

Some Issues in Personalization of Intelligent Systems: An Activity Theory Approach for Meta Ontology Development

Daniel E. O'Leary¹

Abstract. Personalization of systems has been part of the renaissance of artificial intelligence in many domains. This paper investigates some emerging issues in the area of personalization as they impact systems from different perspectives. Particular attention is given to the relationship between explicitly and implicitly gathered information, information gathered from other personalization settings and with the generation of a personalization information ontology, based on an activity theory approach. Finally, some privacy issues are considered, potentially limiting information sharing between applications.

1 Introduction

Personalization has been part of the renaissance of artificial intelligence [1]. There has been substantial research on personalization. For example, personalization of interaction with hardware and software has been occurring in multiple settings such as, in web navigation [2,3], Internet commerce [4] and ambient intelligence [5]. In addition, personalization is being used and proposed in many industries, e.g., consumer electronics [5] and health care [6].

1.1 What is personalization?

Personalization has many characteristics, of which some or all may be employed in any particular setting, including the following. First, personalization tries to limit a user's workload and provide context by facilitating *remembering* key aspects of system use by a particular individual. At one extreme, this means providing a history of what has been examined, and when (e.g., web pages, files, etc.). Second, personalization can facilitate *security and privacy* by ensuring that only a particular user makes use of an application [7]. If the system is personalized

¹ University of Southern California, 3660 Trousdale Parkway, Los Angeles, CA 91011, USA
email: oleary@usc.edu

“enough” it would recognize an intruder. Third, personalization can be used to try and make applications more easy-to-use by presenting results in particular ways, ranging from font size to icon to language presentation to other issues. Fourth, personalization will adapt and evolve as the user's interest's change.

Personalization uses whatever data may be available, that is generated by particular individuals. Specific applications typically gather whatever data they can that will facilitate their need to provide personalization. In order to personalize, that data is then mined in order to generate information to facilitate modeling user behavior and interaction. Generally, personalization is based on what the user does, not what they say they do. Accordingly, data is gathered implicitly, in the background. However, in many settings, systems explicitly gather data regarding what the user wants or say that they want. Seldom do systems interact to share personalization data.

1.2 Purpose of this Paper

The purpose of this paper is to review some emerging issues in personalization systems. In particular, this paper is concerned with generating and sharing personalization information between different applications. In addition, it is concerned about what information is needed from or about the user in order to personalize particular functions and activities. Specifically, it is aimed at developing a theory-based meta-ontology that can be used in development of an ontology for personalization systems.

1.3 Outline of this Paper

The paper proceeds in the following manner. Section 1 has introduced the paper, discussed its purpose and provided an outline of the paper. Section 2 provides a brief review of the previous research. Section 3 investigates the process of personalization. Section 4 provides a brief discussion of activity theory, a theory of behavior that is used here as a basis for personalization. Section 5 uses activity theory as a basis for generating a meta-ontology that could be used to generate an ontology for personalization. Section 6 investigates the importance of information about time. Finally, Section 7 briefly summarizes the paper, discusses an extension and reviews some of the paper's contributions.

2 A Brief Review of Previous Research

In a short paper like this, the extent to which the previous literature can be reviewed is limited. However, there has been substantial previous research in many dimensions regarding the use of personalization. Jeevan and Padhi [8] offer a recent survey of the literature of content personalization, focusing on providing a bibliography of research in the area of personalization. Approaches to capturing

personalization have included using data mining [3], intelligent agents [4], ontologies [9], and other approaches.

Although there has been some investigations of ontologies for personalization, the primary focus of those efforts seems to have been in the generation of ontologies for individuals to facilitate search. However, there has been limited research investigating ontologies for personalization, in general.

3 The Process of Personalization: Data, and Mining the Data

The process of personalization requires that the system responsible for personalization gather data that allows the system to personalize itself or allows the user to personalize the system. This is done by explicitly gathering data from the user in the foreground, or implicitly gathering data in the background or from other sources that also may be gathering personal information.

