Identities Management: An Approach to Overcome Basic Barriers in E-Commerce and Collaboration Applications

Michael Koch[1], Kathrin Moeslein[2], Petra Schubert[3], and Ulrike Lechner[4]

[1] Technische Universitaet Muenchen, Department of Informatics,
Boltzmannstr. 3, D-85748 Garching, Germany, kochm@in.tum.de

[2] AIM - Advanced Institute of Management Research, London Business School,
6-12 Huntsworth Mews, London NW1 6DD, UK, kmoeslein@london.edu

[3] Institute for Business Economics (IAB), University of Applied Sciences Basel (FHBB),
Peter-Merian-Str. 86, CH-4002 Basel, Switzerland, petra.schubert@fhbb.ch

[4] University of Bremen, Department for Mathematics and Computer Science,
PO 33 04 40, D-28334 Bremen, Germany, lechner@informatik.uni-bremen.de

Full Reference:
Koch, Michael; Moeslein, Kathrin; Schubert, Petra; Lechner, Ulrike (2004): Identities Management: An Approach to Overcome Basic Barriers in E-Commerce and Collaboration Applications, in: Proceedings of the EURAM Conference, St. Andrews, 05.-08.05.2004.

Identities Management: An Approach to Overcome Basic Barriers in E-Commerce and Collaboration Applications

Michael Koch[1], Kathrin Moeslein[2], Petra Schubert[3], and Ulrike Lechner[4]

[1] Technische Universitaet Muenchen, Department of Informatics, Boltzmannstr. 3, D-85748 Garching, Germany, kochm@in.tum.de

[2] AIM - Advanced Institute of Management Research, London Business School, 6-12 Huntsworth Mews, London NW1 6DD, UK, kmoeslein@london.edu

[3] Institute for Business Economics (IAB), University of Applied Sciences Basel (FHBB), Peter-Merian-Str. 86, CH-4002 Basel, Switzerland, petra.schubert@fhbb.ch

[4] University of Bremen, Department for Mathematics and Computer Science, PO 33 04 40, D-28334 Bremen, Germany, lechner@informatik.uni-bremen.de

Abstract

The development of the Internet was originally based on the assumption that users remain anonymous. In the real world, however, people always have an identity – often even more than one. The transfer of real world transactions to the online world therefore requires identity information. More and more services, especially in e-commerce and collaboration applications, need to identify the user for providing personalized services or for presenting the user to other users. As in real life, a user in an online environment usually plays different roles and interacts with different services hosted by different providers. Current approaches to provide identity information on the Web still force users to provide and update information about their identity for each service independently. Being contradictory to intuitive user expectation, this proves to be a basic barrier for many e-commerce and collaboration applications, it results in cold-start problems for new services and in inconvenience for the user.

The availability of identity information for user representation will be important for future Internet-based e-commerce and collaboration applications. Information about the users is needed for performing transactions, for providing personalized services, and for presenting users to each other. Identity management is about managing the information and access to that information. Identities management and central user profile repositories might help

- *to motivate users making user profile information available (because they have control and awareness about who is using it), and*
- *to enable services to provide effective personalization without cold-start problems*
- *to build social networks (the so called social capital of many business models)*

Our paper highlights the role of user-centric global identities management for future e-commerce and collaboration applications. It presents a review of the current state of the art in the area of identities management (for Intranets and for the Internet) and discusses needs and possibilities for future developments.

Keywords*: Identities management, User profiles, Personalization, Electronic Commerce, Privacy*

1. Introduction

Digital media facilitate new forms of interaction. Personalization, i.e. the adaptation of information according to the preferences of an individual, and multi-lateral communication of communities are two examples for new forms of interaction that are increasingly important in electronic commerce.

1.1. Personalization

Personalization techniques are used for tailoring information services to the needs of individual users. In marketing, personalization supports one-to-one marketing which is designed to increase the customer share over a lifetime. What used to be possible in the corner shop, since the shopkeeper used to knew her customers, will be extensively possible in the electronic medium by the storage of profiles and the automatic evaluation on the basis of predefined rules.

Technically, personalization is about selecting or filtering information objects or products for an individual by using information about the individual. Different methods are known for performing this selection. These methods range from content based filtering with rules or vector similarities to automated collaborative filtering (see Schubert & Koch, 2002 for more information). Independently of the personalization method the ability to deliver personalization is always based on the acquisition of an electronic representation of the user, a user profile. Depending on the personalization method used, there are different requirements to the content and the representation of this user profile. For content based filtering information about preferred content and relationships to content objects has to be stored. For collaborative filtering relationships to other users and ratings or comments have to be managed.

