Enriching Lifelong User Modelling with the Social e-Networking and e-Commerce "Pieces of the Puzzle"

Demetris Kyriacou

Learning Societies Lab
School of Electronics and Computer Science, University of Southampton
Highfield, Southampton, SO17 1BJ, United Kingdom
dke06r@ecs.soton.ac.uk

Abstract. Lifelong User Modelling (LUM) is a long-term vision which is a critical element in achieving lifelong personalized learning. One of the key factors for enabling LUM is making the user models scrutable, a feature used for allowing users to inspect and alter the modelling process. A Scrutable User Modelling Infrastructure (SUMI) is proposed in this paper, which attempts to include in the LUM picture two unexplored pieces of the puzzle: the social e-networking and the e-commerce domains. SUMI gathers the various user models constructed from several (life-long) user interactions with a diversity of providers from the social e-networking and e-commerce domains, passes the control to the owners of the user models by providing scrutability and privacy privileges, and enables exchanging of user information via SUMI, among services from these two domains and educational personalization systems.

1 Introduction

Lifelong User Modelling (LUM) attempts to gather lifelong user models for enabling lifelong personalized pervasive learning, while allowing the owners of these models to scrutinize the modelling process [1]. What if we could enrich the picture in LUM, by modelling our every day (life-time) interactions with various services from the social e- networking (such as Facebook and MySpace) and the e-commerce domains (such as Amazon and eBay)? What if we could add these pieces of the LUM puzzle, resulting in richer user models for -potentially- improved personalization?

Fig. 1. The missing pieces of the LUM puzzle

My research focuses on identifying the requirements for a Scrutable User Modelling Infrastructure (SUMI) which enables gathering of user models constructed from our daily (lifelong) interactions with services from the social e-networking and e-commerce domains and exporting these models towards educational personalization systems. As the name reveals, special emphasis is given in making such an infrastructure as scrutable as possible which also consists implementing user-controlled privacy enhancements. The user-owner has full control of what happens to his/her information from the moment it is imported in SUMI until it has been exported towards educational systems.

2 Key Components of Lifelong User Modelling

Lifelong User Modelling: LUM was introduced in an attempt to model users' daily - lifelong interactions with several services on the Word Wide Web (WWW) while offering to the users the ability to scrutinize and control the whole personalization process [1]. This huge fast-growing collection of user information can play a significant role in setting the foundations for a vision of personalised lifelong learning when suitably used by adaptive educational systems, such as AHA [2] and InterBook [3]. Offering scrutability privileges to the owners of these models enables user control of how the modelling and following personalization procedure is conducted, while special consideration for privacy options empowers users to express their preferences regarding how their data will be retrieved, stored and processed [4].

Interoperability: Interoperability can be described as "a condition that exists when the distinctions between information systems are not a barrier to accomplishing a task that spans multiple systems". Exchanging user profiles across various sources in pervasive and ubiquitous environments can not be achieved if explicit and widely accepted protocols are not developed and adopted, which will allow description, discovery and exchange of user models coming from several domains, stored in various systems, written in different languages and for different platforms [5].

Scrutability & User Privacy: Offering scrutable solutions when gathering user models in pervasive and ubiquitous environments is part of the LUM vision which focuses on user-centered rather than application-centered models. It allows user control over the personalization process, enables self-awareness and acceptability for sharing one's user models for potential benefits [4]. In addition, privacy of information is a crucial factor which affects modelling decisions and influences users' trust and confidence for allowing educational systems to access and process their information for potentially improved personalization services. As the research shows, there is an essential trade-off between using personal information and the risk of exceeding the boundaries set by privacy rules [6].

3 Identified Problem

Although the LUM vision is technically still not possible, it is the way forward. It is worth finding solutions to overcome the various challenges as there are potential

benefits, especially in terms of effective sharing of user models with educational systems and between people [1]. One important step consists identifying that although User Modelling is moving forward, is still been applied single-dimensionally: Most adaptive systems developed, are only using their own internal models when offering personalization services to their users. In addition, newly introduced frameworks and architectures, while offering a solution in achieving interoperability across peer systems for exchanging fragments of user models, do not meet some special requirements for involving systems beyond the educational domain. Furthermore, most User Modelling Servers, a client-server architecture for allowing central information storing and simultaneously data access and retrieval, although are considering and offering scrutability and privacy options to their users, are mostly designed and developed to meet commercial requirements [7].