3.1 Foreground Information Gathering

The system can gather information directly from the user as part of a configuration process in order to facilitate the personalization process. As an example, multilingual applications typically ask the user which language the user would like to use. Gathering personalization information in the foreground generally is seen as obtrusive. On the other hand, oftentimes, such data gathering is expected by the user or implementer to facilitate implementation. Data gathered in the foreground can be viewed as data that the user *says is true* or characteristic. Ultimately, such data may contrast with data gathered from the user as to what they actually do

3.2 Background Information Gathering

The system also can gather information in the background, as the user works or otherwise makes use of the system or its interface.

As an example, [9] developed a system that “watches” a user and gathers information regarding a user’s search while developing profiles based on concept hierarchies designed to facilitate future search for the user. Such background gathered information is based on what the user *actually* does.

3.3 Gather Personalization Information from Other Sources

It appears that in most settings, systems function independently of other systems, each gathering personalization information. However, since personalization is used in a number of settings and applications, one approach would be to gather personalization information from other applications, minimizing redundant work.

As a result, in these settings it would facilitate personalization if there were a standard set of processes and a standard language relating to personalization information, so that such information could be shared. Alternatively, if personalization data from one source could be captured and translated to another

source, that could also facilitate cross use of personalization information. For example, armed with an ontology of personalization information from one source, it would be possible to cross link that ontology with other ontologies to facilitate integration of both and cross use of personalization information.

There are some potential advantages and disadvantages of using personalization information from other sources. Advantages include less work and less time before a system is personalized. For example, if a system is able to directly import personalization information, then the user will not need to replicate, either explicitly or implicitly, generation of that information. This would speed personalization substantially. Disadvantages include misuse and misunderstanding of the information. Data gathered from other applications may not be as reliable as data gathered by the current application. Unfortunately, there is no guarantee that the data has been gathered in the particular setting etc. Further, there is no guarantee as to when the data was gathered, so the personalization could be dated. As a result, there may be concern over the source of the data. Finally, there may be privacy issues if applications share data. Some applications may be based on a web server or by other means provide personalization information to other sources.

3.4 Comparing What We Do and What We Say We Do

Personalization can allow us to map out differences between what we say and what we do, if both sets of data are captured, analyzed and compared.

If both data sets are available then we can compare the two, and in so doing provide a management capability. For example, data could be gathered with respect to when a particular assignment is due, and that information could be used to help a user make sure that they meet the specific deadline. The key issue is “are we doing what we said we wanted to do?” In this setting, the system becomes our alter ego, there to remind and, potentially cajole us to complete our tasks in a timely manner.

Second, if both sets of information are available, then they can be used to provide a quality check. If the two do not reconcile then that can indicate a difference that may deserve additional context generation.

Third, comparing the two data may provide a security check. In particular, if the two are not in sync, then the actual user may not be a legitimate user [7].

3.5 Forgetting Can be More Important

True personalization also requires knowing when things should be “remembered” and when things should be “forgotten.” People change and their interests change. People forget things and move on to other activities. Accordingly, systems also need to forget. A system that is personalized to include old behaviors is not likely to be effective or personal.

4 Activity Theory

Research on human behavior has resulted in construction of what is referred to as “activity theory” [10,11,12]. Detailed discussion of activity theory is beyond the scope of this particular paper. However, because activity theory provides a model of human behavior in context, it also provides a basis for analyzing ontological requirements of personalization systems.

