the model exists and anyone can view its content,
- *private* – others can see that the model exists, but they have to place a request to the model's owner to view the model's content
- *hidden* – others cannot see that the model exists; therefore, the model's content is accessed only by the model's owner.

More specifically, they were asked to set the 1st category as public, the 2nd category as private and the 3rd and 4th as hidden. In addition, a during-task question verified users' understanding of the three privacy settings offered.

Task 4 - Responding to a viewing request from another SUMI user: Once users had attempted this task, they were informed that another user was requesting permission to view one private category of one of their models. Users had to respond to this request and allow the requester to view the content of their private category. No during-task question existed in this task.

Task 5 - Visiting another user's SUMI collection and placing a viewing request: Similarly to the previous task, users were asked to visit another user's SUMI collection of models and place a request on one private category of that user's models. A during-task question tested users' understanding of SUMI's presentation of public categories' content when participants visited another user's model.

After the completion of all six tasks (including the three tasks for scrutability), a series of three-questions-per-task was presented to the users in order to evaluate the their degree of the following:

- *competence* on completing each task,
- *understanding of the consequences* of their decisions while interacting with each task,
- *acceptance* of each proposed privilege

Table 9-10 summarizes the results for the three proposed user privacy privileges, by presenting the percentage of successful responses to the respective questions of the evaluated categories, namely, competence, consequence and acceptance, for the first evaluation, along with the percentage of the successful responses to the during-task questions.

Proposed Privacy Privilege	**Competence(%) Answers/Actual outcome**	**Consequence(%)**	**Acceptance(%)**	**During(%)**
Setting privacy status	76/68	100	100	71

Responding to viewing request	88/74	100	99	N/A
Visiting another user's model and placing a viewing request	84/100	79	83	100

Table 9-10 Results (% of successful responses) for the proposed User Privacy privileges (1st evaluation)

The statistical evidence for the successful responses to the competence questions, and the understanding of the consequence questions and the during-task questions for the proposed privacy privileges for the first evaluation is as follows:

- Average of successful responses = 87,25
- Standard Deviation = 11,70
- Confidence Interval for 95% confidence level = 8,11

Similarly, the statistical evidence for the positive responses to the acceptance questions for the proposed privacy privileges for the first evaluation is as follows:

- Average of positive responses = 94,00
- Standard Deviation = 3,46
- Confidence Interval for 95% confidence level = 4,26

9.5.2 2nd Evaluation

Task 3 – Setting privacy preferences for previously added profiles: This task was very similar to the corresponding task 3 we provided in the first evaluation. Users had to set the privacy status of all four categories of at least one of their models using the proposed three privacy settings:

- *public* – others can see that the model exists and anyone can view its content,
- *private* – others can see that the model exists, but they have to place a request to the model's owner for viewing the model's content
- *hidden* – others cannot see that the model exists; therefore, the model's content is accessed only by the model's owner.

More specifically, they were asked to set the 1st category as public, the 2nd category as private and the 3rd and 4th as hidden. The extra task we added for the second evaluation was asking the users to proceed further and assign a customised privacy privilege to Maria, another SUMI user. We also gave them a much improved interface, which also allowed them to create group of users and assign customised privacy privileges collectively rather than individually. A during-task question verified users' understanding of the offered three privacy settings and the relation between universal and customised privacy settings.

Task 4 – Responding to viewing requests sent by other SUMI users: Again, similarly to evaluation 1, first, users had to respond to a viewing request by another user and allow the requester to view the content of their private category. Second, they were asked to take back that privilege using a new feature provided in the second evaluation. With the during-task question, we evaluated whether users understood the offered privacy privilege that allowed them to assign customised privacy privileges to other users and then take them back if they wished to.

Task 5 – Visiting another user's SUMI collection and placing a viewing request: In task 5, participants were asked to visit another user's SUMI collection and place a request on one private category of that user's profiles. A during-task question tested users' understanding of the proposed four-category SUMI models' architecture when visiting other users' SUMI collections.

As with the first evaluation, after the completion of all six tasks (including the three tasks for scrutability), a series of three-questions-per-task was presented to the users in order to evaluate their degree of the following:

- *competence* on completing each task,
- *understanding of the consequences* of their decisions while interacting with each task,
- *acceptance* of each proposed privilege

Table 9-11 summarizes the results for the three proposed privacy privileges by presenting the percentage of successful responses to the respective questions of the evaluated categories, namely, competence, consequence and acceptance, for the second evaluation, along with the percentage of the successful responses to the during-task questions.

Proposed Privacy Privilege	**Competence(%) Answers/Actual outcome**	**Consequence(%)**	**Acceptance(%)**	**During(%)**
Setting privacy preferences for previously added profiles	97/90	93	99	88
Responding to viewing request by another SUMI user and managing assigned privilege	90/100 & 87 (users were asked to respond to the viewing request and manage that request afterwards)	92	93	89
Visiting another user's collection and placing a viewing request	96/99	90	94	100

Table 9-11 Results (% of successful responses) for the proposed User Privacy privileges (2nd evaluation)

The statistical evidence for the successful responses to the competence questions, and the understanding of the consequence questions and of the during-task questions for the proposed privacy privileges for the second evaluation is as follows:

- Average of successful responses = 92,78
- Standard Deviation = 4,09
- Confidence Interval for 95% confidence level = 2,67

Similarly, the statistical evidence for the positive responses to the acceptance questions for the proposed privacy privileges for the second evaluation is as follows:

- Average of positive responses = 95,33
- Standard Deviation = 3,21
- Confidence Interval for 95% confidence level = 3,63

9.5.3 Comparison of Results and Conclusions

Competence: Figures 9-10 and 9-11 present a comparison of the two evaluations regarding their competence results. The red colour (and green colour) signifies the actual successful outcome while the blue colour shows the participants' positive responses to the evaluation question "*Do you think you have managed to complete the task…?*

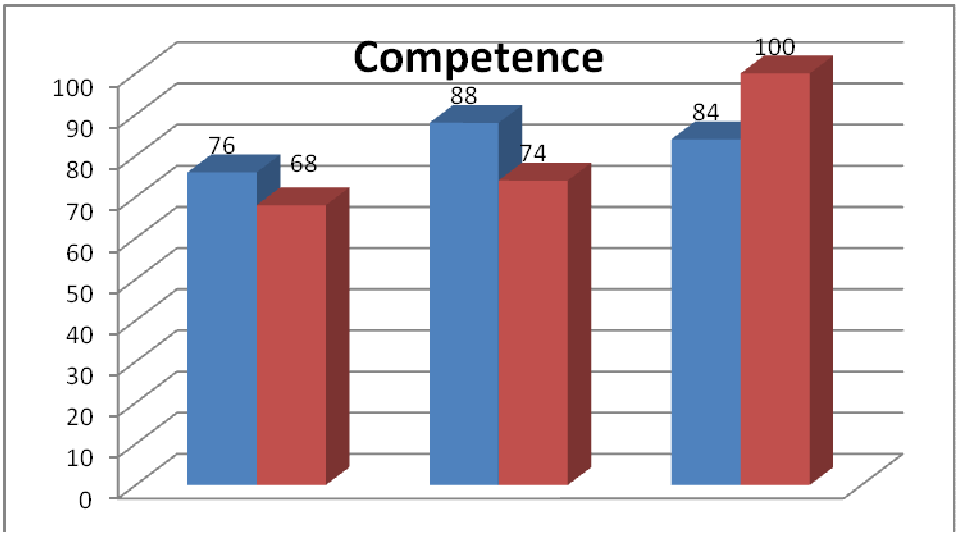

Figure 9-10 Competence results (%) for evaluation 1

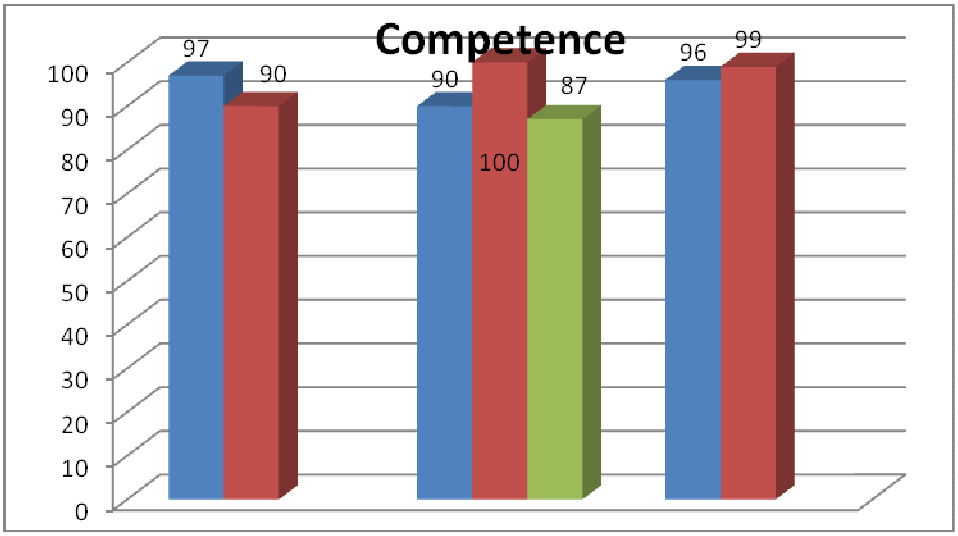

Figure 9-11 Competence results (%) for evaluation 2

Understanding of Consequences: Figures 9-12 and 9-13 present a comparison of the two evaluations regarding the successful responses to the questions that evaluated the participants' understanding of the consequences while attempting to solve the three privacy tasks.

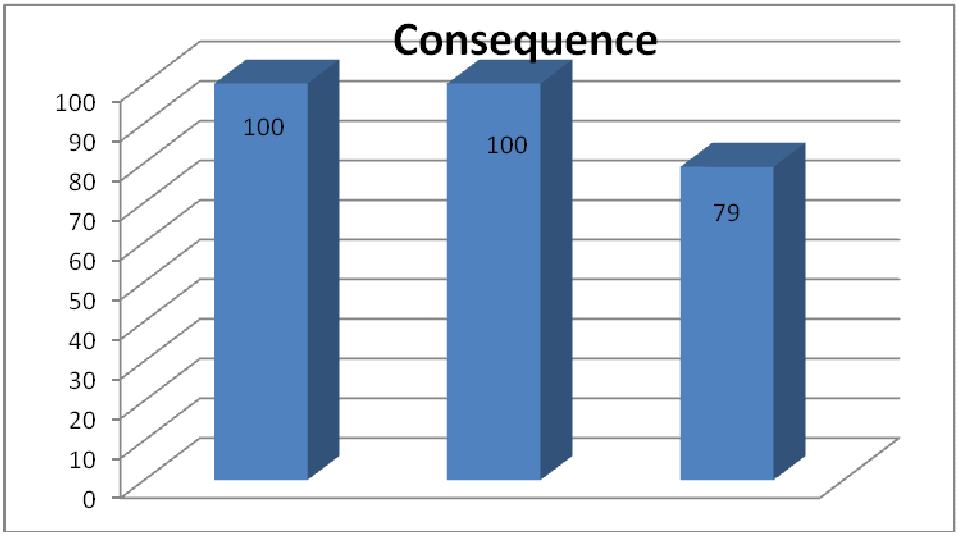

Figure 9-12 Understanding of consequence results (%) for evaluation 1

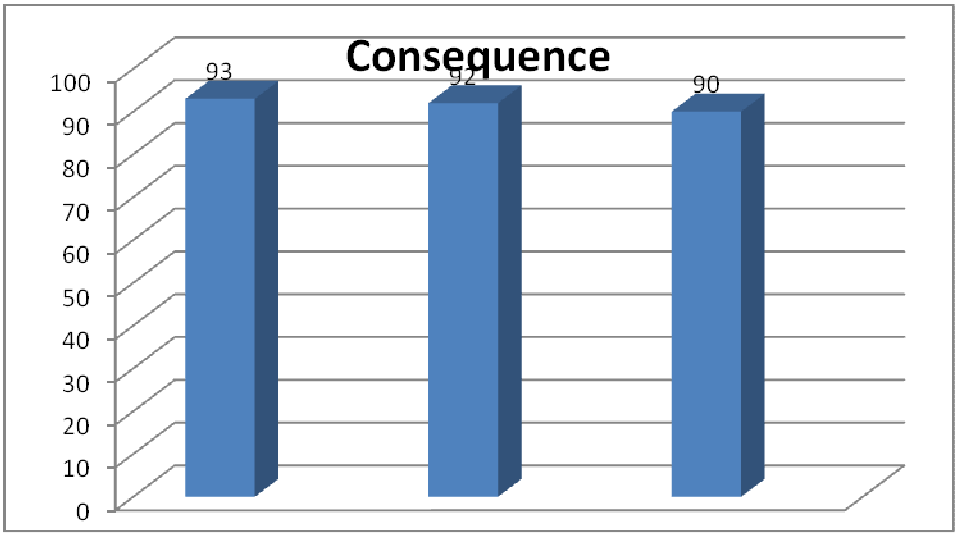

Figure 9-13 Understanding of consequence results (%) for evaluation 2

Acceptance: Figures 9-14 and 9-15 present a comparison of the two evaluations regarding the positive responses to the questions that evaluated the participants' acceptance of the proposed privacy privileges.

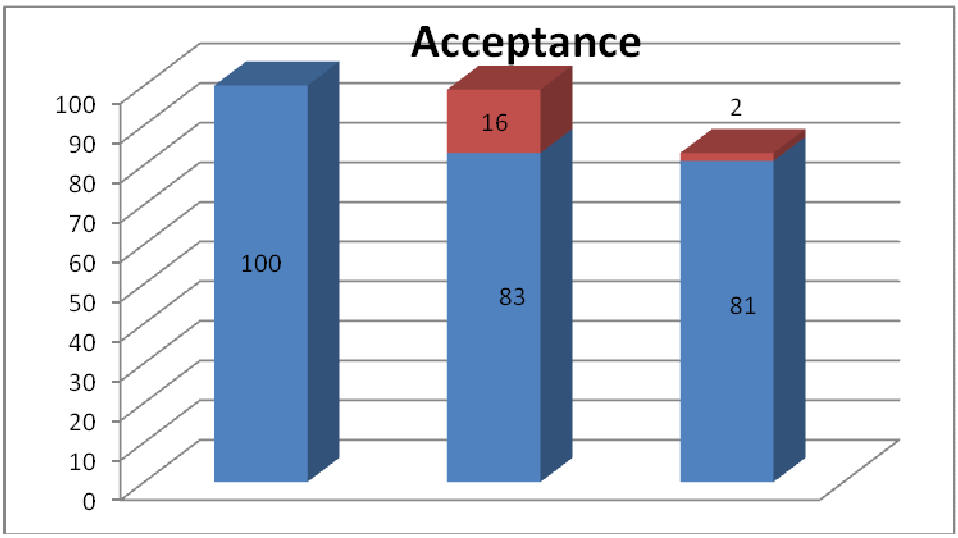

Figure 9-14 Acceptance results (%) for evaluation 1

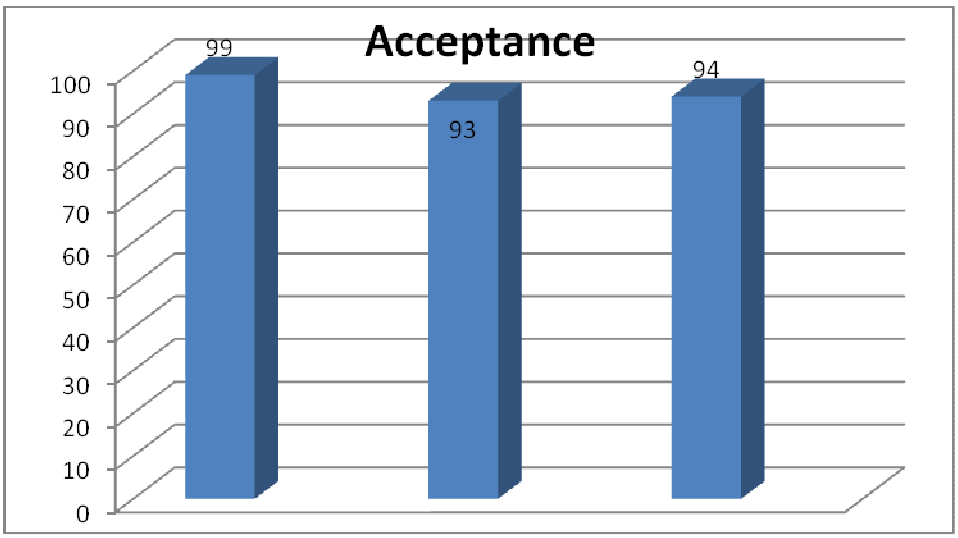

Figure 9-15 Acceptance results (%) for evaluation 2

During-Tasks Questions: Figures 9-16 and 9-17 present a comparison of the two evaluations regarding the successful participant's responses to the during-tasks questions.