1.2. Community Support

Personalization is not the only reason for services to collect information about their users. More and more often Web-based services offer some kind of community support functionality. The users, often the customers, are not supported independently from each other but are put into contact with each other. Users are supported in exchanging information, getting in contact and communicating with each other, and doing transactions among each other. Bringing communities of people together stimulates three major potentials:

(1) the building of trust,

(2) the collection and effective use of (trusted) community information, and

(3) the economic impacts of accumulated buying power.

The economic impacts of communities for accumulating buying power have already been discussed by Hagel and Armstrong (1997). Hagel and Armstrong mainly focus on groups of Internet users that are drawn together around products or companies, and that use the extended possibilities of the online medium to cooperate and gain advantages they would not have if they were acting as isolated customers (better information, discounts). Such groups are often referred to as virtual communities of transaction (Schubert, 1999). In addition to the accumulation of buying power this type of community is a source for valuable data about the products and about community members. The additional information about the users is often the basis for personalization.

In addition to the use of information about users for personalization, in community support platforms user profiles are also needed for presenting users to each other. In communication, which is the primary activity in communities, knowing the identity of those with whom you communicate is essential for understanding and evaluating an interaction and for building trust (Donath, 1998). So the

community members have to be aware of each other and need to know each other. There is no need to have the user representation linked to the real world identity of the user – but the user representation in the community should be persistent.

In sum, for modern applications user representations have to be available for personalization and for presenting users to each other. The management of this information is a complex but crucial task. In the remainder of this paper we will

- elaborate on some of the issues of user representation and identities and highlight some core problems with user representation in e-commerce and collaboration applications (Section 2),
- discuss the concept of (user-centric) "identities management" for handling user representations and for addressing the problems mentioned before (Section 3),
- present examples of identities management in the areas of personalization in e-shops and user representation in peer-to-peer community applications (Section 4), and
- review existing work to highlight the problem of missing user control in identities management (Section 5).

In the conclusions we finally lay out a work agenda for future user representation and identities management (Section 6).

2. User Representation and Identities

2.1. Identity

In the introduction we have used the terms 'user representation', 'user profile' and 'identity' synonymously. Before continuing with our discussion we need to clarify these terms.

The Webster English Dictionary describes identity as: "*1) the condition or fact of being the same or exactly alike (sameness, oneness); 2a) the condition or fact of being a specific person or thing (individuality); b) the condition of being the same as a person or thing described or claimed*" (Webster, 1988). Hence, one important feature of identity is the process of proofing to be a specific person. Another important feature of identity is the information describing a specific person.

In e-commerce applications identity is mainly used for authentication. For personalization and for collaboration support however, identity as information describing a specific person is more important. In this paper we regard identity in the context of a persistent user profile: a set of attributes describing (an aspect of) a person or role in the digital world.

2.2. Identities and Entities

When looking at the dictionary definition of identity we find a claim for equality or 'one-ness'. A person has one identity. However, this is not true for user representations. A person may use different representations or identities at various times during his life and even maintain several at the same moment in time.

Such multiple roles are neither illegal nor used primarily for illegal purposes. Since we intend to keep the term ‘identity’ for user representations we need to introduce a new term for the ‘real-world identity’. Different authors use the term ‘entity’ for this purpose (e.g. Clarke, 2001). ‘Entity’ is used for all kinds or real-world things, including people. An entity does not necessarily have a single identity, but may have many. An identity is a particular presentation of an entity (Clarke, 2001).

See (Clarke, 1994a; Clarke 1994b; Clarke, 2001) for a further discussion of “digital persona” and the differences between entities and identities in the context of (id)entification.

2.3. Modeling Identities

When trying to capture identities in data structures different approaches are taken. The most general approach is to model an identity as a set of attribute value pairs. The attributes present in a user profile and their meaning is usually captured in a user profile model. The information in the profile ranges from the names of the user, demographic attributes and the history of past transactions to dynamic attributes such as the current location of the user as used in location-based personalization.

For the coding of simple user profile information such as address or payment information there are standards available. Examples are the vCARD standard (Howes et al., 1998) or the profile scheme included in World Wide Web Consortiums P3P specification (P3P, 2000). These approaches are mainly based on hierarchically structured sets of attribute value pairs. For more complex information such as interests or browsing histories personalization applications define proprietary codings dependent on the application and on the algorithms operating on the information. In addition to these proprietary codings used in live applications there is some work on user profiles emerging from Artificial Intelligence and Knowledge Management research. See (Fink & Kobsa, 2000) for more information on abstract modeling of user profiles and user profile servers.

2.4. Acquiring Identities

There are various methods for capturing user representations, which require a variable degree of engagement by the user. One usually distinguishes asking the user explicitly for information (fill-in-profile, explicit feedback or ratings), and observing the user (click-stream- or transaction-analysis). While the discussion of these methods is important it does not address some basic issues in user profile acquisition:

(1) Users often do not trust services that collect and use profile information (and therefore try to provide no or false information).