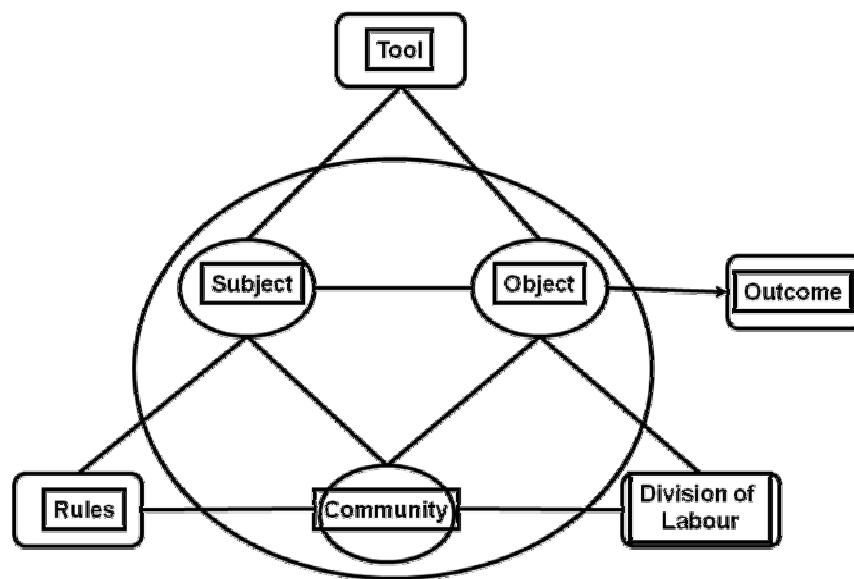


Figure 1 An Activity Theory Model of Behavior [12]

A summary of activity theory is provided in figure 1. Activity theory is based on the notion that subjects (people) perform “activities” or events in a context, using tools (e.g., software and hardware), rules (e.g., rules of interaction or behavior or organization rules), division of labour (each person has their own job to do), with an object of the activity (e.g., something being modified or created by the activity, such as a resource), while based in a community of others. Ultimately, the activity results in an outcome. As an example, in ecommerce, the activity could be a purchase of a pair of jeans, while in knowledge management it could be to produce knowledge artifacts. The subject would be the person making the purchase, the rules could be explicit or implicit (all my friends wear Levi’s jeans, so I need to buy Levi’s). The community could be the reference group, while the tools could be the computer, browser etc. Personalization would need to consider these broad categories, with respect to a particular individual.

5 Personalizing User Information Needs: Selected Ontology Requirements

Ontology requirements have been build to facilitate personalized search (e.g., [9, 13]) however, there has been only limited research structuring an ontology for personalization systems. Accordingly, this section outlines ontological requirements for personalization. In particular, using activity theory, we can anticipate certain characteristics that systems likely need to know ontology-based information about

- Personalization needs more than just history – *Activity Types*
- Personalization needs information about others – *Subject Information*
- Personalization systems need to consider the subject's position in their organization, i.e., the overall *community* in which the activity takes place.
- Some user focused objects are more “persistent” – *Object Types*
- Personalization occurs in a context where there may or may not be related available information – *Tool information*
- Personalization must take into account *rules* that the subject must follow
- Personalization needs to account for the particular portion of the project the subject is responsible for – *Division of labour*

5.1 Personalization Needs More than Just History – *Activity Types*

Some user information needs rapidly start and stop, after particular activities or events occur. Users have specific purposes and needs and when those purposes and needs are met there may not be a need to review any aspect of that process.

Consider a consumer who needs a new pair of jeans, so the outcome of the activity is the purchase of a new pair of jeans. After the jeans have been purchased, there is no longer a need for additional new jeans. For most people that means that they will not be searching for jeans for a long time. As a result, any history regarding the purchase of those jeans is not needed for a long time. After this purchase event there is limited interest in the history of this activity.

As another example, consider a researcher who is interested in examining other research papers on a topic, to first see what has been done in an area, and second, to see if a specific issue has been addressed, with the purpose of doing research on that topic and writing a paper on that topic. For most researchers, this means that once those papers have been identified, they are not interested in constantly reviewing the existing papers or finding new partially related papers. Recent history in these cases is not relevant to unfolding events.

In both of these cases, history is no longer important after some activity has occurred. Accordingly, the system needs to know that with the occurrence of some activities, there is no longer immediate interest in a particular topic. Histories must selectively remember and forget.

5.2 Subject Roles and Goals – Subject Information

The notion that different agents and participants have different roles long has been a part of system architectures. The nature of those roles is likely to vary based on the particular application. For example, [14] used basic economic roles of supply and demand to develop an architecture for intelligent agents in an ecommerce setting. Accordingly, roles and goals are likely to depend on the particular personalization application. However, in general, in personalization settings, a key agent role is one where the agent is required to “spy” on the user in the background, while another agent role may be to analyze data. A third role is likely to be one that makes sure that any appropriate rules, organizational or other are adhered to, as discussed below.