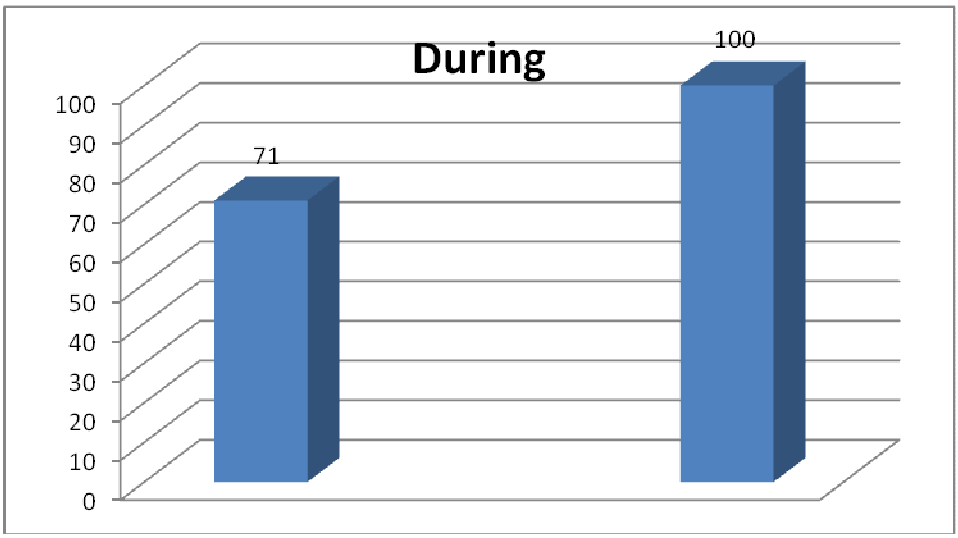

Figure 9-16 During-task questions' results (%) for evaluation 1

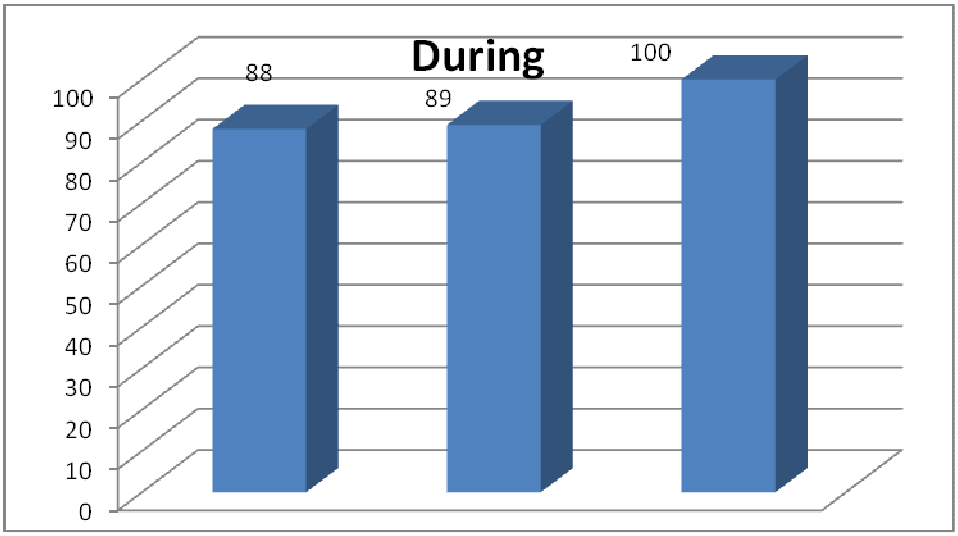

Figure 9-17 During-task questions' results (%) for evaluation 2

Evaluation 1 Summary: Table 9-12 summarizes our conclusions regarding which privilege has been successful in which category – competence, consequence and acceptance for the first evaluation. The X symbol means "not satisfactory", where V marks "success".

Proposed Privacy Privilege	**Competence**	**Consequence**	**Acceptance**
Setting privacy status	X	V	V
	76% of users thought they had managed to complete the task, whereas, in fact, only 68% actually had managed to complete it	All users understood the consequences of their decisions while completing the task, whereas 71% answered the during-task question correctly	All users expressed their acceptance of this privacy privilege
Responding to viewing request	X	V	V

	88% of users thought they had managed to complete the task, whereas, in fact, only 74% actually had managed to complete it	All users understood the consequences of their decisions while completing the task. There was no during-task question for this task	99% of users expressed their acceptance of this privacy privilege
Visiting another user's model and placing a viewing request	V	V	V
	84% of users thought they had managed to complete the task, whereas, in fact, all users actually had managed to complete it	79% actually understood the consequences of their decisions while completing the task, whereas all users answered the during-task question correctly	83% of users expressed their acceptance of this privacy privilege

Table 9-12 Summary of conclusions for proposed privacy privileges for the first evaluation

Evaluation 2 Summary: Table 9-13 summarizes our conclusions regarding which privilege was successful in which category, namely, competence, consequence and acceptance, for the second evaluation. The X symbol means "not satisfactory", while V means "successful".

Proposed Privacy Privilege	**Competence**	**Consequence**	**Acceptance**

Setting privacy preferences for previously added profiles	V	V	V
	97% of users thought they had managed to complete the task, whereas, in fact, 90% actually had managed to complete it	93% actually understood the consequences of their decisions while completing the task, whereas 88% answered the during-task question correctly	99% of users expressed their acceptance of this privacy privilege
Responding to viewing request by another SUMI user and managing assigned privilege	V	V	V
	90% of users thought they had managed to complete the whole task, whereas, in fact, 100% and 87% of users respectively actually had managed to complete the two parts of this task	92% actually understood the consequences of their decisions while completing the task, whereas 89% answered the during-task question correctly	93% of users expressed their acceptance of this privacy privilege

Visiting another user's collection and placing a viewing request	V	V	V
	96% of users thought they had managed to complete the task, whereas, in fact, 99% actually had managed to complete it	90% actually understood the consequences of their decisions while completing the task, whereas all users answered the during-task question correctly	94% of users expressed their acceptance of this privacy privilege

Table 9-13 Summary of conclusions for proposed scrutability privileges for the second evaluation

Conclusions: Similar to the offered scrutability privileges, users approved of the proposed privacy privileges that SUMI offered them in both evaluations. Although it is not as clear as it was with the evaluation on scrutability, users responded better with the holistic approach than with the sequential approach. More specifically, they produced better results on the task for setting the privacy status for their profiles when they were looking the 'whole picture' rather than executing sequential instructions. In addition, a much-improved representation of their user models, which was enhanced with metadata information, could also account for the improved results for this task on the second evaluation. Furthermore, users' competence in responding to a viewing request revealed better results in the second evaluation, when they were also offered the option of managing the customised privacy privileges that were assigned individually to other users. Finally, we observed that users understood the consequences of their decisions in both evaluations. This could be due to the fact that users are familiar with the aspect of privacy of information in contrast to the term 'scrutability', thus producing successful results on understanding the consequences of their decisions, irrespective of the different approaches followed during the two evaluations.

9.6 Post-Questionnaires

The last part of both evaluations was the post-questionnaires. These were presented in the same form as the pre-questionnaires, questions with multiple-choice answers where users had to select one from the available responses. Tables 9-14 to 9-19 show the percentage of each response from the evaluation participants to each question of the post-questionnaires:

Question 1 of both evaluations: Would you use such a service? Do you find it useful?

Available Response	1st Evaluation	2nd Evaluation
Yes, I would use it and I find it very useful	92%	93%
Yes, I would I use it, but I don't know if it's useful	5%	0%
Neutral	3%	7%
No, I wouldn't use it, but I find it useful	0%	0%
No, I wouldn't use it and I don't find it useful	0%	0%

Table 9-14 Post-Questionnaire – Question 1

Question 2 of Evaluation 1: What do you think about SUMI keeping a copy of your information?

Question 2 of Evaluation 2: What do you think of the level of control you had over your SUMI collection while going through the six tasks?

Available Response	1st Evaluation	2nd Evaluation
E1: I find it very useful since SUMI provides two options for importing my information (dynamically and statically) E2: I love it! I had full control over my models inside SUMI	17%	 35%
E1: I am OK with it E2: I am satisfied with the level of control I had inside SUMI	33%	 50%
E1: I don't like it, but I can see the value in doing it	39%	

E2: I am OK with the level of control I had inside SUMI		13%
E1: I understand it, but I don't agree E2: Neutral	8%	 25
E1:I don't like it and I don't see the value in doing it E2: You can do better, I expected more	0%	 0%
E1: I don't understand it E2: I am not satisfied	3%	 0%

Table 9-15 Post-Questionnaire – Question 2

Question 3 (Scrutability) of both evaluations: How much do you value scrutability now, after going through Tasks 1, 2 and 6, when you interact with user modelling providers, such as Facebook and Amazon?

Available Response	**1st Evaluation**	**2nd Evaluation**
Very much	85%	86%
Some value	14%	13%
Neutral	1%	1%
No value	0%	0%
I don't understand the question	0%	0%

Table 9-16 Post-Questionnaire – Question 3

Question 4 (Scrutability) of both evaluations: What other scrutability privileges would you like SUMI to offer to its users?

1st Evaluation (direct quotations)	**2nd Evaluation (direct quotations)**
"I would like subscribed providers to tell me how they have used the data I had given them."	"Ability to remove selected profile that user do not wish to continue being posted on SUMI."

"I would like to be able to add my other models, such as Hi5 and Twitter."	"You guys know everything ...I don't want to give any suggestions to you!!"
"I want to have the option of not allowing SUMI to keep a copy of my data." "I want to have the option of deciding for how long SUMI is allowed to keep my data."	"It would be nice to be able to pick and choose within the sub-categories."
"I don't want to interact with other users. I just want to use the privileges that allow me to add the models and export them to subscribers."	"Exporting content to certain providers, i.e., exporting content that is stored in SUMI from Facebook to Myspace."
"I need more information regarding the security inside this service before allowing you to keep a copy of my information."	"I would like to see music profiles available as well. Possibly PureVolume profiles or a SmartPunk profile. This would allow a user possibly to link with the iTunes store or some other music provider based on their certain musical tastes adding a whole new realm to scrutability in SUMI."
"I would like to be told how my exported information will be used, because that will determine whether I would go on with the transaction or not."	"Hmm ... even though I like simple profiles ... but the features like changing colours and skins of the profiles would surely attract more people..."
	"I think what has been offered is already beyond my expectations. A nicer layout of the pages perhaps?"

Table 9-17 Post-Questionnaire – Question 4

Question 5 (Privacy) of 1st Evaluation: After completing Tasks 3, 4 and 5,how much value does privacy of information have for you now when you interact with user modelling providers, such as Facebook and Amazon?

Question 5 (Privacy) of 2nd Evaluation: User-Controlled Privacy is when users are offered a variety of options for customising their own privacy privileges. After

completing Tasks 3, 4 and 5, how much do you value user-controlled privacy when you interact with user modelling providers, such as Facebook and Amazon?

Available Response	1st Evaluation	2nd Evaluation
Very much	79%	81%
Some value	19%	15%
Neutral	2%	4%
No value	05	0%
I don't understand the question	0%	0%

Table 9-18 Post-Questionnaire – Question 5

Question 6 (Privacy) of both evaluations: What other privacy privileges would you like SUMI to offer to its users?

1st Evaluation	2nd Evaluation
"I would like to be able to create groups of users and assign the same privacy setting to that group."	"I was very satisfied with SUMI and I cannot honestly think of any other privileges I would like to be offered."
"I would like to be able to block people from visiting my SUMI collection – "block" privacy setting"	"I do not know about this one, because I'm not sure I should import my profile data into SUMI."
"I want to be able to take back the privilege I assigned to Maria when I responded to her access request."	"I wish you could also get this on Friendster."
"I want to know who has visited my models and when. "	"Allowing for password protection to profile for any privacy status and to ensure more security."
"I would like to have more privacy settings, like block and ignore. Conditional privileges. For example, if a user visits my model more than 10 times then switch all settings to private for that user."	"Deletion of own account."
"Timed privileges – allow this user to see my profile for 1 hour – then don't allow	"I think that SUMI does a great job of controlling the privacy privileges of your

him anymore – but allow him to request further visits"	profile."
	"Nothing I can think of now." "It's pretty much enough actually."
	"The ability to update you when your profile is viewed, and by whom."

Table 9-19 Post-Questionnaire – Question 6

Finally, a seventh question in the 2nd evaluation asked participants to state whether they had taken part in the first evaluation; 27% answered positively, meaning 73% of the users were not familiar with the SUMI environment during the 2nd evaluation.

Conclusions: The post-questionnaires revealed participants' highly positive attitude towards the SUMI service to which they were exposed; 92%-93% of users approved of our work, a proportion that was reflected in the evaluations, although they did not agree with the feature of SUMI keeping a copy of their information inside its databases. Furthermore, 39% of users in the first evaluation were not convinced of this feature's usefulness, but they did acknowledge that it could be of some value. Therefore, this feature was enhanced with user control mechanisms in the second evaluation to give participants absolute control regarding which elements of their personal information SUMI was allowed to import and keep internally.

Probably the two most important results of this evaluation can be identified in the participants' responses to questions three and five; 85% (1st evaluation) and 86% (2nd evaluation) for scrutability and 79% (1st evaluation) and 81% (2nd evaluation) chose the best answer available when, after completing the evaluation, they were asked how much they valued scrutability and user privacy respectively. If these percentages were to be compared with the responses in the pre-questionnaires, and specifically with the responses to the questions regarding how much users were familiar with the two terms before the evaluation, we would observe a significant increase of percentages in both occasions:

- users that were familiar with the term 'scrutability' in the pre-questionnaires: 11% (1st evaluation) and 10% (2nd evaluation) – users that expressed the best answer available when they were asked in the post-

questionnaires how much they valued scrutability: 85% (1st evaluation) and 86% (2nd evaluation),

- users that were familiar with their privacy options when interacting with social networking and e-commerce providers, in pre-questionnaires: 64% (1st evaluation) and 81% (2nd evaluation) – users that expressed the best answer available when asked in the post-questionnaires how much they valued privacy: 79% (1st evaluation) and 81% (2nd evaluation),

In addition, many useful suggestions emerged from the evaluation, since we specifically asked our users to express their opinions regarding which other scrutability and privacy user privileges they would like SUMI to offer them. This demonstrates the participants' engagement with SUMI, which we think it is beyond satisfactory.

9.7 Summary

In this chapter, we presented the design and results of the two conducted evaluations for the proposed SUMI scrutability and privacy privileges, in order to test our hypotheses, which were stated in sections 6.3.2 and 6.3.3. Results show that users engaged successfully with the idea of having SUMI available to assist them for educational purposes, and this is best reflected in the results of the second evaluation, where we adopted a more holistic approach, which was enhanced with metadata - ontology - information and an example of an adaptive lesson to demonstrate the usefulness of exchanging the various social networking and e-commerce user models with educational systems for potentially improved personalisation services in contrast to the sequential approach of the first evaluation. Furthermore, users favoured customised privacy privileges, which enabled them to have more in-depth control compared to fixed privacy privileges.

Chapter 10 Summary of Contributions and Future Work

10.1 Introduction

This chapter is the concluding chapter for this PhD thesis. First, we refer back to the research questions that were stated in Chapter 6 and verify or dismiss the corresponding hypotheses by summarising our contributions, which have been thoroughly presented in Chapter 7 and verified in Chapters 8 and 9. Furthermore, we explain our own projection regarding the future of Lifelong User Modelling, which is based on the extensive research that was conducted during this PhD. We also describe our future work agenda, which can be considered as a commitment to continue contributing to the area of Lifelong User Modelling.

10.2 Summary of Contributions

In Chapter 7, we presented our proposed solution to the identified problem in the area of Lifelong User Modelling. The various parts of the proposed solution that we introduced were supported by and related to an extensive literature review, and were proved using simulation testing and two hands-on evaluations that were conducted with potential SUMI users. Below we provide the list of our contributions by referring back to the research questions that were presented in Chapter 6 and examining whether our hypotheses are verified or dismissed based on our evaluations' results.

10.2.1 Interoperability

Question 1:

Can we enable the import of user data to occur from social networking and e-commerce providers to the proposed infrastructure?

Hypothesis 1:

We claimed that this would be possible via a web-based service that would allow users to import their information from social networking and e-commerce providers. This service would be responsible for handling the technical requirements when a user initiates import of his/her data from a provider, using the available APIs and/or data portability tools.

Answer 1:

We identified that user modelling standards, such as PAPI and LIP, unfortunately are currently not adopted by social networking and e-commerce providers. Instead, various APIs and data portability projects are offered that can be used to retrieve user data from the providers' websites.

This research question was answered by the affordances of the current state of technology in the two examined domains. With the evaluation we conducted, we confirmed the hypothesis that the offered APIs and data portability tools can be used successfully when retrieving user data for the social networking and e-commerce domains.

Question 2:

Where do services from the two examined domains stand regarding passing the control of information back to its owners?

Hypothesis 2:

It was our hopethat this notion of passing control of information back to its owners will continue to evolve, with newcomers adopting the example of their pioneers, Facebook and OpenSocial in the social networking domain, and Amazon and eBay in the e-commerce domain.

Answer 2:

As shown in Table 7-1, according to the list of the top fifty websites of 2009 based on user enjoyment and ease of navigation (Time.com and CNN, 2009a), 22 social networking and 6 e-commerce websites that made the list have available for use

APIs that can be used to retrieve user data from the respective websites. Looking back to December 2008, where Facebook and OpenSocial were the first to announce their data portability programs (PC World, 2008b), one can identify a movement from the social networking and e-commerce domains towards the adoption of data portability initiatives. Although one can only speculate about any future decisions, which may or may not favour data portability projects even further, currently the evidence clearly shows that services from the two examined domains will continue to introduce features and mechanisms for passing the control of information back to its owners.