(2) Even in the event that the user cooperates, time and effort is needed before enough information is collected to provide appropriate recommendations.

The second issue is called “cold-start problem”. This means that users expect good recommendations from the beginning, but the system is only able to provide recommendations after having asked the user a lot of questions or after having watched the user for some time. This issue is of special importance in the field of Small and Medium-sized Enterprises (SMEs) because the portals of these companies usually have only few and short contacts with a customers. Thus, these small or

infrequently visited sites never reach the point where they acquire enough user profile to effectively work with them.

2.5. Interpreting and Understanding Identities

The information that describes the specific property of an entity, that represents it, needs to be interpreted. The information about a user (provided implicitly or explicitly) represents a current state that very much reflects the past history of the user. Often this information has to be embedded in a context (e.g. of other identities) to be interpreted in a sensible way. Consider as an example the recommendation services. A profile with a transaction history of goods bought at a specific site is of limited use. Only a comparison with other profiles or additional information about relationships among products (e.g. lists of add-on products) facilitate recommendations. Consider also the information about past transactions that is part of the identity of users at many sites. There are sites at which users typically have a small number of transactions (travel sites) and sites with many transactions (today's eBay or Amazon). The interpretation of a piece of information that is part of an identity depends on this context.

3. Identities Management

Managing the availability of identity information for applications is called "identity management" in literature. As outlined before, we are talking about managing *several* identities for one single user or entity for many applications, so we are rather using the term "identities management" instead.

Identities management can contribute to solve the two key problems mentioned in the previous section (distrust and cold-start) by

- giving control of identity information back to users (to solve the distrust issue), and
- allowing the reuse of identity information among the different personalization services (to solve the cold-start problem).

In order to harness these advantages we have to separate

- the use of identity information (in the personalization or collaboration services) and
- the storage of identity information (e.g. in a central user profile or identity server).

This separation of identities use and storing identities opens the possibility of identity reuse and provides a single location where user access control and user awareness can be implemented in a way that is being perceived trustworthy by the profile owners. However, this kind of use of user profiles by different applications also makes the modeling issue more important and raises the need for standards.

Identities management is something we do everyday in our normal conversation when we decide on what to tell one another about ourselves. In interactions with others we consider the situational context and the role we are currently playing as well as the respective relationship with the communication partner. This consideration results in different sets of information being released to different interaction partners. Sometimes this leads to a situation where a person is known under different names in different contexts, e.g. by using special names, nicknames or pseudonyms suiting the

occasion (Köhntopp & Bertold, 2000). The real-world example also shows the separation of identities storage (the person itself) and identities use (the interaction partners of the person).

In the electronic world an identities management system would allow people to define different identities, roles, associate personal data to it, and decide whom to give data and when to act anonymously. An identities management system would empower the user to maintain their privacy and control their digital identity.

3.1. Components of identities management

When looking more closely into identities management, the following tasks can be identified: 1) storing of identity data, 2) authentication of identity owners, 3) definition and evaluation of access rights.

From these basic tasks the following main components can be identified for identities management systems:

- *directory service*: maintaining information about registered users
- *authentication:* establishing the validity of the identity of a user (linking a user to an identity)
- *authorization*: controlling access to user profile information (for other services and other users)

Virtually every service (and device) today holds a directory with information about personal identities. Most services have some authentication functionality. Our vision for the future is a way to reduce this variety from currently many to a few (or just one) directory and add authorization functionality in a user-centric way.

In identities management authorization can be characterized as defining and enforcing access rights to user profile attributes. Users are not inclined to share their information with every service. Some services are more trustworthy to users than others. Requirements for authorization in identities management include flexibility, arbitrary level of detail and granularity, integration of collection limitation and purpose binding of re-quests, control over sharing and distribution of data and dependency on the privacy policy of a service.

3.2. Privacy

Providing means for defining (service-specific) access rights to the user profile is already one aspect of privacy. (Information) privacy refers to the claims of individuals that information about themselves should generally not be available to other individuals or organizations, and that, where data is possessed by another party, the individual must be able to exercise a substantial degree of control over that data and its use (Clarke, 1999). This aspect of privacy is reflected in European law.

Another way to maintain privacy has already been mentioned before: Having different views on a personal entity in form of different identities. This means that making it possible to use different identifiers already contributes to privacy.

In general, different levels of identity can be distinguished:

- *veronymous*: it is possible to derive the real entity or an entifier from the information in the identity
- *pseudonymous*: the identity (identifier) is persistent, i.e. it is used several times to indicate that this is the same person, but it is not possible to derive the entifier or other identifiers of the same entity from the attributes of the identity
- *anonymous*: the identity is only valid for one transaction or page access, it is not possible to derive identifier or entifiers of formerly used identities

Clarke has coined the term 'nym' to distinguish (id)entifiers for (id)entities from keys referencing to pseudonymous or anonymous user representations: "*A 'nym' is one or more data items relating to an (id)entity that are sufficient do distinguish it form other instances of its particular class, but without enabling association with a specific (id)entity*" (Clarke, 2001).