We are loosing user information, which is flowing around the WWW, because we are not thinking multi-'*domai*n'sionally. Two un-explored rich sources for user information are the social e-networking and e-commerce domains. We can enrich LUM if we find a way to model our every day (life-long) interactions with services on the WWW from the social e-networking and the e-commerce domains, in order to enrich user information sets, which are used by educational systems for personalization purposes. But is this feasible? Where do current services/platforms from these two domains stand regarding this issue?

Social e-Networking domain: For the purpose of this research we have analyzed the two best representatives, based on their liquidity, of the social e-networking domain: *Facebook and Google's OpenSocial*, to determine if such a multi-'*domain'*sional vision is possible. A variety of APIs is being offered by these two key platforms, for application development inside participating websites. Developers can take advantage of the API method calls to retrieve user data, and store it after direct consent by the 'user-owner' of that data. In addition, recent data portability initiatives such as Facebook Connect [8] and Google's Friend Connect [9] allow developers to integrate the power and flexibility of these platforms into third-party websites.

E-Commerce domain: Using the same approach, we have analyzed the two best representatives, again based on their liquidity, of the e-commerce domain: *Amazon and eBay*. Just like with the social e-networking providers, a variety of APIs is being offered by these two key representatives, for developers to take advantage in order to retrieve user data and store it, after direct consent by the 'user-owner' of that data, for a small transaction fee.

4 Proposed Solution

Special consideration has been given to collecting the requirements for employing a Scrutable User Modelling Infrastructure (SUMI), in an attempt to enable exchanging of user models from the social networking and e-commerce domains towards educational personalization services [10]. A SUMI model can be considered as a model of models, where users can add models, from the two examined domains to their SUMI collection, import the content of these models using the available providers' APIs, define the privacy status of each one of these models or parts of

them, and should they decide to do so, export their models, or parts of them, to subscribed educational personalization systems.

4.1 Achieving Interoperability across Domains

SUMI Models' Architecture: The first step requires defining the architecture of SUMI models which will "prepare" the infrastructure to accept the various models imported by the SUMI users. The architecture of SUMI models was resulted after a comparative evaluation of the two representatives in each of the examined domains and consists of 4 categories:

Generic User Data: Any input that is being entered by the user manually in any way AND it is common to both domain representatives; 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 it is NOT common to both domain representatives; 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, which is generated by the provider based on previous user input AND it is common to both domain representatives; 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, which is generated by the provider based on previous user input AND it is NOT common to both domain representatives; e.g. Amazon's "Items New Releases" which you do not find in eBay.

SUMI Ontology: A SUMI ontology has been developed, and can be found at [11], while taking in mind the structure of the analyzed representatives from the social networking and e-commerce domains, in order to enable mapping of the various providers' data models for successful communication between them and educational personalization systems via SUMI. The ontology introduces five new categories while adopting some other categories from an existing and well-accepted ontology: the General User Modeling Ontology (GUMO), which was created for uniform interpretation of distributed models in intelligent semantic web enriched environments [12]. The reason for extending GUMO is because of not adequate representation of all attributes, which can be found in the two examined domains, inside GUMO.

Defining meaning with Dictionary Concepts: Defining attributes' meaning using a dictionary, which was written expressly for the purpose of explaining terms to people, can always explain any ontological reasoning and relationships by showing the user the relevant dictionary entries [13]. SUMI follows the same approach and uses the online Oxford English Dictionary [14] to provide to users, while inspecting their SUMI collection, explicit and widely-accepted meaning to the various attributes gathered from providers of the two examined domains.

A RESTful Approach: By analyzing the API versions that are available from the representatives of the two examined domains, one can identify 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 advance solution to an 'easier' problem. A more 'relaxed' approach can fit perfectly to the

requirements for enabling communication among providers and exchanging user information between them and educational personalization 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 by a URI which can be accessed via an HTTP method for retrieving the attribute's value. These URIs are exposed to subscribed educational systems, after direct inspection and approval of the users-owners for enabling export of their information towards the subscribed educational systems.

4.2 Evaluating 3 Scrutability & 3 User-Controlled Privacy Privileges

As already stated, our goal is to make SUMI as scrutable as possible while implementing user-controlled privacy enhancements. A SUMI evaluation was conducted with 107 users, which were exposed to a prototype service, in order to determine whether the proposed scrutability and privacy privileges were acceptable to the users, whether the users were able to complete assigned tasks, and whether they understood the consequences of their interactions with the system while completing the tasks. The results revealed that the users expressed their general approval of the proposed privileges while proposing useful suggestions regarding improvements to the presentation and interface to the system [15].

