Ontologies for Quantified Self: a semantic approach

Federica Cena, Silvia Likavec,
Amon Rapp, Martina Deplano
University of Turin
Corso Svizzera 185, Torino, Italy
{cena,likavec,rapp,deplano}@di.unito.it

Alessandro Marcengo
Telecom Italia – Research and Prototyping Department
Via Reis Romoli 274, Torino, Italy
alessandro.marcengo@telecomitalia.it

ABSTRACT
The spreading of devices and applications that allow people to collect personal information opens new opportunities for user modeling (UM). In this new scenario UM together with personal informatics (PI) can offer a new way for self-monitoring that can provide the users with a sophisticated mirror of their behavior, attitudes and habits and their consequences on their life, on the environment and on contexts in which they live in. These new forms of self-reflection and self-knowledge can trigger and motivate the behavior change. In this paper we describe the first step in this direction, focusing on opportunities offered by semantic web ontologies for data integration and reasoning over data for recommendation purposes.

Categories and Subject Descriptors
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

General Terms
Languages.

Keywords
Ontologies, User model, Personal informatics, Quantified Self.

1. INTRODUCTION
Personalized systems are used to meet individual preferences and needs of each specific user, thus tailoring the system response to these particular requirements. Personalized systems extrapolate users' interests and preferences from explicit user ratings and from the observation of user behavior on the web: the system's assumptions about the user based on these observations are stored in a User Model (UM) [1]. A user model is the repository of personal information that has the potential to drive personalization and learning. The UM contains different types of information: from user demographic data to domain-specific preferences data (interest, knowledge…).

On the other hand, Personal Informatics (PI), also known as Quantified Self (QS), is a school of thought which aims to use the increasingly popular invisible technology means for acquiring and collecting data on different aspects of the daily lives of people. They allow users to self-track a variety of data about their own behavior: these data can be, on the one hand, user physical states (such as glucose level in the blood), psychological states (such as mood), behavior (such as movements), habits (such as food intake, sleep); on the other hand, they can be environmental parameters (such as CO2 content, temperature) and contextual information (such as people meeting) of the places passed through by the users during their everyday life. Thus, with this technology, we have the capability to automatically record at large scale the places that the users have been to, things they have seen, how they sleep, how active they are, etc., creating a constant stream of data that can reveal many aspects of their lives.

However, today all these data are scattered in autonomous silos and not integrated. UM techniques have the potential of aggregating and correlating data not only coming from web browsing but also provided by all these PI systems. A UM enriched with a plethora of personal data (behavioral, psychological, physical and environmental), related to different aspects of a person's daily life, will be able to provide the user with a "mirror" of herself, a sophisticated representation of interests, habits, activities in her life, in a novel way that is not yet achieved by any of the personal informatics tools available today [2]. This can support a new complex form of self-awareness and self-knowledge, which could foster behavior change processes [3], promoting more sustainable or healthier behavior, discouraging bad habits, sustaining therapeutic improvement and managing chronic diseases.

In this new scenario UM together with PI can offer a new way for self-monitoring people's own behavior, where self-monitoring refers to an assessment strategy to increase a person's awareness of targeted behavior [4], in order to promote behavior change [5].

UM and PI can provide users with a sophisticated mirror of their behavior, attitudes and habits, highlighting their consequences on their life, on the environment and on contexts in which they live in, promoting a new form of self-reflection and self-knowledge that can trigger and motivate the behavior change.

Our **goal** is to design a sophisticated UM-based PI system which can:

i) gather heterogeneous types of user data (from PI systems' sensors, from social web activities, from user's browsing behavior) and integrate them in an enhanced UM;

ii) reason on the gathered data in order to find aggregations and correlations among data;

iii) provide users with recommendations and meaningful UM visualizations to support self-awareness and self-knowledge.

The paper is structured as follows. We first present our solutions and then we focus on semantic modeling of the domain in order to allow data integration and reasoning.

2. STATE OF THE ART
Traditionally, **User Models (UMs)** [1,6] have the following features: (i) they are restricted to a single application; (ii) data are derived from the web; (iii) they concern short periods of time.

With the advent of ubiquitous computing technologies we are able to track and store large amounts of various personal information, scattered among applications and not integrated [7] even though it is possible to integrate them with semantic web techniques [8]. This project will advance the UM state of the art in the following:

- the integration of data derived from everyday life, in addition to the data derived from the web;
- reasoning on that data to gain further correlations about user behavior.