The importance of pseudonymous and anonymous identities has been shown in a survey conducted by Ackerman et al. (1999). Respondents were less inclined to provide information when personally identifiable information was requested. In a scenario involving a banking Web site, 58% of respondents said that they would provide information about their income, investments, and investment goals in order to receive customized investment advice. However only 35% said they would also supply their name and address so that they could receive an investment guide booklet by mail (Ackerman et all, 1999, p. 5)

On the technical side the different levels of identity or privacy deal with authorization, with determining what attributes are available to the requesting service.

There is an obvious need for mechanisms allowing users to specify and enforce their personal preferences regarding privacy (authorization). For our work on identities management we collected the following requirements for a privacy infrastructure (see (Wörndl, 2002; Wörndl, 2003) and (Wörndl & Koch, 2003) for more information):

- Flexible access right control system, e.g. through rules and negotiation
- Monitoring access rights and accesses
- Using a pseudonym instead of real identity
- Purpose binding of data accesses
- Allowing access for temporary use
- Revoking granted access rights
- Control whether user data can distributed to other services (and users)
- Integration of cryptographic techniques for anonymous data transfers
- Support from privacy authorities

Online services and businesses could benefit from a powerful, user-centric privacy architecture in identities management because users with less fear of risking their privacy are likely to make more and better personal information available to services (Köhntopp & Bertold, 2000).

4. Application Areas for Personalization and Identities Management

The authors are working on different projects that explore the usage of personalization and identities management in different contexts and from different viewpoints. In this section we will present two research areas in more detail to deliver some insight on where and how the concepts discussed up to now are put to work.

4.1. Personalization in E-Shops

Personalization or tailoring information services to the needs of individual users is an interesting feature for e-business applications, especially for B2B and B2C e-shops.

In the project PersECA (Personalization of E-Commerce Applications) we are currently implementing and testing user profiles and the related identities management for personalized customer services (Schubert & Leimstoll, 2004). The project unites different ERP vendors in Switzerland in an intent to extend their existing ERP systems with special e-business modules for personalization. Personalization addresses both vendors (in e-procurement applications) and customers in B2C and B2B applications.

We have identified the modeling of customer profiles of central importance for the project. Thereby customer profiles include information directly requested from the customer and information implicitly learned from Web activity. Table 1 lists the different types of profile information we have identified.

Profile	Content
Explicit profiles	
Identification Profile	user name, role, contact information, personal browser settings, address, payment information, IP-address, etc.
Preference Profile	self-revealed preferences [e.g. product meta data]
Socio-economic Profile	self-categorization in predefined classes [e.g. age, gender, hobbies]
Ratings	three types of ratings: of products, of reviews, of pages [scale e.g.: I like it – not for me]
Relationships	Relationships to other users/customers [e.g. buddy lists]
Reviews/Opinions	Plain text, images, videos and other material
Implicit profiles	
Transaction Profile	transaction log, product purchases linked to product meta data (purchases, inquiries, payment, etc.)
Interaction Profile	click stream (pages viewed are linked to product meta data [preference categories])
External data	Information procured from other sources [e.g. weather report, local news, events, credit rating]

Table 1: Different types of profile information [following Schubert 1999]

Depending on the personalization goals, there are different requirements regarding the contents and the representation of the profile. In PersECA we have therefore structured the personalization goals or personalization functions that are relevant for E-shops, and have related them to the different types of customer profile information listed in Table 1. The following list shows the different types of personalization functions we have identified.

- *individual offers* – products and/or services are selected individually for the customer, it is possible to hide/show predefined product categories

- *personalized recommendations* – recommendations are adapted to the customer, examples are add-on products for products that have been bought in the past, or new products in the assortment that match the preference categories

- *individual layout* – the customer can configure the user interface (content modules, navigation, design) of the e-shop individually; examples for the customization are personalized menus and form templates

- *individual pricing* – prices and discounts are adapted individually to the customer

- *ranking lists* – lists of the most popular products or services based on all orders of all customers

- *product ratings* – customers can see product ratings published by other customers

- *individual newsletter* (based on preference categories)

- *alerts, scheduler* – the customer can define rules for notifications (e.g. repeat orders at a specific date, due-dates)

- *collaboration* – the customer can contact a customer representative by pressing a button only

- *order process support* (order history, order limits, authorization)

- *(order) tracking and tracing* (delivery status)

- *ECR – efficient customer response* – when the inventory of the customer falls below a specified limit, a new order is triggered automatically

- *proactive customer service* – support, maintenance, replacement of products and services by customer specific criteria

See Schubert & Leimstoll (2004) for more information about the different personalization functions and on the requirements they pose on customer profile representation.