5.3 Some User Information Needs are Persistent, But Others Stop on Fulfillment – Object Types

In ecommerce and other settings, personalization systems need to take into account the type of goods, or objects being pursued. Shopping for staples, such as food, beer and wine and music is persistent over time. Where I have been in the past is helpful in the future.

However, other needs such as that new home loan or information regarding houses disappear as soon as I get that loan or buy that new house. The information needs for all goods are not persistent over time.

As a result, knowledge of the “persistence” of information needs is critical to a user’s information needs about particular resources. Accordingly, if a system is to truly personalize for information needs over time, then there is a need to “understand” which information resource needs are persistent and which are not persistent.

5.4 Community Integration

Generally, personalization focuses on the individual, and does not consider the community in which the subject is based. Such issues can be critical. For example, if the subject is part of an organization, that organization needs to be considered. In addition, the behavior may be heavily influenced by reference groups. If so, perhaps personalization needs can be anticipated by understanding personalization needs of those in the reference groups.

5.5 Tools

Generally, personalization is limited to the software and hardware that the subject has access to. A complete ontology about personalization would need to consider the tools available to the user. Since computer-based tools evolve at such a rapid rate, an ontology of tools also would need the ability to evolve to accommodate such changes.

5.6 Rules

Subjects face a broad base of rules with which they must take into account. For example, in the case of organizations, there generally are rules that limit the amount that any one individual can purchase at any one point in time. These rules need to be accounted for as part of personalization, whether they are limiting the individual or all individuals in the group of which the subject is a member.

5.7 Division of Labour

Subjects often only perform a portion of some activity, as the tasks to complete an activity are assembled. As a result, personalization needs to take into account how labor is divided. Here an ontology would account for different jobs that need to be accommodated. Accordingly, the ontology could employ organization models, or models of different jobs.

5.8 Summary

The review of these different categories of information that map into the personalization context, suggests that no single ontology will meet all of these needs. For example, rules are likely to vary substantially from setting to setting, making it difficult for a single ontology to meet the needs of all personalization needs.

6 Some Information Needs are Time Dependent

Ontological needs also extend to concerns about time. As a result, there have been a number of ontologies developed for time including [15]. Such an ontology is important since personalization needs are dependent on what time it is, whether time of day, or time of year. For example,

- I am a soccer coach for my children. I am more concerned about soccer sites roughly from August through the beginning of December, the span of the soccer season in California.
- We have purchased toys on the Internet, usually in October or November.

How can systems get such time dependent information? There are at least two sources: user actions and user plans.

6.1 Time Dependent - User Actions

As we discussed earlier, data about user actions can be gathered unobtrusively focusing on what the user does at what point in time, and monitoring the time lines. For some settings, such as the examples listed above, multiple years of data would be required before a system could infer such results.

6.2 Time Dependent - User Plans

Gathering planning information for a single application can be difficult, because subjects have multiple simultaneous activities. So how could personalization systems find out about these alternative activities with minimal additional user activity? One approach would be to tie into a calendaring system, and the events and activities inherent in such settings[16]. This would require consistency between any calendaring ontology and any personalization ontology.

7 Summary, Extension and Contributions

This paper has investigated some issues of personalization of systems. Data for such systems can be gathered in the foreground or background. In addition, personalization can be facilitated by integrating personalization information from other systems. However, in order for systems to be able to talk with each other generally requires use of the same ontology or the ability to translate from one ontology to another.

Accordingly, this paper investigated the need for an ontology for personalization systems and laid out an initial outline of a meta ontology, based on activity theory, that could be used to generate such an ontology. Based on activity theory, the requirements for that ontology include knowledge of subjects, objects, community, rules, tools, division of labour and how time relates each.

7.1 Extensions

At some firms, such as 1-800-Flowers, there is a lot of emphasis on a determining and meeting user's specific needs as part of the information gathering process, such as sending flowers for a birthday or anniversary. In particular, the user's needs ultimately are linked to a transaction processing system. The transaction processing system behind generating the order gathers "occasion" data unobtrusively. That data can then be used to come back to the user and remind them of the occasions, and their "need" to send flowers. However, use of this information requires at least two developments. First, there needs to be a "link" between the existing order processing system that gathers the data and the personalization system. Information gathered as part of the transaction processing needs to be used in the personalization system, rather than trying to re-gather the same or related data. Second, the ontologies used by the two systems must be "equivalent," at least for the variables of direct concern, e.g., "event," "occasion" or "birthday." Without equivalence, the system link will be limited. Generating and maintaining consistent ontologies is not an easily solved problem.