The opportunity is related to obtaining a Lifelong user model that stores user information for a long period of time and is able to manage user interest change [9]. This project is a first step in this direction.

According to [10], an **ontology** can be seen as a ‘‘formal, explicit specification of a shared conceptualization’’. With explicit specifications of domain objects and their properties, as well as the relationships between them, ontologies serve as powerful formalisms for knowledge representation, providing exact semantics for each statement and avoiding semantic ambiguities. For these reasons, ontologies are often used for semantic data integration and for resolving semantic conflicts, as in [11,12,13,14,15]. Also, the associated rigorous mechanisms allow for different forms of reasoning (for example, to deduce implicit classes), as in [16,17].

Measuring users' daily affective experiences is an important way to quantify their life. In [18], the authors measure users' emotions at various moments throughout the day. They asked the users to answer demographic and general satisfaction questions, to construct a short diary of the previous day, and then to answer structured questions about each episode. In [19], the authors investigate digital recordings of everyday activities, known as visual lifelogging, and elaborate the selection of target activities for semantic analysis. They investigate the selection of semantic concepts for life logging which includes reasoning on semantic networks using a density-based approach.

Motivating behavior change towards a more active lifestyle is a psychological, social and technological challenge. Several Personal Informatics Systems have been developed in order to try to modify a behavior by means of self-monitoring, such as [20,21,22]

3. A NOVEL SEMANTIC PI SYSTEM

We design a novel enhanced PI system, integrated in people’s everyday lives, able to gather data in a transparent way and to build and maintain a sophisticated user model able to aggregate data and provide meaningful visualization and personalized recommendations to the user for promoting behavior change. To reach this goal, we need the following components:

i) *data integration of different user data* for building a sophisticated model of user behavior, habits, needs and preferences coming from different sources (web and real life behavior)

ii) *advanced forms of reasoning* on user data for correlating different aspects of user daily behavior

iii) *personalized feedback for triggering behavior change* in the users:

- *recommendations* triggered by the correlation of different types of data (e.g., recommendations in accordance with user behavior, attitudes and habits in the UM)
- meaningful *visualization* of data for raising awareness and motivating people in changing their behavior.

In this paper we focus on data integration and reasoning over data (points i) and ii)) exploiting opportunities offered by semantic web ontologies [23]. Another challenging issue, namely gathering user data, is out of scope of this paper

4. ONTOLOGIES FOR QUANTIFIED SELF

In order to be able to:

integrate heterogeneous data coming from different devices and sources

reason on these data in order to provide meaningful visualization and recommendation

we design and develop three ontologies, modeling the three main concepts of the Quantified Self world: time, place and user activities. Vital parameters such as weight, blood pressure or blood sugar content are also important parameters, but we omit them from the preset analysis, since they are used primarily by medical experts and are hard to analyze by ordinary people.

Time ontology. We want to model the time from a user point of view, distinguishing work days, weekends and holidays (religious and civil ones), as well as dividing each day into meaningful slots (morning, afternoon, evening, night).

Place ontology. Again we want to model the place from a user perspective, labeling the places where the user lives, works or does the activities, dividing them into indoor (school, house, gym, work, cinema, restaurant, etc.) and outdoor (park, street..). (See Figure 1: Place ontology.)

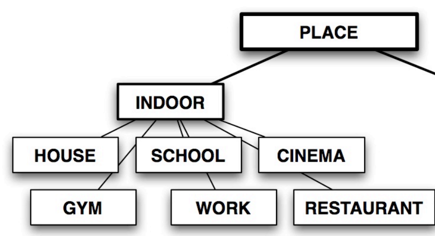

Figure 1: Place ontology

Activities ontology. We tried to model all the user activities, dividing them into two main categories: activities with place change (such as transportation or sports with place change) and activities with no place change (such as sports with no place change, intellectual activity, physical work, resting activity or feelings). Each of these classes has additional subclasses to better describe the performed activity, but we omit them from the picture for better clarity. For example, sports with place change has as its subclasses running, cycling, kayaking or downhill skiing, to name just a few. The design of this ontology was motivated by the categorization of activities in “Moves” application (https://www.moves-app.com). (See Figure 2: Activities ontology.) For lack of space, we included feelings into “Activities ontology”. We actually intend to have an additional “Wellbeing and emotions ontology” to model user’s emotional state and wellbeing, taking inspiration from [24].