Identities management is currently addressed in e-business solutions by simple user management functionality (for administrators and for the customer itself). This includes enabling the customer to edit parts of the profile. However, identities management mainly is limited to one platform. Importing profile information from or sharing profile information with other applications is not supported. The "ownership" of the profiles is clearly at the company.

First attempts towards centralizing the storage of customer profile information (and making them available to several applications) are usually limited to applications in one companies, and mainly address the integration of different sales channels.

4.2. Identities in Decentralized Communities

Communication and, even more, collaboration within virtual communities depends on identities, identity management and the management of social relations between the individuals. Virtual communities – at least the virtual communities on the Internet (not community of practices) – depend on voluntary participation and contributions. The challenge is, first, to motivate users to participate and contribute and, second, to design the interaction within a community so that users adhere to the rules of the community. Identity and communication management are hereby inseparably intertwined. The design of the identity influences the design of the communication and vice versa.

Most approaches in identity and communication management of virtual communities are based on some centralized control and a centralized governance structure. Virtual communities rely on the social network between community members and the value of this network. The benefit that users gain from adhering to the community rules is so important to them that they stick to an identity and to the community rules. Decentralized communities have ways to decentralize the identities, the identity management as well as the communication management. Our example for such decentralized communities are Peer-to-Peer communities, i.e. communities, that utilize mostly Peer-to-Peer networks as interaction platforms.

Peer-to-peer networks for file sharing have renewed identity and communication management in virtual communities. This is to some extent rooted in the fact that the core interest that draws this communities together is on the border of legality. The typical identities in those networks can be characterized as anonyms or nyms. Users prefer to stay anonymous (e.g. in Gnutella) or have a pseudonym to be identified by some centralized service or by buddies in a chat. Moreover, the content and the business model of this content management is of such a nature that all the content is typically only related to a nym.

A closer look at the "first generation" of file sharing systems (e.g. Gutella, the original Napster) revealed the drawbacks of communities with such a low level of identity (and community) management. Only a small number of users contribute the majority of content, and queries are limited to a relatively small scope (Adar & Hubermann, 2000). This is considered to be an undesirable behavior both from the individual and the community point of view. Only few users contribute content and there is hardly any positive feedback from the network for these contributors. This applies to pure and hybrid peer-to-peer systems (Lechner, 2002). The drawbacks of such an architecture of "weak"

identities and consequently weak social networks are imminent and inherent. It is hard to imagine that a professional work context or a significant collaboration could be implemented on such a basis.

A “second generation” of file sharing systems implements a variety of identity and community management systems to overcome the problems of the first generation. The key points are again the identities of the single users and the management of the relations between the single users.

One approach applied by the peer-to-peer networks is standard “management by humans”. Some of the hybrid networks rely on human management to make sure that quality of content and the contribution/sharing ratio meet community specific standards. E.g. at planet-anime.org there is a hierarchy of roles that matches a hierarchy of peer nodes. The core of the network is typically built by powerful servers with good connectivity, lots of content to share and a good reputation. Small nodes, i.e. nodes with little content or resources to contribute are at the edges of the network. Members of the community manage the content, assign roles to members and foster a social network. The basis for this is the existence of permanent identities of peers and users.

A number of Peer-to-Peer networks use some form of implemented identity management to improve availability of content and the ratio contribution/sharing. Partitions of content are distributed to many users. For a download of a piece of content, several peers need to be contacted. This distributes of the load of those who distribute. Also, while downloading a piece of content, the small pieces already downloaded are available for download within the network. Note that the identities themselves are not persistent. They are however attributed with a lot of small pieces of content that are available within the network for download. This increases the availability of content, balances the contribution/sharing ratio and also the quality of content.

Another approach in identities management are virtual monetary systems. E.g. Applejuice and other file sharing systems have their own virtual currency. Users “pay” for fast downloads and top places in download queues and a payment ensures them an upper position in the download lists for certain pieces of content. The monetary approach has the distinctive advantage that users do not need a permanent identity. There are however, severe drawbacks of the approach concerning security. Fraud through copying of pieces of currency is possible. The cold start problem needs to be overcome and there the stability of the currency is problematic. E.g., Applejuice follows the approach that a peer looses all its “money” with disconnecting from the network and peers connecting to the network get a lump sum to start with. Providing contents is more worth than it costs to download them. Note that such an identity and interaction design improves the “behavior” of the community. The individuals tend to stay long and are likely provide some content. Users like only to pay for high-quality download and this benefits the content availability, and the quality of the content within the network. Note that this form of identity and community management works with non-persistent identities and without human identity and communication management.