Question 3:

Is it possible to design a common model architecture that can accommodate the various user models retrieved from social networking and e-commerce providers?

Hypothesis 3:

We claimed that if we were to design a common models' architecture that would fit the four representatives' data models, then based on the market share these providers hold, it would be safe to assume that the proposed models' architecture would fit other current social networking and e-commerce services as well.

Answer 3:

We have introduced a four-category SUMI models' architecture that can accommodate any user model retrieved from current social networking and e-commerce providers.

Question 4:

Can we allow social networking and e-commerce providers of user models to define their data models in the proposed infrastructure, in order to allow the exchange of user data between them and the proposed infrastructure?

Hypothesis 4:

We claimed that ontologies can adequately be used to describe a provider's data model, from both examined domains, assuming we present the necessary interfaces, via the online service, for doing so. In addition, we also claimed that administrative interactions with the online service would be necessary in order to ensure the proper definition and description of each provider's data model.

Answer 4:

We have implemented a SUMI ontology, which extends the GUMO and FOAF ontologies and is capable of handling the data models and their corresponding relationships of current social networking and e-commerce providers. In the ontology, we have already defined the data models of Facebook and OpenSocial from the social networking domain, and Amazon and eBay from the e-commerce domain. In addition, we have offered the necessary interfaces in the SUMI online service to allow any provider from the two examined domains to define its own data model by mapping it to that of its representatives and/or by extending the SUMI ontology if necessary.

Question 5:

How can we map the different data models, from the social networking and e-commerce providers of user models in the proposed infrastructure, in order to overcome the semantic barriers when importing user data from these providers into the proposed infrastructure?

Hypothesis 5:

Similarly to the hypothesis in question 4, ontologies can be used to map the various data models of social networking and e-commerce providers inside the infrastructure, via the necessary interfaces of the web-based service. We also claimed that mapping the representatives' data models inside the infrastructure in advance would assist other services to relate to them, since we assumed, based on the market share that Facebook, OpenSocial, Amazon and eBay hold in the two examined domains, that these representatives' data models are supersets of the other services' data models.

Answer 5:

As we have demonstrated, ontologies written in OWL language can be used to express adequately the relationships between the various data models of the current social networking and e-commerce providers.

Question 6:

Which communication protocol can best serve the process of exporting user data from the proposed infrastructure towards educational personalisation systems?

Hypothesis 6:

We claimed that the REST protocol, with its simple, but elegant and effective design, can adequately serve us as a communication protocol for exporting user data

towards educational personalisation systems, after direct consent has been obtained from the data owners.

Answer 6:

Although a Web Service would adequately meet the requirements to act as a communication protocol, we have demonstrated that the REST protocol can satisfy the requirements, with much less implementation complexity, for exchanging user information from SUMI towards educational personalisation systems.

Question 7:

How can we enable educational personalisation systems to express interest in receiving specific social networking and/or e-commerce user models?

Hypothesis 7:

As already expressed, we thought that the need to develop an interface to the infrastructure, in the form of an online service, would help us to answer this question as well. The necessary interfaces can allow educational personalisation systems to express interest in user modelling providers by subscribing to them.

In addition, we claimed that user information retrieved from social networking and e-commerce services could bring long-term, even lifetime value to personalisation systems, based on the LUM criteria, which were described in Chapter 4, in specific use cases that we aimed to identify.

Answer 7:

Educational personalisation systems can express their interest in receiving models from specific social networking and/or e-commerce providers by subscribing to these providers via the SUMI service. The subscription approach has been adopted by various contributions in the literature of UM (Dolog, 2004; Dolog *et al.,* 2004; Alrifai *et al.,* 2006) and it has been found adequate to serve the same purpose in the SUMI context. The decision to export their data towards the educational systems or not is assigned to the users, who can view the list of subscriptions when they login to the SUMI service.

10.2.2 Scrutability

Question 8:

What scrutability privileges should be offered to users? How can we evaluate whether the proposed scrutability privileges are adequate for such a proposed infrastructure? How can we evaluate whether potential users will understand and accept the proposed scrutability privileges?

Hypothesis 8:

We claimed that the proposed scrutability user privileges, which we offer to potential users and are presented and evaluated in Chapter 9, would be appropriate to be offered in the proposed infrastructure for the social networking and the e-commerce domains, and accepted by users as well.

Assuming we present three tasks to our users, which are presented in Chapter 9, where each task will expose one of the proposed scrutability user privileges, we claimed that:

a) Users would be able to complete each task successfully, irrespective of how the tasks were presented to them, (competence).
b) Users would express their acceptance of having the proposed user privilege available for use in the proposed infrastructure, irrespective of how the tasks were presented to them, (acceptance).
c) Users would be familiar with their decisions during and after the completion of each task, and would understand the consequences of each decision they took, irrespective of how the tasks were presented to them, (consequence).

We claimed that we could assess whether a proposed scrutability user privilege was appropriate to be offered in the proposed infrastructure by evaluating users':

a) Competence for completing each task (compare actual outcome after the completion of the tasks with users' answers to the evaluation questions [have they actually done it versus do they think they have done it]
b) Understanding of the consequences of their decisions while interacting with each task.

We claimed that we could assess whether a proposed scrutability user privilege was accepted by users by evaluating users' acceptance of the proposed scrutability user privilege by asking them directly what they thought about each task and if they would like the proposed infrastructure to offer it to its users

Answer 8:

We evaluated three scrutability privileges that could be offered to potential users to allow them to have control over their own models. These privileges passed the tests for competence and understanding of consequences during the two user evaluations, and were accepted by potential SUMI users:

- Adding any of their social networking and/or e-commerce models into SUMI while initiating import of their data into SUMI as well, using the available APIs
- Customising various SUMI settings regarding which attributes from each model will be imported. Furthermore, users have the right to delete the value of any attribute that SUMI may hold about them
- Exporting any model or part of a model to a subscribed educational system. Users can inspect the transaction details before approving the export process.

Question 9:

How informed are users today regarding the term 'scrutability'? Do they recognise any scrutability options when interacting with social networking and e-commerce providers today?

Hypothesis 9:

Our assumption was that users would not be familiar with the term 'scrutability', although they may have been exposed to some scrutability privileges when interacting with social networking and e-commerce services.

Answer 9:

According to the results of the evaluation pre-questionnaires, which are presented in section 9.3, 90% of participants did not know what the term 'scrutability' meant. Furthermore, after the term had been explained with real examples, 93% of users found the idea of having scrutability privileges available when interacting with various providers a very good idea, which shows the acknowledgment of the importance of scrutability to users, once it had been explained to them.

10.2.3 Privacy

Question 10:

What privacy privileges should be offered to users? How can we evaluate whether the proposed privacy privileges are adequate for such a proposed infrastructure? How can we evaluate whether potential users will understand and accept the proposed privacy privileges?

Hypothesis 10:

We claimed that the proposed privacy user privileges, which we offered to potential users and which are presented and evaluated in Chapter 9, would be appropriate to be offered in the proposed infrastructure for the social networking and the e-commerce domains, and accepted by users.

Assuming we presented three tasks to our users, which are presented in Chapter 9, where each task would expose one of the proposed privacy user privileges:

a) Users would be able to complete each task successfully, irrespective of how the tasks were presented to them, (competence).
b) Users would express their acceptance of having the proposed user privilege available for use in the proposed infrastructure, irrespective of how the tasks were presented to them, (acceptance).
c) Users would be familiar with their decisions during and after the completion of each task, and would understand the consequences of each decision they took, irrespective of how the tasks were presented to them, (consequence).

We claimed that we could assess whether a proposed privacy user privilege was appropriate to be offered in the proposed infrastructure by evaluating users':

a) Competence for completing each task (compare actual outcome after the completion of the tasks with users' answers to the evaluation questions [have they actually done it versus do they think they have done it]
b) Understanding of the consequences of their decisions while interacting with each task.

We claimed that we could assess whether a proposed privacy user privilege was accepted by users by evaluating users' acceptance of the proposed privacy user privilege by asking them directly what they thought about each task and if they would like the proposed infrastructure to offer it to its users.

Answer 10:

We evaluated three privacy privileges that could be offered to potential users to allow them to express their privacy preferences when interacting with such an

infrastructure. These privileges passed the tests for competence and understanding of consequences during the two user evaluations, and were accepted by potential SUMI users.

- Setting the privacy status for any of the four categories of each one of their previously added models. Three privacy settings were provided: public, private and hidden
- Receiving viewing requests by other SUMI users to access the private categories of their models. Users have the ability to manage any customised privacy privileges they have assigned to other users at any time
- Visiting other users' SUMI collections and placing viewing requests to access any of the private categories of any of their models.

Question 11:

What privacy settings will allow users to define their privacy preferences in the proposed context?

Hypothesis 11:

We believed that three categories of privacy settings would satisfy users' needs in such an infrastructure:

a) Category 1: data that their owner is happy to share with everyone and that do not require strong privacy protection,
b) Category 2: data that their owner wants to keep completely hidden from other users
c) Category 3: data to which their owner would like to apply some level of privacy that cannot be satisfied with the privacy settings of the first or the second categories.

Answer 11:

We demonstrated that three privacy settings, specifically, public, private and hidden, which were designed based on the three categories that we claimed would be adequate for the proposed solution, were accepted by the participants in both evaluations. Furthermore, in the second evaluation, users managed to complete the assigned task using these offered privacy settings and understood the consequences of their decisions while undertaking the corresponding task. In addition users suggested

further privacy settings that would like SUMI to offer to its users, as described in Table 9-19, which we have included in our future work agenda.

Additional results:

As the results of the first evaluation reveal, users did not appreciate the fact that SUMI was keeping a copy of their information when they initiated a dynamic import of their data. Their responses clearly showed that, in order to accept this feature, they needed further clarification and explanation. When the second evaluation addressed that issue and offered the option to instruct SUMI regarding what was allowed to be imported and/or stored in its databases, participants approved that privilege and did not raise any of the objections that they had raised in the first evaluation. Being honest with the users by explaining the reasoning behind such a feature while offering them the option to customise the settings for their SUMI collection, is identified as being the best approach. The advantages of keeping a copy of user data inside SUMI can be realised in situations like high levels of network traffic, which can be observed on services when a high number of users are simultaneously accessing the various providers' APIs. Users' lack of trust in having yet another system keeping a copy of personal information is the main disadvantage of such a feature.

In addition, as the results of the second evaluation reveal, we identified that users appreciate customised privacy privileges although they are not familiar with such options when interacting with current social networking and e-commerce services. With the exception of Facebook, currently, users are not offered any customised privacy privileges.

Finally, our interaction with AHA! during the second user evaluation revealed some interesting conclusions regarding which design requirements should be considered by developers of educational personalisation systems, for handling social networking and e-commerce user data. Firstly, it is important to incorporate content adaptive techniques for the purposes of delivering personalised examples to users during an adaptive course. Although, this can also be offered with adaptive navigation, it is more effective when combined with the former. Secondly, event-condition-action rules are crucial for updating the user model based on attributes' values that can be retrieved from social networking and e-commerce providers. Thirdly, allowing programming code inside the pages of an adaptive course will enable communication with social networking and e-commerce providers for retrieving user data while also

allow posting back to them any potential useful outcome from the interactions of users with the personalisation system. Lastly, a scrutable user model is more powerful when combined with scrutable adaptation, thus allowing users to control not only the way they are been modelled but also the process of generating personalisation services based on their user models, could be the key for gaining users' trust.

10.3 Future Work Agenda

Lifelong User Modelling can play a leading role in achieving the Lifelong Learning vision. Although the community is still discovering new challenges, there are some contributions, as presented in Chapter 4, that have brought us a step closer to the ultimate goal of Lifelong Learning. We hope the work presented in this thesis is one of those contributions that will help the community to address some of the questions raised.

It is our critical view that social networking and e-commerce will play a significant role in LUM. Although the first priority is to exchange user data between educational systems, the two domains should not be neglected in such an attempt, since potentially, the contribution these two domains can offer to the area of LUM is immense. APIs and data portability initiatives will continue to evolve, as recently hinted by Facebook (CNET News, 2009a) and Google (Techtree.com, 2009). We believe that more and more services from the two domains will introduce their own API and/or data portability initiative to allow users to retrieve their personal data from their websites. However, it is also our opinion that there will be no agreement between the various services on adopting similar APIs or data portability projects, since the pioneers in these domains will never agree on helping out each other and struggling newcomers (ReadWriteWeb, 2009a; BBC News, 2009a; The Washington Post, 2008). The only issue that has been agreed wholeheartedly among the various social networking services, which was announced by the European Commission, is a pact to work together to safeguard user data that is posted on their platforms (InformationWeek, 2009b). Furthermore, we believe that, in the future, we will observe the merger of social networking and e-commerce features on online services. Rather than specialising in one domain, providers will incorporate both kinds of features into their websites in an attempt to attract even more users (ChannelWeb Network, 2009e). Furthermore, social networking services will introduce generated user information that

will be based on manually-entered user input, whereas e-commerce services will continue to offer personalised services that will be based mostly on generated information rather than on manual user input. LUM's greatest challenge will be to keep track of changes on users' preferences, interests and goals over long periods of time, even lifetime periods. If this is achieved, it will take us one huge step closer towards the lifelong learning vision.

Our future work schedule includes further research (and implementation) that aims to improve SUMI against the already defined criteria. More specifically, we will focus on the following:

1. Incorporating more services from the social networking and e-commerce domains inside SUMI. This will allow further evaluation of the infrastructure to take place and will attract more users to SUMI.
2. Preparing concrete adaptive lessons with educational personalisation systems, like GALE and InterBook. These lessons will take advantage of SUMI and use the rich sets of user information that SUMI can offer to generate personalised examples for their users. This will also allow there to be further evaluation of the infrastructure and will attract more users to SUMI. Furthermore, it will provide a more insightful perspective regarding the value SUMI brings to education.
3. Continuing the design and implementation of scrutability privileges to be offered in SUMI, which will aim to pass even more control to the users/owners. Further focus will also be on finding out more about users' preferences regarding the storing of their information in SUMI, a concern participants had flagged up during the first evaluation for scrutability and privacy.
4. Continuing the introduction of privacy privileges to be offered in SUMI, which will target gaining users' trust and encourage further interaction among SUMI users. Special consideration will be given to customised privacy options for which users had expressed their appreciation during the second evaluation for scrutability and privacy.
5. Introducing a tracking mechanism in SUMI that will be responsible for measuring the various changes in users' preferences as expressed inside SUMI over long periods of time.

6. Evaluating the phenomenon of memory decay. The reason a user has posted some piece of information on a social networking and/or e-commerce website may no longer exist after a long period of time. That reason may be forgotten by its owner, or even altered under various circumstances. If we find a way to model this aspect of user modelling, then we may be able to assign levels of importance and validity to such shared information.

10.4 Summary

In this chapter, we summarised our contributions to the area of Lifelong User Modelling, which have been presented and evaluated during this PhD thesis. Furthermore, we offered our critical analysis of the future of LUM where we detailed our future work agenda, which expresses our intention to continue contributing to the area.

References

Abel, F., Heckmann, H., Herder, E., Hidders, J., Houben, G.J., Krause, D., Leonardi and Kees, E., and van der Sluijs, K., (2009). *A Framework for Flexible User Profile Mashups*. Proc. Workshop Adaptation and Personaliation for Web 2.0 at UMAP 2009, pp. 1-10.

Abel, F., Heckmann, H., Herder, E., Hidders, J., Houben, G.J., Krause, D., Leonardi and Kees, E., and van der Sluijs, K., (2009). *Mashing up user data in the Grapple User Modeling Framework*. Proc. ABIS 2009.

AHA! Home Page. Last accessed on 03/01/2010 at http://aha.win.tue.nl/.

Allen, C., Yaeckel, K. D. and Yaeckel, B., (1998). *Internet World Guide to One-To-One Web Marketing*, John Wiley and Sons, New York.

Allgayer, J., Harbusch, K., Kobsa, A., Reddig, C., Reithinger, N. and Schmauks, D., (1989). *XTRA: A Natural-Language access system to expert systems*. International Journal of Man-Machine Studies, (31) 161-195.

All Headline News, (2008). *Facebook Becomes Most Popular Social Network*. Last accessed on 03/01/2010 at http://www.allheadlinenews.com/articles/7011938817.

Alrifai, M., Dolog, P., and Nejdl, W., (2006). *Learner Profile Management for Collaborative Adaptive eLearning Application*. APS'2006: Joint International

Workshop on Adaptivity, Personalisation and the Semantic Web at the 17th ACM Hypertext'06 conference. Odense, Denmark. ACM Press.

Amazon Web Services Home Page. Last accessed on 03/01/2010 at http://aws.amazon.com/what-is-aws/.

Amazon Web Services Products. Last accessed on 03/01/2010 at http://aws.amazon.com/products/.

Andromedia Home Page. Last accessed on 03/01/2010 at http://www.andromedia.com/products/likeminds/index.html.

Apted, T, J Kay and A Lum., (2004). *Supporting Metadata Creation with an Ontology Built from an Extensible Dictionary*. In P De Bra, W Nejdl (eds), LNCS vol. 3137, Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems: Third International Conference, AH 2004, Eindhoven, The Netherlands, August 23-26, 2004, pp. 4 – 13.

Armstrong, J., and Franklin, T., (2008). *A Review of Current and Developing International Practice in the Use of Social Networking (Web 2.0) in Higher Education*. Last accessed on 03/01/2010 at http://www.franklinconsulting.co.uk/Reports.html#Topic5.