Scrutability Privileges: The examined scrutability privileges which we offered to our users during the evaluation allowed them to determine which of their models would be added to their SUMI collection, which of the four SUMI categories of their added models would import its content and also examined how users reacted when they were asked to inspect the export process before approving it to go through.

User-Controlled Privacy Privileges: Three proposed privacy privileges were designed for controlling how other users would access a user's SUMI collection. The first privilege consisted of setting the privacy status for all 4 categories of all previously added models. Users had the choice of three privacy statuses: *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, and *hidden* – others can not see that the model exists, therefore the model's content is accessed only by the model's owner. The second privilege consisted of responding to viewing requests from other SUMI users to some 'private' categories of previously added models, and the third privilege was about visiting other users' SUMI models and placing viewing requests to some 'private' categories of those users' previously added models.

5 Future Work Schedule

Further work is required to confidently answer our main research question stated in the introduction. Consequently a future work schedule has been developed and it includes: further analysis of the announced data portability initiatives [8, 9] and the available APIs in order to identify any potential challenges for implementing such a proposed SUMI service, working closely with educational personalization systems,

such as AHA![3], for developing real-life adaptive lessons which will demonstrate how SUMI can enrich user models with sets of information from the two examined domains, and analyzing how adopting REST as SUMI's communication protocol can influence any design decisions. Finally, a second user evaluation is already underway, which builds on the feedback from the first evaluation and provides improved scrutability and user-controlled privacy privileges to SUMI users. Until what point can we pass the control to the users, and by adopting which scrutability and privacy privileges can we best achieve this? Can users cope with freedom to inspect and alter the way they are being modeled, or do they need guidance on how to proceed?

References

1. Kay, J., Lifelong learner modeling for lifelong personalized pervasive learning. IEEE Trans on Learning Technologie*s*, 1(4), 215--228 (2008)
2. De Bra, P., Calvi, L., AHA: a Generic Adaptive Hypermedia System, In Proc. of the 2nd Workshop on Adaptive Hypertext and Hypermedia, Pittsburgh, 5--12 (1998)
3. Brusilovsky, P., Eklund, J., Schwarz, E., Web-based education for all: A tool for developing adaptive courseware. In Computer Networks and ISDN Systems, 30 (1-7), 291--300 (1998)
4. Kay, J., Scrutable adaptation: because we can and must. In V. Wade, H. Ashman, and B.Smyth, editors, Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems, 4th International Conference, AH2006, pages 11-19. Springer, (2006).
5. Dolog, P., Schaefer, M., A Framework for Browsing, Manipulating and Maintaining Interoperable Learner Profiles. In Proc. of the 10th International Conference on User Modelling, Edinburgh, UK, (2005)
6. Kobsa, A., Privacy-Enhanced Personalization. In Communications of the ACM, 50(8), 24-33 (2007).
7. Fink J., Kobsa, A., A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web. In User Modeling and User-Adapted Interaction - Special Issue on Deployed User Modeling, 10, 209--249 (2000)
8. Facebook Developers News, Announcing Facebook Connect, Last Retrieved on 22/03/2009 at http://developers.facebook.com/news.php?blog=1&story=108
9. Google Inc, Google Friend Connect, Last Retrieved on 22/03/2009 at http://www.google.com/friendconnect
10. Kyriacou D., A Scrutable User Modelling Infrastructure for enabling life-long User Modelling. In Adaptive Hypermedia and Adaptive Web-Based Systems - Doctoral Consortium, Hannover, Germany, (2008)
11. SUMI Ontology, Last Uploaded on 15/01/2009 at http://mysumi.org/sumiOntology.owl
12. GUMO Ontology and UbisWorld, Last Accessed on 22/03/2009 at http://www.ubisworld.org/
13. Apted, T., Lum, A., and Kay, J. Supporting metadata creation with an ontology built from an extensible dictionary. In: P.D. Bra and W. Nejdl, Editors. Adaptive Hypermedia and Adaptive Web-based Systems. Springer, (2004)
14. The Oxford English Dictionary Online, Last Accessed on 22/03/2009 at http://www.oed.com/
15. Kyriacou, D., Davis, H., and Tiropanis, T., Evaluating Three Scrutability and Three Privacy User Privileges for a Scrutable User Modelling Infrastructure. In: User Modeling, Adaptation and Personalization, 22-26 June 2009, Trento, Italy, (2009).