To sum up the discussion, decentralized communities, like peer-to-peer systems establish a new approach for identity management. Traditional communities employ centralized approaches to identity and identity management. Peer-to-Peer networks have well designed interactions and rather simple identities, since the lack of centralized control and governance makes strong identities and an identity management somewhat obsolete. Communication management takes over some of the identities

management. Without identity management the quality of the interaction deteriorates – which is not desired.

5. Existing Identities Management Solutions

Originally, to the extent that software was aware of identity, it was hard-coded into an application or time-sharing system. An application may have had its own notion of a user, and a password list and authorization. Today, virtually every piece of software and every device you purchase includes a repository for identities.

Work related to exchanging user profiles and identities management can be found both in industry and research. The solutions can be separated in two basic categories. First, there is work on client based profile management. These infomediary solutions store identity information on the user's computer. Second, there are server based identity management solutions.

5.1. Infomediaries

Instead of storing identity information in different services, the information can be stored on the user's computer, and be provided to services when needed. This could lead to higher trust because personal information is located near the user and because the usage of profile information can be controlled and monitored.

Client-side user profile storage is implemented by so-called infomediaries. Infomediaries are (small) applications on the client computer, which manage user profiles and offer services such as automatic fill-in of Web forms or P3P interfaces for exchanging the information with services (Cranor, 1999). Examples for infomediaries are Jotter (www.jotter.com) or Persona (www.persona.com).

The main problem of client-side storage of user profile information is that it is not portable (Mulligan & Schwartz, 2000). Personal information stored on one computer (e.g. at work) cannot be easily transferred to another one (e.g. at home or a mobile device). An additional problem with today's infomediaries is that the definition of access rights is possible but much too complex for everyday usage.

5.2. Single-Sign-On and Server-based User Profile Databases

While the infomediaries focus on user control of identity information (authorization) server-side solutions are often more service-platform-centered and focus on authentication only. These solutions relate back to multi-server authentication solutions like Kerberos (Steiner et al., 1988). In such single-sign-on (SSO) solutions different servers or services share one service to authenticate users.

Today different software vendors offer single-sign-on solutions for Intranets. The solutions are mainly based on (X.500) directory services or at least accessible via the LDAP directory access protocol.

While the single-sign-on solutions mentioned before are tailored for Intranet usage, global solutions like DigitalMe from Novell (www.digitalme.com) or Microsoft Passport (www.passport.com) extend this approach to a service for Internet usage (Dyson, 2002). The core of these services is a central user

profile directory. Users can store and maintain their personal data in these directory servers via Web interfaces. Services that are certified by the profile storage operator can get access to the authentication and profile information when a user tries to log in at these services.

Other central user profile repositories are even more focused on marketing. iFAY (www.ifay.com) or Yodlee (www.yodlee.com) support clustering users and making the information about the affiliation to clusters available to services that pay for it. In addition to the large identity management networks like Microsoft Passport several smaller projects have appeared. Examples are XNS (www.xns.org) and Live-id.org (www.live-id.org). These companies mainly follow a federated approach that allows for different identity servers operated by different companies.

5.3. Microsoft .NET Passport

Passport is the service that made identities management famous. Passport lets you sign up with a minimum of (unverified) personal data: working e-mail address and a pass-word. So Passport provides “persistent pseudonyms” – as many as you want. In its pure form passport is just authentication. It does not assert attributes other than the correspondence between identity and email address. The main service it offers is a single-sign-on which realizes the authentication for different services. Microsoft is using its market power to propagate Passport by making Passport account obligatory for using all Microsoft online services including the Hotmail email service. Therefore, the Passport service currently reports several hundred million accounts and several billion authentications per month.

On the technical side “http redirect” is used for authentication. Thereby, the user’s Web browser is redirected from the service to the identity server if no authentication token is present. The IM system then handles the authentication of the user and sets the authentication token (usually a cookie) that can be used to authenticate the user at further ser-vices.

On the privacy side Passport provides an “opt in”. You have to give explicit permission to have your profile information shared. However, you do no longer have a possibility to control with whom the information is shared when you have opted in nor do you get awareness of how your information is used. Microsoft can forward it to all partners and accept new partners without notifying you.

In addition to the weak privacy Microsoft Passport has been criticized for security problems (Kormann & Rubin, 2000) and lack of privacy considerations. For example, problems with the “http redirect” include potential eavesdropping of the transmission of authentication information and illicit use of stolen authentication tokens.

5.4. Liberty Alliance

Even the hundreds of millions accounts in Microsoft Passport look pale beside today’s large-scale production authentication systems like Visa and Mastercard – which handle not just lightweight authentication for log-ins, but financial transaction.

To allow interoperability, the identity management providers and other companies that are operating these large identification services have joined in the Liberty Alliance to develop a standard for connecting their identification and user profile storage services in a federated way. See

(AberdeenGroup, 2002), (Sun, 2002a; Sun, 2002b) for more information on the industries viewpoint on federated identity services.

