COMPUTERS & SECURITY 29 (2010) 580–591

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Collaborative privacy management

Jan Kolter [a,*], *Thomas Kernchen* [b], *Günther Pernul* [a]

[a] *Department of Information Systems, University of Regensburg, D-93040 Regensburg, Germany*
[b] *Steria Mummert Consulting AG, Französische Str. 48, D-10117 Berlin, Germany*

ARTICLE INFO

Article history:
Received 22 September 2009
Received in revised form
9 December 2009
Accepted 10 December 2009

Keywords:
Privacy
Privacy infrastructure
Privacy-enhancing technologies
Collaboration
Usability

ABSTRACT

The landscape of the World Wide Web with all its versatile services heavily relies on the disclosure of private user information. Unfortunately, the growing amount of personal data collected by service providers poses a significant privacy threat for Internet users. Targeting growing privacy concerns of users, privacy-enhancing technologies emerged. One goal of these technologies is the provision of tools that facilitate a more informative decision about personal data disclosures. A famous PET representative is the PRIME project that aims for a holistic privacy-enhancing identity management system. However, approaches like the PRIME privacy architecture require service providers to change their server infrastructure and add specific privacy-enhancing components. In the near future, service providers are not expected to alter internal processes. Addressing the dependency on service providers, this paper introduces a user-centric privacy architecture that enables the provider-independent protection of personal data. A central component of the proposed privacy infrastructure is an online privacy community, which facilitates the open exchange of privacy-related information about service providers. We characterize the benefits and the potentials of our proposed solution and evaluate a prototypical implementation.

1. Introduction

Today's rich service offer in the World Wide Web increasingly requires the disclosure of personal user data, which poses a growing privacy threat to Internet users. Web site providers utilize these personal data to create and analyze profiles or to trigger personalized advertisements. At the worst, personal information is released or sold to third parties.

Motivated by users who needed technical means to protect their private data, privacy-enhancing technologies emerged (Burkert, 1997; Goldberg and Wagner, 1997). A frequently discussed subject in this area is anonymity on network level. On application level, privacy-enhancing technologies aim for solutions that assist users in controlling and managing the disclosure of personal data. Unfortunately, most approaches rely on the cooperation of service providers who are required to reveal their data handling practices truthfully.

The goal of this paper is the introduction of a collaborative privacy community that facilitates a service provider-independent privacy management. We propose a user-centric privacy architecture and show the functions and the potentials of an inherent collaborative privacy community. Finally, we present a prototypical implementation of our solution.

The remainder of this paper is structured as follows. After describing related work in Section 2, we present an overview as well as the components of a user-centric privacy architecture in Section 3. Section 4 introduces the content, functions as well as the implementation and evaluation of our

* *Corresponding author.*
E-mail address: jan.kolter@wiwi.uni-regensburg.de (J. Kolter).
0167-4048/$ – see front matter
doi:10.1016/j.cose.2009.12.007

collaborative privacy community. Section 5 provides an agenda for the launch of the proposed privacy community, before the main contributions of this paper are highlighted in Section 6.

2. Related work

The Platform for Privacy Preferences (P3P) (Cranor et al., 2006a) represents an early privacy-enhancing technology framework aiming at aiding users in decisions regarding the disclosure of their personal data. Offering an appropriate policy language, P3P gives service providers the opportunity to express their privacy policy in a machine-readable format. When the user visits a Web site, a dedicated P3P privacy agent matches the P3P privacy policy of the service provider with pre-defined disclosure rules (so called privacy preferences) of the user. The matching process results in a recommended disclosure behavior, which is signaled to the user.

In addition to its frequently cited weaknesses (Electronic Privacy Information Center, 2000; Hogben et al., 2002), a crucial factor that prevents the widespread use of P3P is the lagging adoption of P3P privacy policies, which are offered by only a small fraction of service providers (Reay et al., 2007).

Aiming to support users' ability to maintain their privacy, the European PRIME project[1] (Privacy and Identity Management for Europe) developed a privacy-enhancing identity management system, containing a privacy architecture with different design guidelines, protocols and prototypical scenarios (Leenes et al., 2008).