Aroyo, L., Dolog, P., Houben, G.J., Kravcik, M., Naeve, A., Nilsson, M., and Wild, F., (2006). *Interoperability in Personalised Adaptve Learning*. Educational Technology & Society 9(2): 4-18.

Ars Technica, (2009a). *Anonymized data really isn't - and here's why not*. Last accessed on 03/01/2010 at http://arstechnica.com/tech-policy/news/2009/09/your-secrets-live-online-in-databases-of-ruin.ars.

Ars Technica, (2009b). *Netflix Prize 2 (Privacy) - Apocalypse Now*. Last accessed on 03/01/2010 at http://arstechnica.com/tech-policy/news/2009/09/netflix-prize-2-privacy-apocalypse-now.ars.

Assad, M., Carmichael, D. J., Kay, J, and Kummerfeld, B., (2007). *MyPlace Locator: Flexible sharing of location information*. In Proceedings of UM 2007, 11th International Conference on User Modeling, pages 410-414. Springer.

ATG Home Page. Last accessed on 03/01/2010 at http://www.atg.com/en/products/index.jhtml.

BBC News, (2009a). *Facebook in challenge to Google*. Last accessed on 03/01/2010 at http://news.bbc.co.uk/2/hi/technology/8194508.stm.

BBC News, (2009b). *Facebook offers control to users*. Last accessed on 03/01/2010 at http://news.bbc.co.uk/2/hi/technology/7913289.stm.

BBC News, (2009c). *Online time is good for teens*. Last accessed on 03/01/2010 at http://news.bbc.co.uk/2/hi/technology/7740895.stm.

BBC News, (2009d). *Who polices Facebook*. Last accessed on 03/01/2010 at http://www.bbc.co.uk/blogs/technology/2008/11/who_polices_facebook.html.

Bergman, M., (2009). *Moving Beyond Linked Data*. AI^3: Adaptive Information, Adaptive Innovation, Adaptive Infrastructure. Last accessed on 03/01/2010 at http://www.mkbergman.com/802/moving-beyond-linked-data/.

Berners-Lee, T., (2006). *The Four Rules of Linked Data.* Last accessed on 03/01/2010 at http://www.w3.org/DesignIssues/LinkedData.html.

Berners-Lee, T., Hendler, J., and Lassila, O., (2001). *The Semantic Web.* Scientific American, 284(5):34-43.

Billsus, D. and Pazzani, M. J., (2000). *User modeling for adaptive news access*. User Modeling and User-Adapted Interaction 10(2-3), Special Issue on Deployed UserModeling, 147-180.

Borchers, J., (2008). *An ode to tomtom: Sweet spots and baroque phases of interactive technology lifecycles*. In: Interactions of the ACM 15(2), 62–66.

Bowne and Co Home Page. Last accessed on 07/01/2008 at http://www.bowne.com.

Brajnik, G. and Tasso, C., (1994). *A shell for developing non-monotonic user modelling systems.* International Journal of Human-Computer Studies 40, 31-62.

Brusilovsky, P., (1999). *Adaptive and Intelligent Technologies for Web-based Education*. In C. Rollinger and C. Peylo (eds.), Special Issue on Intelligent Systems and Teleteaching, *Künstliche Intelligenz*, 4, 19-25ACM SIG.

Brusilovsky, P., (2001). *Adaptive Hypermedia*. A. Kobsa, ed, User Modelling and User Adapted Interaction, Ten Year Anniversary, 11, p.p. 87-110.

Brusilovsky, P., (1996a). *Adaptive hypermedia, an attempt to analyze and generalize*. P. Brusilovsky, P. Kommers, & N. Streitz (Eds.), Springer, 288-304.

Brusilovksy, P., (2000). *Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-based education.* Proceedings of 5th International Conference on Intelligent Tutoring Systems, ITS 2000, Montreal, Canada.

Brusilovsky, P., (1995). *Intelligent tutoring systems for World-Wide Web.* In R. Holzapfel (Eds.), Poster proceedings of Third International WWW Conference. Darmstadt, April 10-14, 1995 - pp. 42-45.Schwartz, M., and Task Force on Bias-Free Language. Guidelines for Bias-Free Writing. Indiana University Press, Bloomington IN.

Brusilovsky, P., (1996b). *Methods and techniques of adaptive hypermedia.* User Modeling and User-Adapted Interaction, 6 (2-3), p.p. 87-129.

Brusilovsky, P., (1994). *The construction and application of student models in intelligent tutoring systems*. Journal of Computer and System Sciences International, 32(1), 70-89.

Brusilovsky, P., Eklund, J., and Schwarz, E., (1998). *Web-based education for all: A tool for developing adaptive courseware*. Computer Networks and ISDN Systems, 30 (1-7), 291-300.

BusinessWeek, (2009). *More Adults Joining Social Networks*. Last accessed on 03/01/2010 at http://www.businessweek.com/the_thread/techbeat/archives/2009/01/more_adults_joi.html?chan=technology_technology+index+page_top+stories.

CHAI LUM Project. Last accessed on 03/01/2010 at http://chai.it.usyd.edu.au/Projects/LLUM.

ChannelWeb Network, (2009a). *Amazon Settles Kindle Lawsuit, Hopes To Put 'Big Brother' To Rest*. Last accessed on 03/01/2010 at http://www.crn.com/retail/220300747;jsessionid=5VOXEZQQIIUDNQE1GHPCKHWATMY32JVN.

ChannelWeb Network, (2009b). *Facebook Opens Up To OpenID*. Last accessed on 03/01/2010 at http://www.crn.com/software/217500811;jsessionid=BN0HAKAOE2X3YQSNDLPCKH0CJUNN2JVN.

ChannelWeb Network, (2009c). *Facebook Registers 200 Million Users As Social Networks' Influence Grows*. Last accessed on 03/01/2010 at http://www.crn.com/networking/216403582.

ChannelWeb Network, (2009d). *Facebook-Want To Keep That Update Private - Go Ahead*. Last accessed on 03/01/2010 at http://www.crn.com/software/218101325;jsessionid=1PXSBQPSQZIF2QSNDLOSKHSCJUNN2JVN.

ChannelWeb Network, (2008). *Five Social Networks That Are Growing Worldwide*. Last accessed on 03/01/2010 at http://www.crn.com/it-

channel/210004191;jsessionid=TFRXILLLE42W0QSNDLPSKHSCJUNN2JVN?pgno=1.

ChannelWeb Network, (2009e). *Five Ways Businesspeople Can Use Social Networks*. Last accessed on 03/01/2010 at http://www.crn.com/it-channel/212900738.

CNET News, (2009a). *Facebook woos developers with a roadmap*. Last accessed on 03/01/2010 at http://news.cnet.com/8301-27076_3-10385876-248.html.

CNET News, (2009b). *Gartner - Brace yourself for cloud computing*. Last accessed on 03/01/2010 at http://news.cnet.com/8301-30685_3-10378782-264.html.

CNET News, (2008). *Missing girl's father finds help through Twitter*. Last accessed on 03/01/2010 at http://news.cnet.com/8301-17939_109-10065800-2.html.

CNET News, (2009c). *Trapped in a drain...try Facebook for help*. Last accessed on 03/01/2010 at http://news.cnet.com/8301-17939_109-10347498-2.html.

CNET News, (2009d). *Users file privacy lawsuit against Facebook*. Last accessed on 03/01/2010 at http://news.cnet.com/8301-1023_3-10311625-93.html.

ComputerWorld, (2002). *The Web's Best-Seller*. Last accessed on 03/01/2010 at http://www.computerworld.com/s/article/74619/The_Web_s_Best_seller?taxonomyId=071.

Computing Research Association Home Page. Last accessed on 03/01/2010 at http://www.cra.org/.

Cruz I.F., (2003). *Semantic Data Integration*. UMSI: Workshop on Data Mining and Exploration Middleware for Distributed and Grid Computing, Minneapolis, Minnesota. Last accessed on 03/01/2010 at http://www.cs.uic.edu/~ifc/Talks/Minnesota/cruz.pdf.

DataPortability Project Home Page. Last accessed on 03/01/2010 at http://www.dataportability.org/.

DailyTech, (2009). *Facebook Modifies Privacy Controls For Users*. Last accessed on 03/01/2010 at http://www.dailytech.com/Facebook+Modifies+Privacy+Controls+For+Users/article15585c.htm.

De Bra, P., and Calvi, L., (1998). *AHA: a Generic Adaptive Hypermedia System*, Proc. of the 2nd Workshop on Adaptive Hypertext and Hypermedia, pp. 5-12.

Deitel, H.M., Deitel, P.J., DuWaldt, B., and Trees, L.K., (2003). *Web Services: a Technical Introduction*, Prentice Hall, Upper Saddle River, New Jersey.

Deitel, H.M, Deitel, P.J, Nieto, T., Lin, T., and Sadhu, P., (2001). *XML How to Program,* Prentice Hall.

DERI Pipes Home Page. Last accessed on 03/01/2010 at http://pipes.deri.org/.

Devedzic, V., Jovanovic, J., and Gasevic, D., (2007). *The Pragmatics of Current E-Learning Standards*, IEEE Internet Computing 11, 3 (May. 2007), 19-27.

DERI Home Page. Last accessed on 03/01/2010 at http://www.deri.ie/.

Dolog, P., (2004). *Identifying Relevant Fragments of Learner Profile on the Semantic Web*. In Proc. of Workshop on Applications of Semantic Web Technologies for E-learning - SW-EL'04: Semantic Web for E-Learning at International semantic Web Conference, Hiroshima, Japan.

Dolog, P., Henze, N., Nejdl, W., and Sintek, M., (2004). *Personalisation in Distributed eLearning Environments*. In Proc. of WWW2004 - The Thirteen International World Wide Web Conference, New York, USA, ACM Press.

Dolog, P., Kay, J., and Kummerfeld, B., (2009). *Personal lifelong user model clouds*. In Proceeding of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation, and Personalization, pages 1-8.

Dolog, P., and Nejdl, W., (2003). *Challenges and Benefits of the Semantic Web for User Modelling*. In Proc. of AH2003 workshop at 12th World Wide Web Conference, Budapest, Hungary.

Dolog, P., and Schaefer, M., (2005a). *A Framework for Browsing, Manipulating and Maintaining Interoperable Learner Profiles.* In Proc. of UM2005 - 10th International Conference on User Modelling, July, Edinburgh, UK. Springer Verlag.

Dolog, P., and Schaefer, M., (2005b). *Learner modelling on the semantic web*. In Proc. of PerSWeb-2005 Workshop: Personalisation on the Semantic Web at User Modelling 2005: 10th International Conference, July, Edinburgh, UK.

DreamGrow Digital. *Market Share of Social Networking Sites in 2009*. Last accessed on 03/01/2010 at http://www.dreamgrow.com/market-share-of-social-networking-sites-2009/.

D2R Server Home Page. Last accessed on 03/01/2010 at http://www4.wiwiss.fu-berlin.de/bizer/d2r-server/.

eBay API Home Page. Last accessed on 03/01/2010 at http://developer.ebay.com/common/api.

eBay Developers Program. Last accessed on 03/01/2010 at http://developer.ebay.com/products/overview.

Elliot, D., Hopfgartner, F., Leelanupab, T., Moshfeghi, Y., and Jose, JM., (2009). *An Architecture for Life-long User Modeling*. In: Proceeding of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation, and Personalization, June 22-26, 2009, Trento, Italy.

eWeek.com, (2006). *Web 2.0: How High-Volume eBay Manages its Storage*. Last accessed on 03/01/2010 at http://www.eweek.com/c/a/Data-Storage/Web-20-How-HighVolume-eBay-Manages-Its-Storage/.

Facebook Company Profile. Last accessed on 03/01/2010 at http://www.crunchbase.com/company/facebook.

Facebook Connect Home Page. Last accessed on 03/01/2010 at http://developers.facebook.com/connect.php.

Facebook Developers Home Page. Last accessed on 03/01/2010 at http://wiki.developers.facebook.com/index.php/Platform_Core_Components.

Facebook Developers News, (2008). *Announcing Facebook Connect*, Last accessed on 03/01/2010 at http://developers.facebook.com/news.php?blog=1&story=108.

Facebook Official Statistics. Last accessed on 03/01/2010 at http://www.facebook.com/press/info.php?statistics.

Fink, J., and Kobsa, A., (2000). *A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web*. User Modeling and User-Adapted Interaction - Special Issue on Deployed User Modeling 10, 209–249.

Flixster Home Page. Last accessed on 03/01/2010 at http://www.flixster.com/.

FOAF Vocabulary Specification. Last accessed on 03/01/2010 at http://xmlns.com/foaf/spec.

Friend Connect Home Page. Last accessed on 03/01/2010 at http://www.google.com/friendconnect.

Gigaom, (2009). *Europe's Privacy Concerns May Hinder Facebook*. Last accessed on 03/01/2010 at http://gigaom.com/2009/06/18/europes-privacy-concerns-may-hinder-facebook/.

GRAPPLE Project Home Page. Last accessed on 03/01/2010 at http://www.grapple-project.org/.

Hagen, P., Manning, H. and Souza, R., (1999). *Smart Personalization*. Forrester Research, Cambridge, MA.

Heckmann, D., Loskyll, M, Math, R., Recktenwald, P., and Stahl, C., (2009). *UbisWorld 3.0 - a Semantic Tool Set for Ubiquitous User Modeling*. In: First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation (UMAP2009) - Demo, 22-26/06/2009, Trento, Italy.

Heckmann, D., Schwarzkopf, E., Mori, J., Dengler, D., and Kroner, A., (2007). *The user model and context ontology gumo revisited for future web 2.0 extensions*. In P. Bouquet, J. Euzenat, C. Ghidini, D. L.McGuinness, L. Sera_ni, P. Shvaiko, and H. Wache, editors, C&O:RR, volume 298 of CEUR Workshop Proceedings. CEUR-WS.org, 2007.

Hof, R. Green,H. and Himmelstein, L., (1998). *Now it's YOUR WEB*. Business Week (October 5), 68-75.

Huang, X., McCalla, G. I., Greer, J. E. and Neufeld, E., (1991). *Revising deductive knowledge and stereotypical knowledge in a student model*. User Modeling and User Adapted Interaction, 1(1) 87-115.

IEEE Computer Society, (2002). Last accessed on 03/01/2010 at http://www.computer.org/portal/web/csdl/doi/10.1109/NISS.2009.271.

IEEE PAPI Standard. Last accessed on 03/01/2010 at http://www.ieee.org/web/standards/home/index.html.

IMS LIP Specification. Last accessed on 03/01/2010 at http://www.imsglobal.org/profiles/lipbest01.html.

InformationWeek, (2009a). *Facebook Promises Stronger Privacy Controls*. Last accessed on 03/01/2010 at

http://www.informationweek.com/news/internet/social_network/showArticle.jhtml?articleID=219500358.

InformationWeek, (2009b). *Social Networking Sites Ink Safety Pact*. Last accessed on 03/01/2010 at http://www.informationweek.com/news/internet/social_network/showArticle.jhtml?articleID=213402607&subSection=Privacy.

InterBook Home Page. Last accessed on 03/01/2010 at http://www.contrib.andrew.cmu.edu/~plb/InterBook.html.

Internet Retailer, (2009). *The Top 500 Guide*. Last accessed on 03/01/2010 at http://www.internetretailer.com/article.asp?id=30594.

IT Professional Blog, (2008). *Representational State Transfer (REST).* Last accessed on 03/01/2010 at http://mauriziostorani.wordpress.com/2008/07/27/rest-representational-state-transfer-and-restful-web-services-concepts-and-examples/.

Jameson, A., and Gabrielli, S., (2009). *A Preference Evolution Perspective on Lifelong User Modelling*. In: Proceeding of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation, and Personalization, June 22-26, 2009, Trento, Italy.

Kass, R., (1988). *Acquiring a Model of the User's Beliefs from a Cooperative Advisory Dialog*. Ph.D. Thesis, Dept. of Information and Computer Science, University of Pennsylvania, Philadelphia, PA.

Kay J., (1997). *Learner know thyself: Student models to give learner control and responsibility*. In AACE, International Conference on Computers in Education, pages 17–24.

Kay, J., (2006). *Scrutable adaptation: because we can and must*. In Proceedings of 4th International Conference for Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2006, in Dublin, Ireland; Wade, V., Ashman, H., Smyth, B. (eds.), Springer (pub.), pp. 11-19.

Kay, J., (2000). *Stereotypes, Student Models and Scrutability*. In Proceedings of the 5th international Conference on intelligent Tutoring Systems (June 19 - 23, 2000). G. Gauthier, C. Frasson, and K. VanLehn, Eds. Lecture Notes in Computer Science, vol. 1839, 19-30, London. Springer-Verlag.

Kay, J., (1995). *The um toolkit for cooperative user modelling*, User Modelling and User-Adapted Interaction 4(3), 149-196, Kluwer.

Kay, J., and B. Kummerfeld, B., (2009a). *Lifelong user modelling goals, issues and challenges*. In Proceeding of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation, and Personalization, pages 27-34.

Kay, J., and Kummerfeld, B., (2009b). *Proceeding of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation, and Personalization*. Last accessed on 03/01/2010 at http://www.cs.usyd.edu.au/~llum/2009_UMAP09_llum/2009_UMAP_LLUM.pdf.