Finally privacy issues may limit the ability to share personalization information between applications. Alternatively, not using the same ontology across

applications can limit some potential problems with privacy concerns, and facilitate privacy preservation.

7.2 Contributions

This paper has noted that using personalization information from other systems generating personalization information can speed and ease personalization. In addition, a meta ontology for generating personalization ontologies was developed. Activity theory can provide a basis for knowing what kind of detailed ontology information will be needed in personalization systems. It also provides us with the insight that it is unlikely that a single ontology for personalization can be generated, because of broad ranges of diverse sets of rules that individuals function under, organizational settings, rapidly changing tools and other issues.

References

1. O'Leary, D., The Internet, Intranets and the AI Renaissance, *Computer*, Jan 1997.
2. O'Leary, D., AI and Navigation on the Internet and Intranet, *IEEE Expert*, Apr 1996, pp. 8-10.
3. Mulvenna, M., Anand, S., Bucher, A., Personalization on the Net using Web Mining, *Communications of the ACM*, Aug 2000, 43(8), pp. 122-123.
4. Lee, W-P, Liu, C-H and Lu, C-C, Intelligent Agent-based Systems for Personalized Recommendations in Internet Commerce, *Exp Sys with Appl*, Vol 22, 2002, pp. 275-284.
5. Aarts, E., Ambient Intelligence: A Multimedia Perspective, *IEEE Multimedia*, Vol 11(1), Jan – Mar 2004, pp. 12 -19.
6. Riva, G., Ambient Intelligence In Health Care, *CyberPsychology & Behavior*,” 6(3), 2003, pp. 295-300.
7. O'Leary, D., Intrusion-Detection Systems, *J of Inf Sys*, Vol 6, 1, Spring, 1992, pp. 63-74.
8. Jeevan, V.K.J., Padhi, P., A selective review of research in content personalization, *Library Review*, 2006, Vol 55 (9), pp. 556 – 586.
9. Pretschner, A. and Gauch, S., Ontology Based Personalized Search, *Proc. 11th IEEE Intl. Conf. on Tools with Artificial Intelligence*, pp. 391-398, Chicago, Nov 1999.
10. Engestrom, Y. “Learning by Expanding an Activity Theoretical Approach to Developmental Research,” Accessed on Apr 9, 2008, [Orientakonsultit,Helsinki, 1987. http://lchc.ucsd.edu/MCA/Paper/Engestrom/expanding/toc.htm](http://lchc.ucsd.edu/MCA/Paper/Engestrom/expanding/toc.htm),
11. Nardi. B.(Ed), 1996, *Context and Consciousness: Activity Theory and Human Computer Interaction*, MIT press, Harvard Mass.
12. Sierhuis, M. and W. J. Clancey, Modeling and Simulating Work Practice: A Method for Work Systems Design, *IEEE Intelligent Systems*, pp. 32 – 41, Sept – Oct 2002.
13. Trajkova, J. and Gauch, S., Improving Ontology-based User Profiles, 2004, <http://www.ittc.ku.edu/keyconcept/publications/RIAO2004.pdf>, accessed on April 9, 2008.
14. Brown, C., Gasser, L., O'Leary, D., and Sangster, A., AI on the WWW: Supply and Demand Agents, *IEEE Expert*, Aug 1995, pp. 50-55.
15. Hobbs, J. and Pan, F., An Ontology of Time for the Semantic Web, 2004, *ACM*, <http://www.isi.edu/~hobbs/hobbs-pan-TALIP.pdf>, Accessed on Apr 9, 2008.
16. Cost, R., et al., “ITalks: A Case Study in the Semantic Web and DAML +OIL,” *IEEE Intelligent Systems*, Jan – Feb 2002, pp. 40 – 47.