The PRIME architecture allows users to control the disclosure and the usage of their personal data (Leenes et al., 2008; Sommer et al., 2008). A significant element of the architecture is the PRIME Toolbox, which needs to be installed both on the client side and the provider side. The PRIME Toolbox incorporates all necessary components for privacy-enhancing identity management and enables users to manage and use multiple digital identities with varying personal data.

An additional element of the PRIME architecture is the PRIME Middleware that integrates all PRIME components and coordinates the communication between PRIME interaction parties. The PRIME console serves as a graphical interface enabling users to define privacy-related preferences that are used to negotiate data handling practices with service providers. Furthermore, an overview of already disclosed data is provided. The architecture is capable of enforcing negotiated policies, utilizing the installed PRIME components at the service provider side.

In order to make use of the described PRIME functionality, both users and service providers need to install the PRIME Middleware and the PRIME Toolbox. From a user perspective the attractiveness of PRIME rises, if the majority of service providers adapt their service infrastructure. Hence, the success of PRIME highly relies on the service providers' willingness to integrate the described PRIME components into their applications.

[1] https://www.prime-project.eu/.

3. User-centric privacy architecture

In the previous section we proved that existing privacy solutions seeking to protect personal user data strongly rely on the cooperation of service providers. This dependency, however, is responsible for the low practical applicability of many promising solutions. Even though threats of personal data misuse are growing, the example of the P3P specification shows that service providers do not contribute to the widespread availability of accurate P3P policies voluntarily. Likewise, from today's perspective it seems unlikely that service providers will fundamentally change their internal back-end infrastructures, as required for realizing the ideas of the PRIME project.

Acknowledging the conflicting interests of service providers as well as the need for usable tools, we introduce a user-centric, service provider-independent privacy architecture (Kolter et al., 2009), which is depicted in Fig. 1. A collaborative privacy community facilitates Internet users to share privacy-related information about service providers. The community is maintained by all participating members. Three privacy components on the user side offer user-friendly tools that assist users in controlling potential, actual and past information flows, utilizing service provider information of the privacy community.

Unlike provider-dependent privacy technologies our proposed privacy architecture does not require the direct support of service providers. We rather accept today's service landscape of the World Wide Web and offer a more practical privacy infrastructure.

The user is supported by a browser plug-in, which serves as user interface of the privacy architecture. The browser plug-in displays privacy-related information and functions, which are provided by three local privacy components. The Privacy Preference Generator component assists users in controlling potential information flows of personal data, while the Privacy Agent component helps users check and control actual information flows. Finally, the Data Disclosure Log provides an overview of past personal information flows. All local privacy components interact with the collaborative privacy community.

In the following, we briefly specify the main purposes and functions of the local privacy component, before the collaborative privacy community is introduced in Section 4.

3.1. *Privacy preference generator*

Enabling users to control potential personal data flows, the proposed privacy architecture provides a Privacy Preference Generator component that captures individual privacy preferences. Privacy preferences define individual conditions of personal data disclosures. As the Privacy Agent component matches privacy preferences with a service provider's privacy policy, the resulting recommendation highly depends on the accuracy of individual preferences.

Catering the needs of predominantly inexperienced users, our solution offers a tool that allows users to define privacy preferences in a user-understandable way (Kolter and Pernul,