Kay, J., Kummerfeld, R.J., and Lauder, P., (2002). *Personis: a server for user models*, De Bra, P., P. Brusilovsky, R. Conejo (eds), Proceedings of AH'2002, Adaptive Hypertext 2002, 203 – 212, Springer-Verlag.

Knutov, E., De Bra, P., and Pechenizkiy, M., (2009). *AH - 12 years later: a comprehensive survey of adaptive hypermedia methods and techniques*. New Review of Hypermedia and Multimedia, 15(1), 5-38.

Kobsa, A., (1985). *Benutzermodellierung in Dialogsystemen*. Springer Verlag, Berlin, Heidelberg.

Kobsa, A., (2001). *Generic User Modelling Systems*. User Modelling and User-Adapted Interaction 11(1-2), 49-63.

Kobsa, A., (1990). *Modeling the user's conceptual knowledge in BGP-MS, a user modeling shell system*. Computational Intelligence, (6) 193-208.

Kobsa, A., (2007). *Privacy-Enhanced Personalization* (cover article). Communications of the ACM 50(8), 24-33.

Kobsa, A., (1995). *Using Modeling and User-Adapted Interaction* 4(2), Special Issue on User Modeling Shell Systems, iii-v.

Kobsa, A., and Pohl, W., (1995). *The BGP-MS user modeling system*. User Modeling and User-Adapted Interaction 4(2), 59-106.

Kobsa A, and Wahlster, W., (1989). *User Models in Dialog Systems*. Springer Verlag, Heidelberg, Berlin.

Kono, Y., Okeda, M. and Mizoguchio, R., (1994). *THEMIS: A nonmonotonic inductive student modeling system*. Journal of Artificial Intelligence in Education, 5(3) 371-413.

Kyriacou E, D., (2008). *A Scrutable User Modelling Infrastructure for enabling life-long User Modelling.* In: Adaptive Hypermedia and Adaptive Web-Based Systems - Doctoral Consortium, 29 July - 1 August 2008, Hannover, Germany.

Kyriacou E, D., (2009). *Enriching Lifelong User Modelling with the Social e-Networking and e-Commerce "Pieces of the Puzzle"*. In: First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation (UMAP2009) - Doctoral Consortium, 22-26/06/2009, Trento, Italy.

Kyriacou E, D. and Davis, H., (2008). *Moving towards life-long User Modeling*. In: The 8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008), July 1st- July 5th, 2008, Santander, Cantabria, Spain.

Kyriacou E, D., Davis, H. and Tiropanis, T., (2009a). *A (multi'domain'sional) Scrutable User Modelling Infrastructure for Enriching Lifelong User Modelling*. In: Lifelong User Modelling Workshop (in conjunction with conference UMAP 2009), 26 June 2009, Trento, Italy.

Kyriacou, D., Davis, H. and Tiropanis, T., (2009b). *Evaluating Three Scrutability and Three Privacy User Privileges for a Scrutable User Modelling Infrastructure*. In: First and Seventeenth International Conference on User Modeling, Adaptation, and Personalisation (UMAP2009), 22-26/06/2009, Trento, Italy.

Lacy, L., (2005). *OWL: Representing Information Using the Web Ontology Language*, Trafford Publishing.

Leonardi, E., Houben, G.J., Van der Sluijs, K., Hidders, J., Herder, E., Abel, F., Krause, D., and Heckmann, D., (2009). *User Profile Elicitation and Conversion in a Mashup Environment*. Paper at the International Workshop on Lightweight Integration on the Web.

Linked Data Home Page. Last accessed on 03/01/2010 at http://linkeddata.org/.

Manna Inc Home Page. Last accessed on 03/01/2010 at http://www.mannainc.com.

MediaPost News, (2009). *Facebook To Wind Down Beacon To Resolve Privacy Lawsuit*. Last accessed on 03/01/2010 at http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=113848.

Murray, T., (1999). *Authoring Intelligent Tutoring Systems: an analysis of the state of the art*. International Journal of Artificial Intelligence in Education, 10, 98-129.

Net Perceptions Home Page. Last retrieved on 25/12/2009 at http://www.netperceptions.com.

Netflix Home Page. Last accessed on 03/01/2010 at http://www.netflix.com/.

OAuth Home Page. Last accessed on 03/01/2010 at http://oauth.net/.

OpenID Home Page. Last accessed on 03/01/2010 at http://openid.net/.

OpenSocial API Home Page. Last accessed on 03/01/2010 at http://www.opensocial.org/Technical-Resources/opensocial-spec-v09/OpenSocial-Specification.html.

OpenSocial Containers. Last accessed on 03/01/2010 at http://wiki.opensocial.org/index.php?title=Containers.

OpenSocial REST API Home Page. Last accessed on 03/01/2010 at http://www.opensocial.org/Technical-Resources/opensocial-spec-v09/REST-API.html.

OpenSocial Home Page. Last accessed on 03/01/2010 at http://wiki.opensocial.org/index.php?title=Main_Page#Container_Information.

Orwant, J., (1995). *Heterogenous learning in the Doppelganger user modeling system*. User Modeling and User-Adapted Interaction 4(2), 107-130.

Ounnas, A., Davis, H. and Millard, D., (2008). *A Framework for Semantic Group Formation*. In: The 8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008), July 1st- July 5th, 2008, Santander, Cantabria, Spain. pp. 34-38.

Paget Home Page. Last accessed on 03/01/2010 at http://code.google.com/p/paget/.

Paiva, A., and Self, J., (1995). *TAGUS - A user and learner modeling workbench*. User Modeling and User-Adapted Interaction 4(3), 197-226.

Pandia Search Engine News. *Top 5 Semantic Search Engines*. Last accessed on 03/01/2010 at http://www.pandia.com/sew/1262-top-5-semantic-search-engines.html.

Passin, T., (2004). *Explorers Guide to the Semantic Web*. Manning, Greenwich CT.

PC Magazine, (2008a). *Facebook Connect Goes Live*. Last accessed on 03/01/2010 at http://www.pcmag.com/article2/0,2817,2335707,00.asp.

PC Pro, (2009a). *Facebook pays $9.5 million to bury Beacon*. Last accessed on 03/01/2010 at http://www.pcpro.co.uk/news/351721/facebook-pays-9-5-million-to-bury-beacon.

PC World, (2008b). *Facebook, Google Launch Data Portability Programs to All*. Last accessed on 03/01/2010 at http://www.pcworld.com/article/154988/facebook_google_launch_data_portability_programs_to_all.html.

PC World, (2009b). *Facebook to Tighten Privacy After Canadian Investigation*. Last accessed on 03/01/2010 at http://www.pcworld.com/businesscenter/article/170947/facebook_to_tighten_privacy_after_canadian_investigation.html.

PC World, (2009d). *Netflix Unveils $1 Million Prize Winners*. Last accessed on 03/01/2010 at http://www.pcworld.com/article/172370/netflix_unveils_1_million_prize_winners.htm.

PC World, (2009e). *New Irresponsible Netflix Contest May Violate Customer Privacy*. Last accessed on 03/01/2010 at http://www.pcworld.com/businesscenter/article/172373/new_irresponsible_netflix_contest_may_violate_customer_privacy.html.

Personal Reader Home Page. Last accessed on 03/01/2010 at http://www.personal-reader.de/wp/.

Powers, S., (2003). *Practical RDF*, 1st edition. Beijing; Köln: O´Reilly.

Product Advertising API Home Page. Last accessed on 03/01/2010 at http://docs.amazonwebservices.com/AWSECommerceService/2009-10-01/DG/.

ReadWriteWeb, (2009a). *Google's New Social Search Is A Big Chess Move Against Facebook*. Last accessed on 03/01/2010 at http://www.readwriteweb.com/archives/google_social_search_facebook.php.

ReadWriteWeb, (2009b). *Interview with Tim Berners-Lee, Part 1: Linked Data*. Last accessed on 03/01/2010 at http://www.readwriteweb.com/archives/interview_with_tim_berners-lee_part_1.php.

Reuters, (2009). *Online networking more popular than email*. Last accessed on 03/01/2010 at http://www.reuters.com/article/newsOne/idUSTRE52A6E420090311.

Shibboleth Home Page. Last accessed on 03/01/2010 at http://shibboleth.internet2.edu/.

Simon, B., Dolog, P, Miklos, Z., Olmedilla, D, and Sintek, M., (2004). *Conceptualising smart spaces for learning*. Journal of Interactive Media in Education. Special Issue on Educational Semantic Web, (9):221–239, May 2004.

Sleeman, D. (1985). *UMFE: A user modelling front-end subsystem*. International Journal of Man-Machine Studies, (23): 71-88.

sqlREST Home Page. Last accessed on 03/01/2010 at http://sqlrest.sourceforge.net/.

Tabulator Home Page. Last accessed o 25/12/2009 at http://www.w3.org/2005/ajar/tab.

Taylor, J. A., Carletta, J. and Mellish, C., (1996). *Requirements for belief models in co-operative dialogue*. User Modeling and User-Adapted Interaction, 6(1) 2368.

TechCrunch, (2009). *A map of Social (Network) Dominance*. Last accessed on 03/01/2010 at http://www.techcrunch.com/2009/06/07/a-map-of-social-network-dominance/.

TechCrunch, (2007). *Details Revealed: Google OpenSocial To Launch Thursday*. Last accessed on 03/01/2010 at http://www.techcrunch.com/2007/10/30/details-revealed-google-opensocial-to-be-common-apis-for-building-social-apps/.

Techtree.com, (2009). *Google Wants to Know Your Preferences*. Last accessed on 03/01/2010 at

http://www.techtree.com/India/News/Google_Wants_to_Know_Your_Preferences/551-97934-643.html.

Thang, M.D., Dimitrova, V., Djemame, K., (2007). *Personalised Mashups: Opportunities and Challenges for User Modelling*. In Proceedings of User Modeling 2007, p.415-419.

The Register, (2009). *Canucks crack whip on Facebook privacy*. Last accessed on 03/01/2010 at http://www.theregister.co.uk/2009/08/27/canadian_privacy_commissioner_facebook_changes/.

The Washington Post, (2008). *Biggest Battle Yet For Social Networks - You, Your Identity And Your Data On The Open Web*. Last accessed on 03/01/2010 at http://www.washingtonpost.com/wp-dyn/content/article/2008/11/30/AR2008113002453.html.

TG Daily, (2009). *Facebook privacy now a user-controlled option*. Last accessed on 03/01/2010 at http://www.tgdaily.com/html_tmp/content-view-41750-108.html.

Time.com and CNN, (2009a). *Fifty Best Websites of 2009*. Last accessed on 03/01/2010 at http://www.time.com/time/specials/packages/0,28757,1918031,00.html.

Time.com and CNN, (2009b). *The 25-Best E-Commerce Sites*. Last accessed on 03/01/2010 at http://www.time.com/time/digital/reports/ecommerce/25best.html.

UK Computing Research Committee Home Page. Last accessed on 03/01/2010 at http://www.ukcrc.org.uk/.

Vabour Home Page. Last accessed on 03/01/2010 at http://vapour.sourceforge.net/.

Van der Sluijs, K., Hidders, J., Leonardi, E., and Houben. G.J., (2009). *GAL: A Generic Adaptation Language for describing Adaptive Hypermedia*. In: Proceedings of the International Workshop on Dynamic and Adaptive Hypertext: Generic Frameworks,

Approaches and Techniques (DAH'09) in conjunction with the 20th ACM Conference on Hypertext and Hypermedia (Hypertext 2009), Torino, Italy, 29 June 2009.

Van der Sluijs, K., and Hover, M., (2009). *Integrating adaptive functionality in a LMS.* In: Proceedings of the International Conference on E-Learning in the Workplace (ICELW2009) , 10 - 12 June 2009, New York, USA.

Vernero, F., Petromilli, A., Cena, F., and Gena, C., (2009). *An Experiment to Evaluate how to Better Present User Models to the Users*. In: Proceedings of the User-Centred Design and Evaluation of Adaptive Systems at UMAP'09 User Modeling Adaptation, and Personalization, pages 21-33, June 22-26, 2009, Trento, Italy.

Wang, Y., Vassileva, J., (2009). *A User-Centric Authentication and Privacy Control Mechanism for User Model Interoperability in Social Networking Sites*. In: Proceedings of Adaptation and Personalization for Web 2.0 at UMAP'09 User Modeling Adaptation, and Personalization, pages 110-119, June 22-26, 2009, Trento, Italy.

W3C Home Page. Last accessed on 03/01/2010 at http://www.w3.org/.

W3Schools – Introduction to OWL. Last accessed on 03/01/2010 at http://www.w3schools.com/RDF/rdf_owl.asp.

W3Schools – Architecture of Web Services. Last accessed on 03/01/2010 at http://www.w3.org/TR/ws-arch/.

Yahoo Pipes Home Page. Last accessed on 03/01/2010 at http://pipes.yahoo.com/pipes/.

Zitgist Home Page. Last accessed on 03/01/2010 at http://fgiasson.com/blog/index.php/2007/06/20/zitgists-rdf-browser-browse-the-semantic-web/.

Zukerman, I, and Litman, D., (2001). *Natural language processing and user modeling: synergies and limitations*. User Modeling and User-Adapted Interaction, 11(1-2) 129-158.

Appendix A Description of Facebook & OpenSocial APIs

Facebook RESTful API:

http://wiki.developers.facebook.com/index.php/Users.getInfo

Users:

- *facebook.users.getInfo:*
    - User ID: “uid”
    - About Me: “about_me”
    - Activities: “activities”
    - Network Affiliations: “affiliations”
        - Network ID: “nid”
        - Year: “year”
        - Type: “type”
            - College network: “college”
            - High school network: “high school”
            - Work network: “work”
            - Geography network: “region”
        - Graduate status: “status”
        - Name of network: “name”
    - Birthday: “birthday”
    - Favorite Books: “books”
    - Current Location: “current_location”
        - City: “city”
        - State: “state”
        - Country: “country”
        - Zip Code: “zip”
    - School Information: “education_history”
        - Year: “year”
        - Name of school: “name”
        - List of courses: “concentration”
        - Degree field: “degree”

    - Name: “name”
        - First name: “first_name”
        - Last name: “last_name”
    - Hometown: “hometown_location”
        - City: “city”
        - State: “state”
        - Country: “country:
        - Zip Code: “zip”
    - High School information: “hs_info”
        - High School 1: “hs1_name”
        - High School 2: “hs2_name:
        - Graduation Year: “grad_year”
    - Interests: “interests”
    - Looking for: “meeting_for”
    - Interested in: “meeting_sex”
    - Favorite Movies: “movies”
    - Favorite Music: “music”
    - Number of notes written by user: “notes_count”
    - Profile Picture: “pic”
        - 200x600: “pic_big”
        - 50x150: “pic_small”
        - 50x50: “pic_square”
    - Political Views: “political”
    - Profile Last Updated: “profile_update_time”
    - Favorite Quotes: “quotes”
    - Relationship Status: “relationship_status”
        - Significant Other: “significant_other_id”
    - Religious Views: “religion”
    - Gender: “sex”
    - Status: “status”
    - Timezone: “timezone”
    - Favorite TV shows: “tv”
    - Number of wall posts: “wall_count”
    - Work History: “work_history”
        - Location: “location”
        - Company Name: “company_name”
        - Description: “description”
        - Position: “position”
        - Start Date: “start_date”
        - End Date: “end_date”
- *facebook.friends.get:* returns the ids of the current user’s friends
- *facebook.notifications.get:* returns the notifications of a user

Events:

- *facebook.events.get*: can be used to find all events of a user
    - User ID to find events: “uid”
    - Events IDs: “eids”
    - Start time: “start_time”
    - End time: “end_time”

    - RSVP status: “rsvp_status”
    - Event photo
        - 100x300: “pic”
        - 200x600: “pic_big”
        - 50x150: “pic_small”
- *facebook.events.getMembers*: membership list associated with an event

Groups:

- *facebook.groups.get*: to find all groups a user is a member
    - User ID to find groups: “uid”
    - Groups IDs: “gids”
    - Group photo
        - 100x300: “pic”
        - 200x600: “pic_big”
        - 50x150: “pic_small”
- *facebook.groups.getMembers*: returns membership list associated with a group

Photos:

- *facebook.photos.get*: returns all visible photos of a user
    - Photos sizes:
        - 130x130: “src”
        - 604x604: “src_big”
        - 75x225: “src”small”
- *facebook.photos.getAlbum:* returns all albums of a user
- *facebook.photos.getTags:* returns the set of user tags on all photos specified
    - Photos IDs: “pids”

Marketplace:

- *facebook.marketplace.getCategories*: returns the top level marketplace categories
    - *facebook.marketplace.getSubcategories:* returns the selected marketplace’s subcategories
- *facebook.marketplace.getListings:* returns all marketplace listings of a user
- *facebook.marketplace.search:* searches a user’s networks for listings matching the category, subcategory or query provided to the method
    - Category: “category”
    - Subcategory: “subcategory”
    - Query: “query”

OpenSocial JavaScript API

http://www.opensocial.org/Technical-Resources/opensocial-spec-v09/OpenSocial-Specification.html#opensocial.Person.Field