The Liberty Alliance is currently on the way to define an open standard for the representation of identities, for the authentication of users and for authorizing access to user profile information. The goal is to make it easy for services that are storing user profile information to exchange the information among each other (Liberty, 2002). The idea of Liberty Alliance is a kind of balance-of-powers notion where all the companies compete to be the customers first point of contact, whereas the Microsoft approach simply takes it for granted that Microsoft is the primary point of contact (for authentication, at least). The main idea is not to create a platform for sharing personal data, but rather for passing and linking unique identifiers and confirming that they have been authenticated. So, the identities will not be available to all members instantly, but information is forwarded from one member to another one based on bilateral contracts (circle of trust).

However, the focus of Liberty Alliance still is on authentication only. There is no real user control in authorization built into the proposal yet.

5.5. Profile Information Exchange

Other related work in the commercial field is about exchange and synchronization of user profile information among users or among applications of one user.

Examples for the replication of user data among users are business card exchange ser-vices (see www.cardxchange.net for one example). In these services users can store their contact information (and any additional attributes) and make a subset (view) of this information explicitly available to other users. When the information is changed by the owner the electronic business card changes at all places or the people replicating the information are notified by email. Similar functionality is often built into Community Support platforms. With the FOAF (Friend-of-a-Friend) standard the World Wide Web Consortium is currently evaluating a RDF based representation of identities that allows business card exchange and more sophisticated applications.

Exchange and synchronization of PIM data (Personal Information Management – calendar, address and todo lists, notes) up to now has been restricted by the large number of proprietary protocols on the market, each focusing on only a small number of devices, applications and data types. SyncML (www.syncml.org) - an open industry standard for the synchronization of remote personal data across multiple networks, applications, plat-forms and devices - resolves this issue by providing a level of interoperability that is not possible with the industry's current proprietary synchronization protocols. In the context of SyncML also different data standards like vCard and vCal are promoted.

6 Summary and Conclusions

The availability of identity information for user representation will be important for future Internet based E-Commerce and Collaboration applications. Information about the users is needed for performing transactions, for providing personalized services, and for presenting users to each other.

Identities management and central user profile repositories might help

- to motivate users making user profile information available (because they have control and awareness about who is using it), and
- to enable services to provide effective personalization without cold-start problems.

These two effects could help to boost the use of personalization in online services.

Ultimate benefit of activities like Microsoft Passport or the Liberty Alliance project will be to make authentication and data sharing practices open and visible.

To gain trust from the profile owners a solution has to clearly support

- definition and handling of different access rights and/or sub-identities
- provision of awareness of access to the profile information

The functionality has to be provided in an intuitive way and has to cover the emerging mobile applications that also need user profile information for performing their services. Therefore, different (trusted) operators for identity management servers are needed to choose from.

We have built a tool called IDRepository that extends the possibilities of these solutions by providing a user centric identities management while preserving interoperability with emerging networks where possible (see Koch 2002, Koch & Wörndl 2001).

However, there are still some challenges to be addressed. From the technical point of view the most important issues are:

- How to specify (and enforce) access rights (especially including usability and user interface issues – see Jendricke & Gerd tom Markotten (2000) for some information on this issue).
- How to represent user profile data to make it usable by different services (up to now nobody has dealt with user profile structure very much).

Some of the issues cannot be solved through technology alone. Especially the issue of access right enforcement. As already discussed in the P3P project of the World Wide Web Consortium (W3C) a certification of services is needed to ensure that the services make correct statements about planned user profile usage. An issue linked with the ser-vice certification is the selection of a trusted operator for the identities management service. Here we have taken an approach that allows different providers to operate identity servers and allows the user to select.

Another topic that has to be addressed is business models for future identity management providers. First ideas for such business models are drawing from analogies of “user profile banks” with classical banks that have gained trust and are providing access to money from everywhere.

When this issue is solved we can also extend the scope of central identities management solutions to appliances, i.e. to have personal appliances load user profile information for personalization from the central repositories.

References

AberdeenGroup (2002): Federated Identity Systems – An Executive White Paper, Technical Report, Aberdeen Group, Boston, MA, Jun 2002.

Ackerman, M. S.; Cranor, L.F. and Reagle, J. (1999): Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences, Proc. ACM Conference on Electronic Commerce, Nov. 1999.

Adar, E. and Huberman, B. (2000): Freeriding on Gnutella. Firstmonday 5 (10).

Clarke, R. (1994a): The Digital Persona and its Application to Data Surveillance. The Information Society, 10(2) June 1994.

Clarke, R. (1994b): Human Identification in Information Systems: Management Challenges and Public Policy Issues. Information Technology & People, 7(4) December 1994, pp 6-37.

Clarke, R. (1999): Internet Privacy Concern Confirm the Case for Intervention, Communications of the ACM, Vol. 42, No. 2, Feb. 1999, pp. 60 – 67.