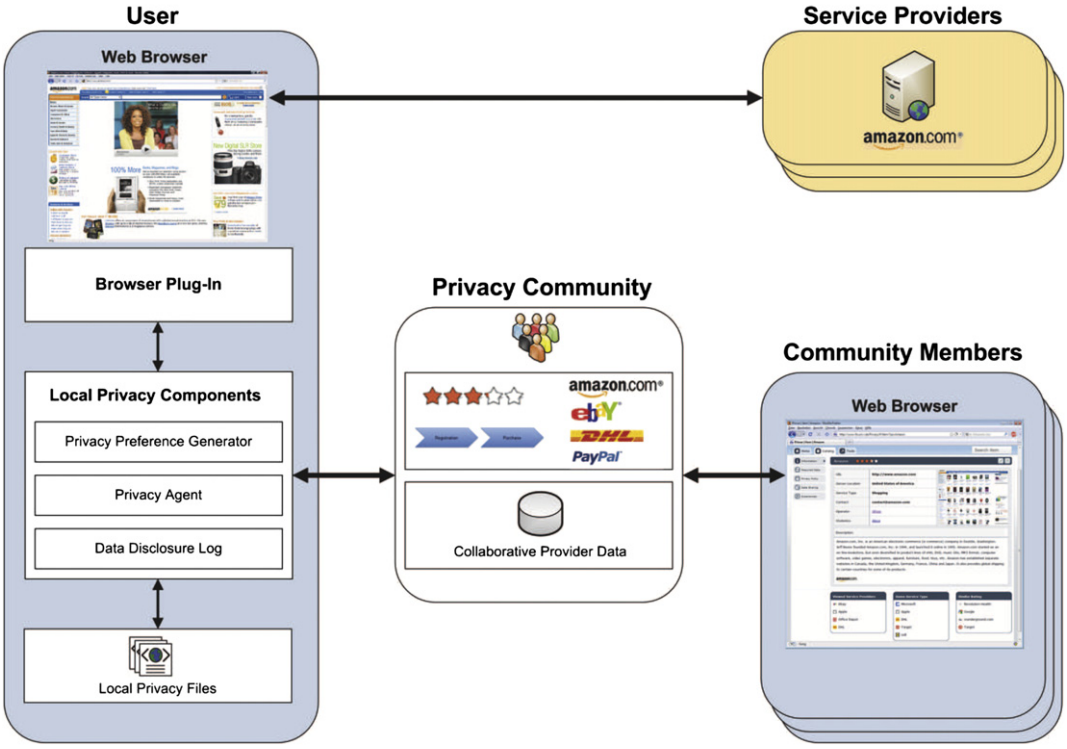

Fig. 1 – Collaborative, provider-independent privacy architecture.

2009). We facilitate the individual definition of privacy preferences for pre-defined Internet service types, guaranteeing more realistic and practical results. In addition to a user-friendly configuration wizard, our solution provides a clear configuration summary and evaluation.

3.2. *Privacy agent*

The Privacy Agent component assists users in making an informed decision about the actual disclosure of personal data. With regard to the visited Web site, this support involves the presentation of relevant information of the privacy community and the matching of a published P3P privacy policy with user's pre-defined privacy preferences. Returning users benefit from information about linkable partial identities and already disclosed personal data.

3.3. *Data disclosure log*

The Data Disclosure Log component records personal data transfers and provides a clear overview of past personal data flows. Such an overview enables users to know the recipients of past personal data transactions at any time (Pettersson et al., 2006). This knowledge, for instance, represents a prerequisite for an ex post revocation of personal data. The Data Disclosure Log component requires both a tracking tool that monitors personal data disclosures in the Web browser as well as usable interfaces that illustrate logged data transactions in a comprehensible way. Ideally, the disclosure log allows users to directly access, change or remove disclosed personal data stored by a service provider.

4. Collaborative privacy community

The privacy community marks the central element of our proposed privacy architecture and enables users to collaboratively exchange privacy-relevant information, ratings and experiences about service providers (Kolter et al., 2009). This information includes, for instance, the required amount of personal data for the fulfillment of a service and third parties the provider shares personal user data with. A collaborative privacy community represents a valuable data source for all three local privacy components and facilitates a provider-independent privacy protection.

All service provider information is maintained and organized in a Wiki-like Web front-end (Leuf and Cunningham, 2001). For each service provider privacy-related information is grouped into an article, allowing all Internet users to view and edit articles in the Web browser. In addition, open Web service interfaces allow the flexible integration of privacy-related service provider information into the local privacy components.

4.1. *Contents and functions*

This section defines the provided content and functions of our proposed privacy community. The selection was primarily driven by the needs of the local privacy components, but also by the potential of a provider-independent privacy environment.

4.1.1. *Static information about service providers*

When an unknown Web site is visited, users generally have the option to trust a service provider at face value or to look for

information about its reputation and data handling practices. A survey shows that many users do not look up reputational information, but rather judge service providers' trustworthiness by estimating the Web site's "Look and Feel", considering questionable factors (Fogg et al., 2001). As collecting information about a service provider is time-consuming, this behavior of especially inexperienced users seems understandable. Addressing this fact, the privacy community provides users with an easily available overview of static service provider data.