Person: *Class opensocial.Person*

- ID: “id
- Name: “name”
    - Family Name: “familyName”
    - Given Name: “givenName”
    - Additional Name: “additionalName”
    - Prefix: “honorificPrefix”
    - Suffix: “honorificSuffix”
    - Unstructured Name: “unstructured”
- Nickname: “nickname”
- Photo: “thumbnailUrl”
- Profile: “profileUrl”
- Current Location: “currentLocation”
- Addresses: “addresses”
    - Type of Address (work, home, etc.): “type”
    - Unstructured Address if the container does not support format for address: “unstructuredAddress”
    - PO Box: “poBox”
    - Street Address: “streetAddress”
    - Region: “region”
    - Locality: “locality”
    - Postal Code: “postalCode”
    - Country: “country”
    - Lattitude: “latitude”
    - Longitude: “longitude”
- Emails: “emails”
    - Type of Email: “type”
    - Email Address: “address”
- Phone Numbers: “phoneNumbers”
    - Type of Phone: “type”
    - Phone Number: “number”
- About Me: “aboutMe”
- Status: “status”
- Profile Song: “profileSong”
- Profile Video: “profileVideo”
- Gender: “gender”
    - Male: “male”
    - Female: “female”
- Sexual Orientation: “sexualOrientation”
- Relationship Status: “relationshipStatus**”**
- Age: “age”
- Date of Birth: “dateOfBirth”
- Body Type: “bodyType”
    - Build: “build”
    - Height: “height”
    - Weight: “weight”
    - Eye Color: “eyeColor”
    - Hair Color: “hairColor”
- Ethnicity: “ethnicity”

- Is Smoker: “smoker” -> Enum
- Is Drinker: “drinker” -> Enum
- Description of user’s Children: “children”
- Description of user’s Pets: “pets”
- Description of user’s Living Arrangement: “livingArrangement”
- User’s Time Zone: “timeZone”
- Languages Spoken: “languagesSpoken”
- Jobs the user has held: “jobs” -> Organisation
- Job Favorite Jobs: “jobInterests”
- Schools the user has attended: “schools” -> Organisation
- Interests: “inerests”
- URLs: “urls”
    - Type of URL: “type”
    - Text of the Link: “linkText”
    - Address the URL points to: “address”
- Music: “music”
- Movies: “movies”
- TV Shows: “tvShows”
- Books: “books”
- Activities: “activities”
- Sports: “sports”
- Heroes: “heroes”
- Favorite Quotes: “quotes”
- Favorite Cars: “cars”
- Favorite Food: “food”
- Turn Ons: “turnOns”
- Turn Offs: “turnOffs”
- Any user Tags: “tags”
- Comments about Romance: “romance”
- Living Arrangements: “livingArrangement”
- Profile Song: “profileSong”
- Profile Video: “profileVideo”
- Scared Of: “scaredOf”
- Happiest When: “happiestWhen”
- Fashion Thoughts: “fashion”
- Humor Thoughts: “humor”
- Looking For: “lookingFor”
- Religion or Religious Views: “religion”
- Political Views: “politicalViews”

Organization:

- Name of Organization: “name”
- Title of user in organization: “title”
- Description of user’s work: “description”
- Field of Organization: “field”
- Sub-field of Organization: “subField”
- Start Date for user: “startDate”

- End Date for user: “endDate”
- User Salary: “salary”
- Address of Organization: “address”
- Webpage of Organization: “webpage”

Enum

- Yes: “yes”
- No: “no”
- Socially: “socially”
- Occasionally: “occasionally”
- Regularly: “regularly”
- Heavily: “heavily”
- Quitting: “quitting”
- Quit: “quit”

Appendix B Description of Amazon & eBay APIs

Amazon Product Advertising API

http://docs.amazonwebservices.com/AWSECommerceService/2009-10-01/DG/

The Product Advertising API includes the following response groups:

- *Accessories:* It returns up to five ASINs and titles of accessories associated with items in the response.

- *AlternateVersions:* It returns all of the available media formats for a book title. Sample formats include Paperback, Audio CD, Audio Cassette, and Hardcover.

- *BrowseNodeInfo:* For a given browse node ID, the BrowseNodeInfo response group returns the browse node name and ID of the child and parent browse nodes.

- *BrowseNodes:* It returns the browse node names and IDs associated with the items returned in the response.

- *Cart:* It provides information about a specified remote shopping cart and the items in it.

- *CartNewReleases:* It returns the ASINs and titles of the top five new releases in the root category of the item specified in the cart operation. For example, when adding a television to a cart the top five new releases in the root category, electronics, are returned.

- *CartTopSellers:* It returns the ASINs and titles of the top five, best sellers in the root category of the item specified in the cart operation. For example, when

adding a television to a cart, the five top sellers in the root category, electronics, are returned, for example, the top selling computers, MP3 players, or digital cameras.

- *CartSimilarities:* Returns the title and ASINs of items that are similar to the item specified in the request, have been viewed by customers who also viewed the item specified in the request, can be found in other categories that are similar to the item specified in the request.

- *Collections:* For every item returned in a response, the items associated with it are also returned if the Collections response group is used in the request. Items in collections are related thematically. For example, all of the linens that go into a bedroom might be associated in a bedding collection. The Collections response group returns the ASINs and titles of the items in a collection.

- *CustomerFull:* It returns all of the content created by a customer including: id, name, nickname, location (city and state), WishList IDs, reviews, AboutMe message.

- *CustomerInfo:* For each customer in the response, the CustomerInfo response group returns the customer's Nickname and CustomerId. This response group will only return information that customers have chosen to make public through www.amazon.com.

- *CustomerLists:* It returns the WishList IDs created by each customer in the response. The WishList IDs are returned only if the customer has choosen to make the WishList public.

- *CustomerReviews:* For each customer in the response, the CustomerReviews response group returns his/her reviews/ratings.

- *EditorialReview:* For each item in the response, the EditorialReview response group returns Amazon's review of the item.

- *Fitments:* This response group returns the vehicles a specified part works in.

- *HasPartCompatibility:* It returns, for each vehicle part, a HasPartCompatibility element with a value of 1 (has) or 0 (doesn't have) depending whether the part has a compatibility chart.

- *Help:* It returns information about operations and response groups.

- *Images:* The Images response group returns the URLs to all available images of an item in three sizes: small, medium, and large.

- *ItemAttributes:* It returns a potentially large number of attributes that describe an item. All search indices can return all item attributes. The number of item attributes returned, however, varies by ASIN.

- *ItemIds:* The ItemIds response group returns the ASINs for all items returned in a response.

- *Large:* It returns a great deal of information about items in the response.

- *ListFull:* It provides comprehensive information about a list and the items on it.

- *ListInfo:* It provides descriptive information about a list.

- *ListItems:* It describes the items on a list.

- *ListmaniaLists:* It returns the Listmania list IDs and names of those lists that items, in the response, belong to.

- *ListMinimum:* The ListMinimum response group returns, for each list in the response, the list's ID, list name, number of items on the list, the number of pages of list items. There are ten list items per page.

- *Medium:* It returns a great deal of information about the items in a response. The response group is ideally suited for creating lightweight, product detail pages.

- *MerchantItemAttributes:* It returns merchant-specific information about a merchant's items for sale.

- *MostGifted:* It returns the ASINs and titles of the ten items given as gifts most within a specified browse node.

- *MostWishedFor:* It returns the ASINs and titles of the ten items given as the items listed on the greatest number of wishlists within a specified browse node.

- *NewReleases:* It returns the ASIN and title of newly released items in a specified browse node (category or products).

- *OfferFull:* It returns comprehensive information about an offer.

- *OfferListings:* It returns the offer listings for items returned in the response.

- *Offers:* The Offers response group is a parent response group that returns the contents of the OfferSummary response group plus, by default, seller and offer listing information.

- *OfferSummary:* It returns for each item in the response, the number of offer listings and the lowest price for each condition type.

- *PartBrandBinsSummary:* This response group returns a list of brands that satisfy the specified year, make, model, trim, and vehicle options.

- *PartBrowseNodeBinsSummary:* It returns a list of child nodes of the top Automotive browsenode (or the `BrowseNodeId` if supplied in the request) with the name, ID, and a total count of parts in that node that would work in the specified year, make, model, trim, and vehicle options.

- *PromotionDetails:* It returns detailed information about promotions (if any) related to items in a response. Included is information about the merchant offering the promotion, claim codes for the promotion, allowed promotion combinations, the type of promotion, beginning and ending dates of the promotion, the promotion ID, eligibility requirements, and text that describes the specifics of the promotion.

- *PromotionSummary:* It returns summary information about a promotion, including the type of promotion, beginning and ending dates of the promotion, the promotion ID, eligibility requirements, and text that describes the specifics of the promotion.

- *RelatedItems:* It returns items related to an item specified in an ItemLookup request.

- *Request:* It returns all of the parameters and their values that were submitted in a request. Use this information to debug requests.

- *Reviews:* It returns for each item in the response a list of customer's reviews, average review rating and total number of reviews.

- *SalesRank:* It returns the sales rank for each item in the response. One is the highest rating; a large number means the item has not sold well.

- *SearchBins:* The SearchBins response group groups the items returned by ItemSearch into bins. A set of bins, for example, can be a set of price ranges for a product.

- *Seller:* It returns information about sellers, including the seller ID, nickname, seller rating, and location for each seller in the response.

- *SellerListing:* The SellerListing response group returns information about items for sale by sellers in the Marketplace.

- *Similarities:* It returns titles and ASINs of items that are similar to the one specified in the request.

- *Small:* It returns basic information about items in a response.

- *Subjects:* It returns a book's Subject description, which characterizes the book's content.

- *TaggedGuides:* It returns all guides labeled by a specified tag.

- *TaggedItems:* It returns information about all items labeled by a specified tag.

- *TaggedListmaniaLists:* It returns all Listmania lists labeled by a specified tag.

- *Tags:* It returns complete information about tags associated with specified items.

- *TagsSummary:* It returns the items tagged by a specified tag.

- *TopSellers:* It returns the ASINs and titles of the ten best sellers within a specified browse node.

- *Tracks:* It returns the title and number of each track on each CD in the response.

- *TransactionDetails:* It returns information about customer transactions.

- *VariationMinimum:* It returns all of the children ASINs of each parent ASIN in the response. For example, if the response contains the parent ASIN, Brand X T-Shirt, this response group will return the child ASINs for Brand X T-Shirt, including: Small – Blue - Brand X T-Shirt, Medium – Blue – Brand X T-Shirt, etc.

- *Variations:* It is a parent response group that returns the contents of the VariationSummary and VariationMinimum response groups plus other variation details, such as item attributes, offers, and offer listings for each variation in the response.

- *VariationImages:* It displays different image variations of the same item in four sizes: swatch, small, medium, and large, where the swatch image is smaller than the small image.

- *VariationMatrix:* It returns, for a given parent ASIN, the variation dimension name and value of each child ASIN.

- *VariationOffers:* It enables you to retrieve the offers for the children of a parent ASIN.

- *VariationSummary:* It provides the lowest price, highest price, lowest sale price, and highest sale price for all child ASINs in a response.

- *VehicleMakes:* This response group returns, for a given year, all of the makes of vehicles manufactured. If you use this response group, you must specify the `Year`.

- *VehicleModels:* It returns, for a given year and make, all of the models of the vehicles manufactured. If you use this response group, you must specify the `Year` and `MakeId`.

- *VehicleOptions:* It returns, for a given year, make, model, and trim a list of available options.

- *VehiclePartFit:* This response group returns, for a given year, make, model, and part (`ItemId`), a boolean value, YES or NO, which specifies whether the specified part works with the specified car.

- *VehicleParts:* It returns, for a given year, make, model, and trim, up to fifteen parts that would fit the vehicle. The number of parts returned is specified by the `Count` parameter in the `VehiclePartSearch` request.

- *VehicleTrims:* This group returns, for a given year, make ID, and model ID, a list of available trims. A trim is a package of vehicle options, such as power steering and power seat.

- *VehicleYears:* This response group returns all of the years a car was made.

eBay Shopping API:

http://developer.ebay.com/DevZone/shopping/docs/CallRef/index.html

The Shopping API includes the following response groups:

- **Item Search**
  - *FindItems:* Searches for items based on a query or a seller ID. If you use keywords, this call returns items that contain the keywords in the title. A maximum of 50 items is returned.
  - *FindItemsAdvanced:* Advanced search for items on eBay via keyword, ProductID, SellerID and several search filters.
  - *FindProducts:* Searches for stock product information (stock description and Item Specifics), such as information about a particular kind of DVD or camera. Also, retrieves up to 200 eBay listings associated with a product.
  - *FindHalfProducts:* Searches Half.com for stock product information (stock description and Item Specifics), such as information about a particular kind of DVD or book. Also, retrieves up to 30 Half.com listings associated with a product.
- **Item Data**
  - *GetSingleItem:* Gets publicly visible details about one listing. This gives you most of the data that eBay shows on the View Item page (title, description, price, and other details).
  - *GetItemStatus:* Allows you to get the status for a group of items. Returns status information such as ListingStatus and End Time for all items that are listed in the request.
  - *GetShippingCosts:* Gets shipping costs for an item.
  - *GetMultipleItems:* Retrieves publicly available data for one or more listings.
- **Category Information**
  - *GetCategoryInfo:* This call will give you the ability to retrieve high level Category information, relevant for a buy-side application.
- **User Reputation**

    - *GetUserProfile:* Retrieves user information based on the user ID you specify. The response contains detailed information about a user. You can specify the types of user information you want in the response.

- **eBay Pop**
    - *FindPopularSearches:* Finds the words more frequently used by eBay users when searching for listings. If you use keywords, this call returns available alternative keywords in addition to popular related keywords.

    - *FindPopularItems:* Searches for popular items based on a category or keyword. Returns WatchCount in addition to item information.

- **Search**
    - *FindReviewsandGuides:* Searches reviews and guides based on product, category, or user. The response provides information about each user or product's reviews and guides.

- **eBay Time**
    - *GeteBayTime:* Gets the official eBay system time in GMT.

eBay Trading API:

http://developer.ebay.com/DevZone/XML/docs/Reference/eBay/index.html

The Trading API includes the following response groups:

- *AddDispute*: Enables a seller to create a new Unpaid Item dispute. (Item Not Received disputes can only be created via the eBay web site.)

- *AddDisputeReponse:* Adds a response or comment to a dispute, or closes a dispute.

- *AddItem:* Sends data defining a new item (specified by a seller) to eBay, where it becomes a new listing.

- *AddLiveAuctioItem:* Available to eBay Live Auctions sellers. Sends data defining a single new lot item to the eBay Live Auctions site, where it becomes a new lot in a seller's Live Auction catalog. The listing also appears on the main eBay site.

- *AddMemberMessageAAQToPartner:* Enables a buyer and seller in a transactional relationship to send messages to each other's My Messages Inboxes within 90 days of the creation of the transaction.

- *AddMemberMessageRTO:* Enables a seller to reply to a question about an active item listing. The reply is sent to the user's My Messages inbox.

- *AddMemberMessagesAAQToBidder:* Enables a seller to send up to 10 messages to bidders and users who have made offers (via Best Offer) during an active listing. Messages to a user appear in the user's My Messages inbox.

- *AddOrder:* Combines two or more transactions into a single order, enabling a buyer to pay for all of those transactions with a single payment (and, if so arranged, ship all of the items together).

- *AddSecondChanceItem:* Creates a new Second Chance Offer (that is, an offer for an unsold item) for one of that item's non-winning bidders.

- *AddToItemDescription:* Appends a horizontal rule, then a message about what time the addition was made by the seller, and then the seller-specified text.

- *AddToWatchList:* Adds one or more items to the user's My eBay watch list.

- *AddTransactionConfirmationItem:* Ends the listing specified by ItemID (if listed for at least 24 hours) and creates a new Transaction Confirmation Request (TCR) for an item, enabling the TCR recipient to purchase the item.

- *ApproveLiveAuctionBidders:* Provides Live Auction sellers with the ability to approve, decline, and set the bidding limit of the bidders that have signed up for a catalog.

- *CompleteSale:* Enables a seller to do various tasks after a transaction has been created. A seller can leave feedback for the buyer, change the paid status, or set shipment tracking information (or any combination of these).

- *DeleteMyMessages:* Removes selected alerts and messages for a given user.

- *EndItem:* Ends the specified item listing before the date and time at which it would normally end (per the listing duration).

- *FetchToken:* Retrieves a user token. Also can be used to retrieve a REST token.

- *GetAccount:* Enables a seller to retrieve his or her own account data.

- *GetAdFormatLeads:* Retrieves sales lead information for a lead generation listing.

- *GetAllBidders:* Provides three modes for retrieving a list of the users that bid on a listing.

- *GetApiAccessRules:* Reports how many calls your application has made and is allowed to make per hour or day.

- *GetAttributesCS:* Retrieves an XML string that describes how to present Item Specifics to a seller who is creating a new listing.

- *GetAttributesXSL:* Retrieves the Item Specifics SYI XSL stylesheet for use with the GetAttributesCS and GetProductSellingPages response. You use the stylesheet to render Item Specifics in a user interface, as applicable within a particular category.

- *GetBestOffers:* Retrieves the best offers associated with an ItemID according to the BestOfferStatus filter, where Active is the default value. Specify a best offer ID to retrieve the details for a specific best offer.

- *GetBidderList:* Retrieves all items on which the user is currently bidding or which the buyer has won or purchased.

- *GetCart:* Retrieves information about an eBay Express shopping cart.

- *GetCategories:* Retrieves the latest eBay category hierarchy for a given eBay site. Information returned for each category includes the category name and the unique ID for the category.

- *GetCategory2CS:* Retrieves mappings between categories and characteristic sets that are available for an eBay site.