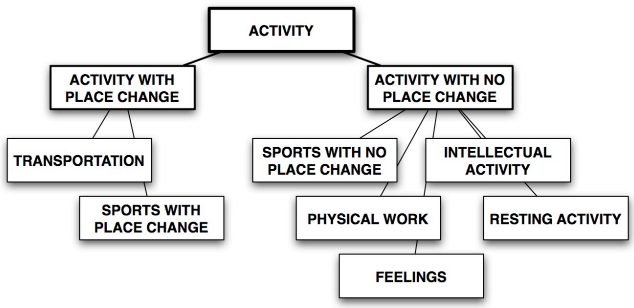

Figure 2: Activities ontology

Then, we use these ontologies in two ways.

First, we use ontologies to solve the possible data value and schema conflicts occurring among the data gathered from PI tools. As an example of data value conflicts, we gather “steps” both from the pedometer on the smart phone and from the smart bracelet and the collected numbers can differ: thus, in this case, we calculate an average number of steps. Even more challenging would be to deal with contradicting or seemingly unrelated data. For example, a pedometer might suggest that you were sedentary, while at the same time having the gym as your location. Pedometer forgotten in the locker or sitting in the gym bar? Another example concerns the mood levels: from an ad hoc app on the smart phone, we gather 4 mood values, whereas from the tangible channel we gather 6 mood values. Hence, the values should be normalized.

Schema conflicts are more complex: for example, what is modeled as an attribute in one relational schema may be modeled as an entity in another schema (e.g. "hour" as an attribute for the entity "sleep" and "hour" as an entity that has a relationship with "sleep"). As another example, two sources may use different names to represent the same concept (e.g. "running" and "jogging"), or the same name to represent different concepts, or two different ways for conveying the same information (e.g. "date of birth" and "age"). We solve these conflicts by mapping the data to our ontologies.

Second, we use these ontologies to make inferences useful for recommendation, in conjunction with Data Mining techniques for discovering correlations among data, where various forms of generalization can make correlations more powerful. For example, data mining techniques might provide a correlation between headache and running or biking activities. Since the two activities are two types of "outdoor activities" in the Activities Ontology, we can indicate a correlation between outdoor activities and headache. Alternatively, if we know that a certain user has a headache on December 24th, January 1st and August 15th, and from the Time Ontology we know that these are holidays, we can infer a correlation between holidays and headache.

Moreover, we could suggest a behavior that is similar or different but somehow related to what the user is used to doing. For example, if we know that the user loves running, but according to our data, we discovered that this is correlated with bad sleep, we might suggest some similar activities (in the same category) such as hiking or walking..

5. CONCLUSIONS

In this paper we tackle an important problem of long term management of users' data in PI systems and address a number of challenges including the need for data integration and interpretation. We motivate the introduction of suitable ontologies for modeling the core aspects of user behavior which would help overcome these problems.

This work is still at its early stage. We aim at experimentally evaluating our proposal by means of user tests to see short and long term effects of recommendations and visualizations on user behavior, as well as the acceptability of the solution.

6. ACKNOWLEDGMENTS

Our gratitude goes to Telecom Italia for their support.

7. REFERENCES

[1] Brusilovsky, P.: Methods and techniques of adaptive hypermedia. *User Modeling and User-Adappted Interaction*, v 6, n 2-3, pp 87-129, 1996

[2] Li, I., Dey, A. K., Forlizzi, J.: A Stage-Based Model of Personal Informatics Systems. In *Proceedings of SIGCHI Conf. on Human Factors in Computing Systems*, pp. 557-566. ACM, NY, USA, 2010

[3] Bandura, A.: Social Cognitive Theory of Self-Regulation. *Organizational Behavior and Human Decision Processes*, 50, 248--287 1991

[4] Burke, L. E., Wang, J., Sevick, M. A.: Self-monitoring in weight loss: a systematic review of the literature. *Journal of the American Dietetic Association*, 111(1), 92-102, 2011