This general information is primarily utilized by the local Privacy Agent component and is displayed on demand, enabling users to easily get a first impression of a service provider.

In particular, the privacy community offers the following static service provider information:

- The service provider's URL
- The physical location of the server
- The offered service type
- Information about the revocation of already transferred personal data
- Contact information
- A short textual description of the service provider

The URL is required to uniquely identify service providers. The server location clarifies legal matters, as different privacy laws apply in different countries. The service provider's service type is queried by the Privacy Agent component. This capability facilitates the individual generation of privacy preferences for each service type and their application during policy matching.

In addition, the privacy community offers information (e.g. a link or an e-mail address) about the removal of disclosed personal data that have been transferred to a service provider. This knowledge helps users exercise their rights to control already transferred data and is utilized by the Data Disclosure Log component. Exact contact information facilitates prosecution, if personal information is misused, or if users want to enforce their rights to revoke their personal data. Furthermore, a short textual description specifies the main characteristics of a service provider.

4.1.2. *Required amount of personal data*

In addition to static provider information the privacy community enables users to know in advance what personal data are requested by a certain service in the World Wide Web.

Users generally understand the necessity to disclose, for example, name, address and payment information for a product order at an online shop. If the service provider asks for additional information, such as the marital status, the date-of-birth or the annual salary, users tend to abort the process, if they feel uncomfortable releasing these excessive data. An online survey we conducted with 350 persons revealed that 77% of all test persons cancel registration and purchasing processes if too much personal information is requested. Unfortunately, with today's technical means users are unable to determine in advance, what personal information is necessary to use a specific service. In an effort to find out, users have to start the process of filling a set of Web forms. In many cases the most privacy-sensitive information is requested on the last form page. If the user decides not to proceed, he/she wasted valuable time and disclosed the already transferred information with no use.

The introduced privacy community spares users from this negative experience and enables them to exchange the amount and type of personal data required for each process a service provider offers. In this context, a process refers to each separate service offer, such as *Purchase* or *Newsletter Subscription*. In addition, the community stores the reliance of a process on the completion of a different process. The process *Purchase* could, for instance, require the completion of the process *Registration*. In addition, the privacy community performs an automatic evaluation of the required personal data with regard to the offered process and service type, which further assists users in assessing personal data requests of service providers. The amount of required personal data as well as the results of the automatic evaluation is employed and displayed to the user by the local Privacy Agent component.

The amount of required personal data represents a fundamental element of privacy policies. Its online availability in a privacy community facilitates the local Privacy Agent component to retrieve this information and match it with individual privacy preferences, if no sufficient machine-readable privacy policy is offered by a service provider.

4.1.3. *Third party recipients*

The decision to disclose personal information to a service provider not only relies on the amount of data, but also on the service provider's data handling practices. Here, the forwarding of user data to third parties is a considerably privacy-sensitive factor.

While the P3P specification only defines third party categories, the proposed privacy community allows the exchange of individual third parties the service provider shares personal data with. These parties can include affiliated companies and other business partnerships. This information is displayed to the user by the local Privacy Agent component on demand. Again, information about third party releases can be utilized to replace a machine-readable privacy policy of the service provider.

4.1.4. *Collecting and explaining privacy policies*

In general, a service provider's textual privacy policy is the only available information source about its data handling practices. Studies show, however, that privacy policies are not regarded as understandable and are read by only a small fraction of Internet users (Jensen et al., 2005; Pollach, 2007).

Addressing the needs of the majority of Internet users, the privacy community allows experienced users to write and share an understandable explanation of a provider's privacy policy. As privacy experts comprehend all aspects of a policy, they have the ability to paraphrase important elements in a form that – compared to a published policy as well as automatic privacy policy summaries (Arshad, 2004; Cranor et al., 2006b) – is easy to understand.