- *GetCategoryFeatures:* Returns information about certain features that may only be applicable to certain categories on the site, such as particular listing durations, shipping term requirements, and Best Offer support.

- *GetCategoryListings:* Returns items in a specified category. A number of inputs are provided for filtering the item listings returned using such criteria as the listing type and whether the item is listed in an eBay Store.

- *GetCategoryMappings:* Retrieves a map of old category IDs and corresponding active category IDs defined for the site to which the request is sent.

- *GetCategorySpecifics:* Retrieves the most popular custom Item Specifics that sellers use when they list items in certain categories. Item Specifics are optional in listings.

- *GetChallengeToken:* Retrieves a botblock token and URLs for an image or audio clip that the user is to match.

- *GetCharities:* Searches for non-profit charity organizations that meet the criteria specified in the request.

- *GetContextualKeywords:* Retrieves top-ranked contextual eBay keywords and categories for a specified web page.

- *GetCrossPromotions:* Retrieves a list of upsell or cross-sell items associated with the specified item ID.

- *GetDescriptionTemplates:* Retrieves the DescriptionTemplates for a category.

- *GetDispute:* Requests the details of a dispute corresponding to the given dispute ID, any time after the dispute was opened and up to five years after it was closed.

- *GeteBayDetails:* Retrieves eBay IDs and codes (e.g., site IDs and shipping service codes), enumerated data (e.g., payment methods), and other common eBay meta-data.

- *GeteBayOfficialTime:* Gets the official eBay system time in GMT.

- *GetExpressWishList:* Retrieves eBay Express user wish lists.

- *GetFeedback:* Retrieves the accumulation of feedback left for the specified user by other users.

- *GetHighBidders:* Retrieves a list of high bidders for the Dutch auction specified in the ItemId property of the request. A seller can use this list to determine which buyers are winning bidders and how many items each can purchase.

- *GetItem:* Requests data for a specific item identified by item ID. Data returned includes title, description, minimum bid price, seller information, high bidder information (if there currently is a high bidder), and shipping specifications (if the seller elected to have the buyer pay for shipping).

- *GetItemRecommendations:* Examines potential item data that a seller has specified and returns recommended changes or opportunities for improvement.

- *GetItemsAwatingFeedback:* Returns items for which feedback needs to be left.

- *GetItemShipping:* Returns shipping cost estimates for an item for every calculated shipping service that the seller has offered with the listing. This is analogous to the Shipping Calculator seen in both the buyer and seller web pages.

- *GetItemTransactions:* Retrieves transaction information for a specified itemID.

- *GetLiveAuctionBidders:* Gets a list of users who have signed up to participate in their auctions, including the bid approval status of each user.

- *GetLiveAuctionCatalogDetails:* Retrieves all the eBay Live Auctions catalogs and upcoming sale schedules that the user has created.

- *GetMemberMessages:* Retrieves messages posted to the 'Ask Seller A Question' messaging system.

- *GetMessagePreferences:* Returns a seller's Ask Seller a Question (ASQ) subjects.

- *GetMyeBayBuying:* Returns items from the All Buying section of the user's eBay account, including items the user is watching, bidding on, has won, has not won, or has made best offers on.

- *GetMyeBayReminders:* Requests totals of various reminder types from the user's eBay account.

- *GetMyeBaySelling:* Returns a summary and details of items a user is selling from the user's eBay account.

- *GetMyMessages:* Retrieves information about the messages and alerts sent to a given user.

- *GetNotificationPreferences:* Retrieves the requesting application's notification preferences.

- *GetNotificationUsage:* Retrieves usage information about platform notifications for a given application.

- *GetOrders:* Retrieves all of the orders for which the user is a participant (as either buyer or seller) that meet the criteria specified in the request.

- *GetOrderTransactions:* Retrieves information about one or more orders or one or more transactions (or both).

- *GetPictureManagerDetails:* Requests information about folders or pictures in a Picture Manager account or the account settings.

- *GetPictureManagerOptions:* Requests a list of Picture Manager options and allowed values, such as subscription type and picture display.

- *GetPopularKeywords:* Retrieves the words more frequently used by eBay users when searching for listings.

- *GetProductFamilyMembers:* If the user wants to see more versions (editions) of the product, you can use GetProductFamilyMembers to retrieve all versions of the product.

- *GetProductFinder:* Retrieves data that you use to construct valid "product finder" queries (queries against multiple attributes) against catalog products or listed items.

- *GetProductFinderXSL:* Retrieves the Product Finder XSL stylesheet for use with the XML returned from GetProductFinder.

- *GetProducts:* Searches for stock product information (Pre-filled Item Information), such as information about a particular DVD or camera. Also, retrieves top reviews, buying guides, and up to 200 listings associated with a product. Primarily useful for buying use cases.

- *GetProductSearchPage:* Retrieves the attributes a seller can use to form a query when searching for Pre-filled Item Information to use in a listing for a category that is catalog-enabled.

- *GetProductSerchResults:* Searches for Pre-filled Item Information within one or more characteristic sets.

- *GetProductSellingPages:* Retrieves information that describes how to present catalog product information to a seller.

- *GetPromotionalSaleDetails:* Used to obtain information about promotional sales

- *GetPromotionRules:* Requests a list of the cross-promotion rules associated with a specific referring item or store category.

- *GetReturnURL:* The URLs that eBay should redirect users to after they complete the authentication and authorization process.

- *GetRuName:* Returns a globally unique runame (unique identifier for an authentication data entry).

- *GetSearchResults:* Retrieves item listings based on keywords you specify.

- *GetSearchResultsExpress:* Retrieves brief details about active items that are listed on eBay Express.

- *GetSellerEvents:* Retrieves price changes, item revisions, description revisions, and other changes that have occurred within the last 48 hours related to a seller's eBay listings.

- *GetSellerList:* Returns a list of items a seller has listed for auction.

- *GteSellerPayments:* Retrieves a summary of pending or paid payments that Half.com created for the seller identified by the authentication token in the request.

- *GetSellerTransactions:* Retrieves transaction information for the user for which the call is made (and not for any other user), where a transaction is the information about the sale of one or more items by one buyer from a single listing.

- *GetSessionID:* Retrieves a session ID that identifies a user and your application.

- *GetShippingDiscountProfiles:* Returns details of the shipping discount profiles defined by the user, along with other combined payment-related details such as packaging/handling cost and shipping insurance.

- *GetStore:* Retrieves configuration information for the eBay store owned by the specified user or caller.

- *GetStoreCategoryUpdateStatus:* Returns the status of the processing progress for category structure changes specified with a SetStoreCategories request.

- *GetStoreCustomPage:* Retrieves the custom page or pages for the user's Store.

- *GetStoreOptions:* Retrieves the current list of Store configuration options.

- *GetStorePreferences:* Retrieves a user's Store preferences.

- *GetSuggestedCategories:* Requests a list of up to 10 categories with the highest percentage of listings whose titles or descriptions contain the keywords you specify.

- *GetTaxTable:* Retrieves the tax table for a user on a given site or retrieves the valid jurisdictions (if any) for a given site.

- *GetTokenStatus:* Requests current status of user token.

- *GetUser:* Retrieves data for one eBay user. Input fields control what user data is returned.

- *GetUserContactDetails*: Returns contact information for a specified user if a bidding relationship exists. Both sellers and bidders can use this call.

- *GetUserDisputes:* Requests a list of disputes the requester is involved in as buyer or seller. The list includes both Unpaid Item and Item Not Received disputes.

- *GetUserPreferences:* Retrieves some or all of a user's preferences.

- *GetVeROReasonCodeDetails:* Retrieves details about VeRO reason codes for a given site or all sites. You must be a member of the Verified Rights Owner (VeRO) Program to use this call.

- *GetVeROReportStatus:* Retrieves status information about VeRO reported items. You must be a member of the Verified Rights Owner (VeRO) Program to use this call.

- *GetWantItNowPost:* Retrieves data for a specific, active Want It Now post identified by a post ID.

- *GetWantItNowSearchResults:* Retrieves a list of active Want It Now posts that match specified keywords and/or a specific category ID.

- *IssueRefund:* Issues a refund for a single Half.com transaction.

- *LeaveFeedback:* Enables a user to leave feedback about another user.

- *PlaceOffer:* Enables the requesting user to submit a bid (make an offer) on the item listing specified in ItemID.

- *RelistItem:* Enables a seller to take a single item (or a single multi-item listing) that ended without bidders or buyers and re-lists it on a specified eBay site.

- *RemoveFromWatchList:* Enables a user to remove item from the user's eBay watch list.

- *RespondToBestOffer:* Enables the seller of a Best Offer item to accept, decline, or counter offers made by bidders.

- *RespondToFeedback:* Enables users to reply to feedback left by the other party in the transaction and/or follow up on the feedback left for the other party.

- *RespondToWantItNowPost:* Enables a seller to respond to a Want It Now post with an item listed on the eBay site.

- *ReviseCheckoutStatus:* Enables a seller to update the payment status of an item.

- *ReviseItem:* Enables a seller to change the properties of a currently active listing.

- *ReviseLiveAuctionItem:* Enables an eBay Live Auctions seller to modify a single lot item on the eBay Live Auctions site.

- *ReviseMyMessages:* Sets the read state for messages and alerts, the flagged state of messages, and moves alerts and messages into and out of folders.

- *ReviseMyMessagesFolders:* Renames, removes, or restores specified My Messages folders for a given user.

- *RevokeToken:* Voluntarily revokes a token before it would otherwise expire, typically because of security concerns or because the user has unsubscribed.

- *SellerReverseDispute:* Enables a seller to "reverse" an Unpaid Item dispute that has been closed, for example, if buyer and seller reach an agreement. The seller's Final Value Fee credit and the buyer's strike are both reversed. If applicable.

- *SendInvoice:* Enables a seller to send an invoice to a buyer involved in the transaction.

- *SetCart:* Enables a third party developer to create a new shopping cart in the eBay Express domain or to modify the contents of an existing shopping cart by adding or removing items or changing item quantity.

- *SetMessagePreferences:* Enables a seller to add custom Ask Seller a Question (ASQ) subjects to display on the seller's Ask a Question page.

- *SetNotificationPreferences:* Manages a user's notification preferences.

- *SetPictureManagerDetails:* Creates, updates, or deletes Picture Manager account settings, folders, or pictures.

- *SetPromotionalSale:* Creates or modifies a promotional sale. Promotional sales enable sellers to apply discounts across many listings.

- *SetPromotionalSaleListings:* Used to change which item listings are affected by a promotional sale.

- *SetReturnURL:* Configures your application's authentication and authorization preferences and other data, such as the URLs that eBay should redirect users to after they complete the authentication and authorization process.

- *SetShippingDiscountProfiles:* Enables a seller to define shipping cost discount profiles and a few additional parameters related to combined payment, such as shipping insurance and packaging handling costs.

- *SetStore:* Sets the configuration of the eBay store owned by the caller.

- *SetStoreCategories:* Requests changes to the category structure for a store.

- *SetStoreCustomPage:* Sets a custom page for a user's eBay Store.

- *SetStorePreferences:* Sets the preferences for a user's eBay Store.

- *SetTaxTable:* Sets the tax table for a seller on a given site.

- *SetUserNotes:* Enables users to add, replace, and delete notes for items that are being tracked in the My eBay All Selling and All Buying areas.

- *SetUserPreferences:* Sets the user's preferences to those specified in the request.

- *UploadSiteHostedPictures:* Uploads a picture for a listing and returns a URL for the picture.

- *ValidateChallengeInput:* Validates the user response to a GetChallengeToken botblock challenge.

- *ValidateTestUserRegistration:* Requests to enable a test user to sell items in the Sandbox environment.

- *VerifyAddItem:* Enables a seller to specify the definition of a new item and submit the definition to eBay without creating a listing.

- *VerifyAddSecondChanceItem:* Simulates the creation of a new Second Chance Offer listing of an item without actually creating a listing.

- *VeROReportItems:* Reports items that allegedly infringe your copyright, trademark, or other intellectual property rights. You must be a member of the Verified Rights Owner (VeRO) Program to use this call.

eBay Research API:

http://developer.dataunison.com/pages/developers_area/ebay_research_api/api_call_reference.html

The ResearchAPI includes the following response groups:

- *GetCategoryHotList:* It calculates eBay categories that are doing well based on their bids to listings.

- *GetCategoryItems:* It retrieves a list of all items from a given category allowing you to browse the category.

- *GetCategoryTrends:* It provides trend data over an entire category for specific time spans.

- *GetPriceResearch:* This API is a free call that quickly returns basic pricing statistics based on a keyword search.

- *GetResearchItems:* It can be used to retrieve the raw items associated with a normal API search. It will return up to 200 items at a time, and can be provided with an offset to return entire result sets with multiple calls.

- *GetResearchResults:* It's used to retrieve basic pre-calculated statistics for a specific set of keywords, along with other filters.

- *GetResearchTrendData:* It is used to retrieve trends based on the data similar to the ResearchResults call.

- *GetSellerResearchResults:* It can be used to retrieve research statistics about a specific user, or without a specific user, on eBay. It takes in eBay seller IDs as well as all the other filters found in the GetResearchResults call, and returns the same results as that call.

- *GetSellerResearchTrendData:* It s used to retrieve trends based on the data similar to the SellerResearchResults call.

- *GetSellerTopTitles:* It retrieves the top titles of a given seller.

- *GetTitleBuilderResults:* It returns hot keywords to use based on both closed and live listings.

- *GetUsageStats:* It can be used by clients to retrieve call volume statistics for their account.

Appendix C Properties of SUMI Ontology

List of DatatypeProperties in SUMI ontology

Name	Domain	Range
hasUserName	User	String
hasUserURI	User	String
hasAttributeName	Attribute	String
hasAttributeURI	Attribute	String
hasAttributeType	Attribute	String
hasAttributeDescription	Attribute	String
hasAttributeValue	Attribute	String
hasProviderDescription	Provider	String
hasProviderName	Provider	String
hasProviderURI	Provider	String
hasCategoryDescription	AttributeCategory	String
hasCategoryName	AttributeCategory	String
hasCategoryURI	AttributeCategory	String
dictionaryConcept	DictionaryConcept	String

Table C-1 DatatypeProperties in SUMI Ontology

List of ObjectProperties in SUMI ontology

Name	Domain	Range
userIsPerson	User	foaf:Person
interactsWithSocialNetworkingProvider	User	SocialE-Networking
interactsWithBusinessProvider	User	E-Commerce
attributeForUser	Attribute	User
sameAsDictionaryConcept	Attribute	DictionaryConcept
attributeForProvider	Attribute	Provider
hasGenericUserData	Provider	GenericUserData
hasServiceSpecificUserData	Provider	ServiceSpecificUserData
hasServiceGenericGeneratedInfo	Provider	ServiceGenericGeneratedInfo
hasServiceSpecificGeneratedInfo	Provider	ServiceSpecificGeneratedInfo
belongsToGenericUserData	CommonUserInput	GenericUserData
belongsToServiceSpecificUserData	SpecificUserInput	ServiceSpecificUserData
belongsToServiceGenericGeneratedInfo	CommonGeneratedInformation	ServiceGenericGeneratedInfo
belongsToServiceSpecificGeneratedInfo	SpecificGeneratedInformation	ServiceSpecificGeneratedInfo
hasGUMORelation	CommonUserInput, SpecificUserInput, CommonGeneratedInformation, SpecificGeneratedInformation	gumo:SituationalElements
hasCommonUserInputAttr	GenericUserData	CommonUserInput

ibutes		
hasSpecificUserInputAttributes	ServiceSpecificUserData	SpecificUserInput
hasCommonGeneratedAttributes	ServiceGenericGeneratedInfo	CommonGeneratedInformation
hasSpecificGeneratedAttributes	ServiceSpecificGeneratedInfo	SpecificGeneratedInformation

Table C-2 ObjectProperties in SUMI Ontology

Appendix D User Evaluation Structures

FIRST USER EVALUATION STRUCTURE

Scrutability User Privileges

Multiple-Choice Evaluation Questions and List of Available Answers

Scrutability User Privilege 1	Competence Question	Acceptance Question	Consequence Question	Between task Question
Task 1: Adding models to the SUMI collection of models.	Do you think you have managed to complete the task and add at least one model to your SUMI collection?	What do you think about this proposed user privilege - Adding models to your SUMI collection?	Do you understand the consequences of adding your models in SUMI?	Which category of the social networking provider you have just added contains the "Favourite Movies" section which you can find in both Facebook and MySpace?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I can now access them via SUMI	Generic User Data

			and I am in control	
	No	I like it but I do NOT want to have it available for use	I can access them via SUMI but SUMI is in control of my added models	Service Specific User Data
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	Service Generic Generated Info
		I do NOT like it and do NOT want to have it available for use		Service Specific Generated Info

Table D-1 Evaluation 1 - Scrutability User Privilege 1

Scrutability User Privilege 2	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 2: Importing content for all 4 categories of all previously added models both dynamically and statically.	Do you think you have managed to complete the task and import the content of one category of at least one of your models, both dynamically and statically?	What do you think about this proposed user privilege - Importing the content of your previously added models?	Do you understand the consequences of importing the content of your models in SUMI?	1) In which colour is the actual content of your pre-selected category/model presented? 2) Which of the following statements is wrong?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**