[5] Bertram, C. L., Wang, J. B., Patterson, R. E., Newman, V. A., Parker, B. A., Pierce, J. P.: Web based self monitoring for weight loss among overweight/obese women at increased risk for breast cancer: the HELP pilot study. *Psycho-Oncology*, 22(8), 1821-1828, 2013

[6] Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M., von Wilamowitz-Moellendorff, M.: Gumo - the general user model ontology. In *User Modeling* 2005, pp. 428-432, Springer, 2005

[7] Aroyo, L., Dolog, P., Houben, G.-J., Kravcik, M., Naeve, A., Nilsson, M., Wild, F.: Interoperability in personalized adaptive learning. *Journal of Educational Technology and Society* 9(2), 4–18, 2006

[8] Sluijs, K.van der, Houben, G.-J.: A generic component for exchanging user models between web-based systems. *International Journal of Continuing Engineering Education and Life-Long Learning* 16(1–2), 64–76, 2006

[9] Kay, J., Kummerfeld, B. eds. Proceedings of the Lifelong User Modelling Workshop, at User Modeling Adaptation and Personalization Conf. UMAP '09, 2009

[10] Gruber, T. R.: A translation approach to portable ontology specifications. *Knowledge Acquisition Journal* 5 (2), 199–220, 1993.

[11] Arens, Y., Ciiee, Y., Knoblock. A.: SIMS Integrating data from multiple information sources. Information science institute, University of Southern California, U.S.A, 1992

[12] Goh, C.H., Bressan, S., Madnick, S. and Siegel. M.: Context interchange New features and formalisms for the intelligent integration of information. *ACM Transaction on Information Systems*, 17(3):270–290, 1999

[13] Beneventano, D., Bergamaschi, S., Guerra, F. Vincini. M. 2001: The MOMIS approach to information integration. In *ICEIS 2001, Proceedings of the 3rd Int. Conf. on Enterprise Information Systems*.

[14] Visser, P. R., Jones, D. M., Beer, M. , Bench-Capon, T., Diaz, B. and Shave, M.: Resolving ontological heterogeneity in the KRAFT project. In 10th Int. Conf. and Workshop on Database and Expert Systems Applications DEXA'99, 1999

[15] Abel, F., Herder, E., Houben, G. J., Henze, N., Krause, D.: Cross-system user modeling and personalization on the social web. *User Modeling and User-Adapted Interaction*, 23(2-3), pp.169-209, 2013

[16] Wang, X. H., Zhang D. Q., Gu, T., Pung, H., K.: Ontology Based Context Modeling and Reasoning using OWL. In *Proceedings of the 2nd IEEE Ann. Conf. on Pervasive Computing and Communications Workshops (PERCOMW '04)*. IEEE Computer Society, 2004

[17] Eiter, T., Ianni, G., Polleres, A., Schindlauer, R., Tompits, H.: Reasoning with rules and ontologies (2006), Reasoning Web 2006

[18] Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N., Stone, A.A.: A survey method for characterizing daily life experience: The day reconstruction method, Science 306, pp. 1776–1780, 2004

[19] Wang, P., Smeaton, A.F.: Semantics-based selection of everyday concepts in visual lifelogging, *International Journal of Multimedia Information Retrieval* 1, pp 87–101, 2012

[20] Shumaker, A., Ockene, J. K., Riekert, K.: *The Handbook of Health Behavior Change*, Springer, 2008

[21] Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.: UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. In *Proceedings of the SIGCHI Conf. on Human Factors in Computing Systems*. pp. 1043-1052. ACM, New York, USA, 2009

[22] Kay, M., Choe, E. K., Shepherd, J., Greenstein, B., Watson, N., Consolvo, S., And Kientz, J. A.: Lullaby: a capture & access system for understanding the sleep environment. In *Proceedings of 2012 ACM Conf. on Ubiquitous Computing*. pp. 226--234, ACM, New York, USA, 2012

[23] Guarino, N.: Formal ontology and information systems. In Proceedings of the 1st Int. Conf. on Formal Ontology in Information Systems, FOIS '98, IOS Press, pp 3-15, 1998.

[24] Patti, V., Bertola, F.: Organizing Artworks in an Ontology-based Semantic Affective Space, Proceedings of the 1st Int. Workshop on Emotion and Sentiment in Social and Expressive Media: approaches and perspectives from AI, ESSEM@AI*IA, 2013.