Clarke, R. (2001): Authentication: A Sufficiently Rich Model to Enable e-Business, Working Paper, 2001, http://www.anu.edu.au/people/Roger.Clarke/EC/AuthModel.html

Cranor, L.F. (1999): Agents of Choice: Tools that Facilitate Notice and Choice about Web Site Data Practices, Proc. 21st Intl. Conf. on Privacy and Personal Data Protection, Hong Kong, China, 1999.

Dyson, E. (2002): Digital Identity Management. Release 1.0, 6(20), June 2002.

Fink, J.; 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 10, pp. 209 – 249, 2000.

Hagel, J.; Armstrong, A. (1997): Net Gain: Expanding markets through virtual communities, Boston, MA: Harvard Business School Press, 1997.

Howes, T.; Smith, M.; Dawson, F. (1998): MIME Content-Type for Directory Information (vCARD Specification), RFC 2425, 1998.

Jendricke, U.; Gerd tom Markotten, D. (2000): Usability meets Security – The Identity-Manager as your Personal Security Assistant for the Internet, Proc. of the 16th Annual Computer Security Applications Conference, 2000.

Koch, M. (2002): Interoperable Community Platforms and Identity Management in the University Domain, International Journal on Media Management, 4(1), pp 21-30.

Koch, M.; Wörndl, W. (2001): Community-Support and Identity Management. In: Proc. European Conf. on Computer-Supported Cooperative Work (ECSCW2001), Bonn, Germany, pp. 319-338.

Köhntopp, M.; Bertold, O. (2000): Identity Management Based on P3P, Proc. Workshop on Design Issues in Anonymity and Unobservability, Berkeley, CA.

Kollock, P. (1999): The Production of Trust in Online Markets. In: Advances in Group Processes (Vol. 16), Lawler, E.J.; Macy, M.;Thyne, S.; Walker, H.A. (eds.), JAI Press.

Kormann, D. P.; Rubin, A. D. (2000): Risks of the Passport Single Signon Protocol. IEEE Computer Networks, Vol. 33.

Lechner, U. (2002): Peer-to-Peer beyond Filesharing. In Second Conference on Innovative Internet Computing Systems, H. Unger, T. Boehme, and A.Mikler, editors, Lecture Notes in Computer Science 2346, pp. 153-162. Springer-Verlag.

Liberty (2002): Liberty Architecture Overview – Version 1.0, Technical Report, Liberty Alliance Project, Jul. 2002.

Mulligan, D.; Schwartz, A. (2000): Your place or mine? Privacy Concerns and Solutions for Server and Client-side Storage of Personal Information, Proc. Computers, Freedom and Privacy, Toronto, ON, Canada.

P3P (2000): The Platform for Privacy Preferences 1.0 (P3P1.0) Specification, World Wide Web Consortium (W3C) Candidate Recommendation, Dec. 2000.

Schubert, P. (1999): Virtuelle Transaktionsgemeinschaften im Electronic Commerce: Management, Marketing und Soziale Umwelt, Lohmar - Köln: Josef Eul Verlag.

Schubert, P.; Koch, M. (2002): The Power of Personalization: Customer Collaboration and Virtual Communities, In: Proc. Americas Conference on Information Systems (AMCIS2002), Dallas, TX, pp. 1953 – 1965.

Schubert, P.; Leimstoll, U. (2004): Personalization of E-Commerce Applications in SMEs: Conclusions from an Empirical Study in Switzerland, in: Journal of Electronic Commerce in Organizations (JECO), 2 (3), July-Sept 2004, pp. 22-40.

Steiner, J.G.; Neuman, B.C.; Schiller, J.I. (1988): Kerberos: An Authentication Service for Open Network Systems. Proc. Winter 1988 Usenix Conference.

Sun (2002a): Strategic Implications of Network Identity, Technical Report, Sun Microsystems, Palo Alto, CA, wwws.sun.com/software/sunone/wp-identity.pdf.

Sun (2002b): How to Implement Network Identity, Technical Report, Sun Microsystems, Palo Alto, CA, wwws.sun.com/software/sunone/wp-implement_ni.pdf.

Webster (1988): Webster's New World Dictionary of American English, Third College Edition, Cleveland : Webster's New World.

Wörndl W. (2003): Privatheit bei dezentraler Verwaltung von Benutzerprofilen. PhD Thesis, Department of Informatics, Technische Universitaet Muenchen, Germany, Aug. 2003.

Wörndl, W.; Koch, M. (2003): Privacy in Distributed User Profile Management. Proc. 12th Intl. World Wide Web Conference (WWW2003), Budapest, Hungary, May 2003.

Wörndl, W. (2002): Using P3P to Negotiate Access Rights to User Profiles. Proc. Work-shop on the Future of P3P, Washington DC, Nov. 2002.