	Yes	I like it and I want to have it available for use	I can select which category of which of my models to import its content and I am in control in which way the content is being imported (either dynamically or statically)	1) Red 2) A dynamic import will always return the actual content of the selected category and model
	No	I like it but I do NOT want to have it available for use	I can access them via SUMI but I am not in control	1) Brown 2) A static import will always return the actual content of the selected category and model
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	1)Blue 2) SUMI keeps a copy of my content ONLY when I initiate a dynamic import
		I do NOT like it and do NOT want to have it available for use		

Table D-2 Evaluation 1 - Scrutability User Privilege 2

Scrutability User Privilege 3	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 6: Export content of all 4 categories of all previously added models to subscribed	Do you think you have managed to complete the task and export the content of	What do you think about this proposed user privilege - Exporting the content of	Do you understand the consequences of exporting the content of	Which of the following statements is correct?

providers after inspecting and approving transaction details.	one of your models to a subscribed provider?	your models to subscribed providers?	your models to a subscribed provider?	
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I am allowing access to the provider for retrieving the content of a category of one of my models - I inspected and approved which attributes will go through with the transaction before it happened.	I have NOT inspected NOR approved the transaction details before exporting my content
	No	I like it but I do NOT want to have it available for use	I am allowing access to the provider for viewing the content of my models	I have inspected but NOT approved the transaction details before exporting my content
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	I have NOT inspected but approved the transaction details before exporting my content
		I do NOT like it and do NOT want to have it available for use		I have inspected and approved all the transaction details before exporting my content

Table D-3 Evaluation 1 - Scrutability User Privilege 3

Privacy User Privileges

Multiple-Choice Evaluation Questions and List of Available Answers

Privacy User Privilege 1	Competence Question	Acceptance Question	Consequence Question	Between task Question
Task 3: Setting the privacy status for all 4 categories of all previously added models.	Do you think you have managed to complete the task and set the privacy status for all 4 categories of the first model of your SUMI collection?	What do you think about this proposed user privilege - Setting the privacy status of the categories of your models?	Do you understand the consequences of setting the privacy status of the categories of your models?	Which category of your first model will now other users be able to see but not access its content?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I can determine who gets to see the content of each of my models' categories.	Generic User Data
	No	I like it but I do NOT want to have it available for use	I can leave anyone to do whatever they want with my models	Service Specific User Data
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	Service Generic Generated Info
		I do NOT like it and do NOT want to have it available for use		Service Specific Generated Info

Table D-4 Evaluation 1 - Privacy User Privilege 1

Privacy User Privilege 2	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 4: Responding to viewing/accessing requests from other SUMI users to all 'private' categories of all previously added models.	Do you think you have managed to complete the task and reply to one request by another user?	What do you think about this proposed user privilege - Responding to other users' requests for your private categories of your models?	Do you understand the consequences of replying to other users' requests for the private categories of your models?	N/A
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I am allowing access to other users for viewing the content of the private categories of my models	N/A
	No	I like it but I do NOT want to have it available for use	I am allowing access to other users to do anything they want with my models	
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	
		I do NOT like it and do NOT want to have it available for use		

Table D-5 Evaluation 1 - Privacy User Privilege 2

Privacy User Privilege 3	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 5: Visiting other users' requests and placing viewing/accessing requests to all 'private' categories of those users' previously added models	Do you think you have managed to complete the task and place a viewing request while visiting Maria's SUMI collection of models?	What do you think about this proposed user privilege - Visiting other users' SUMI collection of models and placing viewing requests?	Do you understand the consequences of visiting other users' SUMI models and placing viewing requests on private categories of that user's models?	Which is one of Maria's favourite books?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I am asking from another user access to the requested private categories of his/her models	The Da Vinci Code
	No	I like it but I do NOT want to have it available for use	I am asking from another user to give me access to the content of all of his/her models	Angels & Demons
	I don't know	I do NOT like it but I want to have it available for use	I do not understand the consequences	Alchemist
		I do NOT like it and do NOT want to have it available for use		

Table D-6 Evaluation 1 - Privacy User Privilege 3

Pre-Questionnaire

Multiple-Choice Questions and List of Available Answers

Questions	Answers
Which social networking websites do you *mostly* use (Facebook, MySpace, Hi5, Twitter, etc...)?	An empty text box is provided
Which e-business website do you *mostly* use (Amazon, eBay, Barnes & Nobles, etc...)?	An empty text box is provided
Are you familiar with the term scrutability?Do you think the various providers (e.g. Facebook, Amazon, etc.) offer you some scrutability user privileges?	I am familiar with the term and I recognize some scrutability user privileges when I interact with various providers
	I am NOT familiar with the term but I do recognize some scrutability user privileges when I interact with various providers
	I am familiar with the term but I do NOT recognize any scrutability user privileges when I interact with various providers
	I am NOT familiar with the term NOR I recognize any scrutability user privileges when I interact with various providers
Would you like to have the option of been allowed to inspect and alter, in a variety of ways, the way you are been modeled by various providers (such as MySpace and eBay)?	Yes, that would be great
	Sounds good
	I don't really mind
	No
	I don't understand the question
Are you familiar with your privacy options when you interact with various providers (e.g. MySpace, eBay, etc.)? Do you take advantage of the offered privacy options?	I am familiar with my privacy options and I do take advantage of them
	I am familiar with my privacy options but I do NOT take advantage of them
	I am NOT familiar with my privacy options but I do take advantage of them
	I am NOT familiar with my privacy options NOR I take advantage of them
Would you mind if the various providers	Yes, I would mind a lot

which you interact with (such as MySpace and eBay) do not allow you to set the privacy status of your information which you have personally entered in a previously stage?	It would bother me
	I don't really mind
	No
	I don't understand the question
Finally, please let us know your studies' major (e.g. Computer Science, Psychology, Music & Arts, ...)	An empty text is provided

Table D-7 Evaluation 1 - Pre-Questionnaire

Post-Questionnaire

Multiple-Choice Questions and List of Available Answers

Questions	Answers
Would you use such a service? Do you find it useful?	Yes I would use it and I find it very useful
	Yes I would I use it but I don't know if it's useful
	Neutral
	No I wouldn't use it but I find it useful
	No I wouldn't use it and I don't find it useful
What do you think about the fact of SUMI keeping a copy of your information?	I find it very useful since SUMI provides two options for importing my information (dynamically and statically)
	I am OK with it
	I don't like it but I can see the value in doing so
	I understand it but I don't agree
	I don't like it and I don't see the value in doing so
	I don't understand it
How much do you value scrutability now, after going through Tasks 1,2 and 6, when you interact with user modeling providers	Very much
	Some value

such as Facebook and Amazon?	Neutral
	No value
	I don't understand the question
What other scrutability privileges would you like SUMI to offer to its users?	An empty text box is provided
How much value privacy of information has for you now, after completing Tasks 3, 4 and 5, when you interact with user modeling providers such as Facebook and Amazon?	Very much
	Some value
	Neutral
	No value
	I don't understand the question
What other privacy privileges would you like SUMI to offer to its users?	An empty text box is provided

Table D-8 Evaluation 1 - Post-Questionnaire

SECOND USER EVALUATION STRUCTURE

Scrutability User Privileges

Multiple-Choice Evaluation Questions and List of Available Answers

Scrutability User Privilege 1	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 1: Adding models to the SUMI	Do you think you have managed to	What do you think about this proposed	Do you understand the	In which colour was the actual content of your profiles

collection of models & Importing real-time data of the added models	complete the task and add at least one profile to your SUMI collection while importing its content too?	scrutability user privilege - Adding your profiles in SUMI & Importing your profiles' content?	consequences of adding your profiles in SUMI?	presented?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I can now access them via SUMI and I am in control	Red
	No	I like it but I do NOT want to have it available for use	I can access them via SUMI but SUMI is in control of my added models	Brown
	I don't know	I do NOT like it but I want to have it available for use	I can access them via SUMI and both SUMI and myself have control over my added profiles	Blue
		I do NOT like it and do NOT want to have it available for use	I do not understand the consequence	Black

Table D-9 Evaluation 2 - Scrutability User Privilege 1

Scrutability User Privilege 2	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 2: Customising SUMI settings for the collection of models and instructing SUMI what is allowed to be imported from each model. In addition, deleting any values that SUMI holds about each user and the owner prefers not to be kept inside SUMI.	Do you think you have managed to complete the task and customise the SUMI settings for your collection of profiles?	What do you think about this proposed scrutability user privilege - Customising SUMI settings for your collection of profiles while inspecting and altering what SUMI holds about you?	Do you understand the consequences of customising the SUMI settings for your collection of profiles?	SUMI allows me to:
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I can determine how SUMI handles my previously added profiles and what SUMI is allowed to import and store for my profiles	Delete anything I do not want SUMI to hold about me
	No	I like it but I do NOT want to have it available for use	I can determine how SUMI handles my previously added profiles but NOT what SUMI is allowed to store for my profiles	Instruct SUMI not to import again any attribute of any of my profiles

	I don't know	I do NOT like it but I want to have it available for use	I can NOT determine how SUMI handles my previously added profiles NOR what SUMI is allowed to store for my profiles	Inspect what SUMI holds about each and every one of my profiles
		I do NOT like it and do NOT want to have it available for use	I do not understand the consequences	All of the above Only a and b are correct

Table D-10 Evaluation 2 - Scrutability User Privilege 2

Scrutability User Privilege 3	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 6: Export content of all 4 categories of all previously added models to a subscribed educational system after inspecting and approving transaction details. In addition, going through an example of how the exported	Do you think you have managed to complete the task and export the content of one of your profiles to a subscribed educational system?	What do you think about this proposed user privilege - Exporting the content of your profiles to subscribed educational systems?	Do you understand the consequences of exporting the content of your profiles to subscribed educational systems?	Which one of the following statements is WRONG?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I am allowing access to the educational system for retrieving the content of my	I have inspected the transaction details before exporting my information

information could be used by the educational system			profiles	
	No	I like it but I do NOT want to have it available for use	I am allowing access to the educational system for retrieving specific information from selected profile(s) - I inspected and approved the details of the transaction before allowing it to go through	I have approved the transaction details before exporting my information
	I don't know	I do NOT like it but I want to have it available for use	I am allowing access to the educational system for viewing the content of my profiles	I have NOT inspected NOR approved the transaction details before exporting my information
		I do NOT like it and do NOT want to have it available for use	I do not understand the consequences	By exporting my information to AHA, I have allowed this system to use my information and provide a personalised service to me

Table D-11 Evaluation 2 - Scrutability User Privilege 3

Privacy User Privileges

Multiple-Choice Evaluation Questions and List of Available Answers

Privacy	Competence	Acceptance	Consequence	Between task

User Privilege 1	Question	Question	Question	Question
Task 3: Setting the privacy status for all 4 categories of all previously added models. Users have the option of creating groups of users and assigning a common privacy setting to all group members as preferred	Do you think you have managed to complete the task and set the privacy status for all 4 categories of the social networking profile of your SUMI collection?	What do you think about this proposed privacy user privilege - Setting universal and customised privacy settings for your profiles?	Do you understand the consequences of setting the universal and customised privacy settings for you profiles?	Which one of the following statements is WRONG?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I can only determine who gets to see the content of the private categories of each one of my profiles	Customised privacy settings override universal privacy settings
	No	I like it but I do NOT want to have it available for use	I can only determine who gets to see the content of the hidden categories of each one of my profiles	If I set a category of one of my profiles as -hidden- then no user will be able to see the content of that category
	I don't know	I do NOT like it but I want to have it available for use	I can determine who gets to see the content of every category of each one of my profiles	Universal privacy settings are only applied when I set the privacy status of my profiles to -public-
		I do NOT like it and do NOT want	I do not understand the consequences	If I set a category of one of my profiles as

		to have it available for use		-private- then no user will be able to see the content of that category

Table D-12 Evaluation 2 - Privacy User Privilege 1

Privacy User Privilege 2	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 4: Responding to viewing requests sent by other SUMI users to all 'private' categories of all previously added models. In addition, inspecting any previously assigned privacy privileges and altering as preferred	Do you think you have managed to complete the task and reply to one request by another user while managing the assigned viewing privilege afterwards?	What do you think about this proposed user privilege - Responding to other users' requests for your private categories of your profiles and managing all previously assigned privileges?	Do you understand the consequences of replying to other users' requests for the private categories of your profiles?	Which one of the following statements is WRONG?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**
	Yes	I like it and I want to have it available for use	I am allowing access to other users for viewing the content of the private AND hidden categories of my profiles	In this task I have the option of managing the customised privacy settings I have previously assigned in Task 3
	No	I like it but I do NOT want to have it available for use	I am allowing access to other users for viewing the content of the private categories of my profiles	When I respond to a viewing request from another user is the same as setting a customised privacy setting for the same

				user back in Task 3
	I don't know	I do NOT like it but I want to have it available for use	I am allowing access to other users for viewing the content of all the categories of my profiles	If I set the status of all the categories of all of my models to -hidden- then no user will be able to send me any viewing requests
		I do NOT like it and do NOT want to have it available for use	I do not understand the consequences	If I respond to a request from another user, the status I decide to assign to that specific user is permanent and I can NOT change it afterwards

Table D-13 Evaluation 2 - Privacy User Privilege 2

Privacy User Privilege 3	**Competence Question**	**Acceptance Question**	**Consequence Question**	**Between task Question**
Task 5: Visiting other users' requests and placing viewing requests to all 'private' categories of those users' previously added models	Do you think you have managed to complete the task and place a viewing request while visiting Maria's SUMI collection of profiles?	What do you think about this proposed user privilege - Visiting other users' SUMI collection of profiles and placing viewing requests?	Do you understand the consequences of visiting other users' SUMI collection and placing viewing requests on private categories of that user's profiles?	Which of the following is NOT one of Maria's favourite books?
	Competence Answers	**Acceptance Answers**	**Consequence Answers**	**Between task Answers**

	Yes	I like it and I want to have it available for use	I am asking from another user access to the requested private categories of his/her profiles	Veronica Decides to Die
	No	I like it but I do NOT want to have it available for use	I am asking from another user to give me access to the content of all of his/her profiles	The Alchemist
	I don't know	I do NOT like it but I want to have it available for use	I am asking from another user access to the requested public categories of his/her profiles	The DaVinci Code
		I do NOT like it and do NOT want to have it available for use	I do not understand the consequences	Everything Happens for a Reason

Table D-14 Evaluation 2 - Privacy User Privilege 3

Pre-Questionnaire

Multiple-Choice Questions and List of Available Answers

Questions	Answers
Are you familiar with the term scrutability?Do you think the various providers (e.g. Facebook, Amazon, etc.) offer you some scrutability user privileges?	I am familiar with the term and I recognize some scrutability user privileges when I interact with various providers
	I am NOT familiar with the term but I do recognize some scrutability user privileges when I interact with various providers
	I am familiar with the term but I do NOT recognize any scrutability user privileges when I interact with various providers
	I am NOT familiar with the term NOR I recognize any scrutability user privileges when I interact with various providers

Would you like to have the option of been allowed to inspect and alter, in a variety of ways, the way you are been modeled by various providers (such as MySpace and eBay)?	Yes, that would be great
	Sounds good
	I don't really mind
	No
	I don't understand the question
Are you familiar with your privacy options when you interact with various providers (e.g. MySpace, eBay, etc.)? Do you take advantage of the offered privacy options?	I am familiar with my privacy options and I do take advantage of them
	I am familiar with my privacy options but I do NOT take advantage of them
	I am NOT familiar with my privacy options but I do take advantage of them
	I am NOT familiar with my privacy options NOR I take advantage of them
Do you think you have the freedom of customising your privacy settings when interacting with various providers (e.g. Facebook, Amazon, etc.) in order to control how the offered privacy privileges are applied to your profiles?	Yes, I have the option of customising my privacy settings when I interact with some providers
	I recognise some options for customising my privacy privileges but I would like to have more offered
	I don't recognise any options for customising my privacy settings
	No such options are offered when I interact with various providers
	I don't understand the question

Table D-15 Evaluation 2 - Pre-Questionnaire

Post-Questionnaire

Multiple-Choice Questions and List of Available Answers

Questions	Answers
Would you use such a service? Do you find it useful?	Yes I would use it and I find it very useful
	Yes I would i use it but I don't know if it's useful
	Neutral
	No I wouldn't use it but I find it useful
	No I wouldn't use it and I don't find it useful
What do you think of the level of control you had on your SUMI collection while going through the six tasks?	I love it! I had full control of my models inside SUMI
	I am satisfied with the level of control I had inside SUMI
	I am OK with the level of control I had inside SUMI
	Neutral
	You can do better, I expected more
	I am not satisfied
How much do you value scrutability now, after going through Tasks 1,2 and 6, when you interact with user modeling providers such as Facebook and Amazon?	Very much
	Some value
	Neutral
	No value
	I don't understand the question
What other scrutability privileges would you like SUMI to offer to its users?	An empty text box is provided
User-Controlled Privacy is when users are offered a variety of options for customising their own privacy privileges. After	Very much
	Some value

completing Tasks 3, 4 and 5, how much do you value user-controlled privacy when you interact with user modeling providers such as Facebook and Amazon?	Neutral
	No value
	I don't understand the question
What other user-controlled privacy privileges would you like SUMI to offer to its users?	An empty text box is provided
Finally, can you please let us know if you have participated in SUMI's 1st evaluation or is this the first time you have heard about SUMI?	Yes, I have participated in SUMI's 1st evaluation
	No, this is the first time I am participating in SUMI's evaluation process

Table D-16 Evaluation 2 - Post-Questionnaire