Furthermore, as privacy policies change over time, the privacy community maintains a history of privacy policies, containing both textual policies as well as machine-readable P3P policies. Such a policy history enables users to determine

the policy that has been valid, when personal data have been disclosed. The privacy community also allows users to rate current and past privacy policies of service providers with regard to the stated data handling practices.

4.1.5. *Adherence to privacy policies*

As the presence of a privacy-friendly privacy policy is no guarantee that a service provider follows that expressed policy, the presented privacy community enables users to rate the policy adherence of service providers. Based on their individual experiences users evaluate, whether or not a service provider processes personal data as stated in the privacy policy. For example, if not expressed in the privacy policy, a received e-mail that promotes a product would justify a negative policy adherence rating of that service provider. Displayed by the local Privacy Agent component, this information is of considerable importance for a disclosure decision.

4.1.6. *Individual experiences*

Finally, the offered service provider information is complemented by individual user experiences. These open postings can contain any privacy-related positive or negative experiences and are not related to a specific aspect of the provider's data handling practices.

Integrated into the Privacy Agent component, the individual experiences are utilized for the presentation of reputational information about a service provider.

4.1.7. *Sharing privacy preferences with connected users*

In Section 3.1 we pointed out the purpose and usage of individual privacy preferences. The Privacy Preference Generator component allows the definition of these disclosure rules, which are in turn used by the Privacy Agent component to calculate disclosure recommendations. The quality of these recommendations strongly relies on the accuracy of privacy preferences. Even though the Privacy Preference Generator component should alleviate this challenge by offering a usable and understandable user interface, building accurate privacy preferences is a critical task. This especially applies to inexperienced users, as they are not familiar with service providers' data handling practices and the used privacy-related language.

For this reason, the privacy community facilitates the exchange of privacy preferences among users. Using an integrated social networking component (Boyd et al., 2007), users have the option to upload privacy preferences and share them with selected members.

Imported privacy preferences of a trusted privacy expert or organization represent valuable assistance for inexperienced users, resulting in improved disclosure recommendations of the local Privacy Agent component.

4.2. *User management*

The internal user management of the privacy community administers three user roles. Offering an open information source, the basic user role is assigned to every unregistered user and grants access to all available information about service providers. Furthermore, it permits users to edit articles collaboratively and to create new service provider articles. In order to prevent vandalism, the privacy community provides adequate backup and versioning functionality.

If users want to directly exchange information with connected members, a simple registration is necessary. Registration only requires a username and a password. The community does not request any additional personal user information. Unlike basic users, registered users have the option to upload and share generated privacy preferences with connected members. Likewise, privacy preferences of connected members can be downloaded and imported into the Privacy Preference Generator. We point out that the involved social networking component does not have the purpose of maintaining social contacts, but only to exchange privacy experiences and privacy preferences. Users can self-assess their level of knowledge and experience, helping inexperienced users to estimate the quality of advises and preferences.

Finally, users holding the administrator role specify available processes as well as the appropriate amount of personal data for each process. If necessary, administrators are able to block users.

4.3. *Prototype*

We implemented a prototype of our proposed privacy community. In the following, we present the prototype's system architecture, the used frameworks as well as the graphical user interface.

4.3.1. *System architecture*

As mentioned earlier, the Web front-end and the local privacy components on the user side simultaneously access the community. The integration of heterogeneous client applications requires the specification of standardized interfaces, which is ideally realized by a Service-oriented Architecture (SOA) (MacKenzie et al., 2006). Implementing the concept of a SOA, the privacy community encapsulates the offered information pieces and actions into fine-grained Web services. Each Web service provides a machine-readable WSDL (Chinnici et al., 2007b,a) service definition, which clearly defines its interface. The communication of the privacy community with its clients via SOAP messages (Gudgin et al., 2007a,b) guarantees a consistent data exchange format. Fig. 2 shows the privacy community's interaction with the components of the privacy architecture.

For the community's Web front-end we utilize an Ajax (Garrett, 2005) Web architecture, allowing asynchronous, interactive communications between the Web front-end and the community server. The Ajax engine transforms JavaScript (Flanagan, 2006) requests of the user into SOAP requests, which are forwarded to the community back-end on the server side. A Web service server receives and processes requests querying the provider database, before requested data are sent back to the client via SOAP. The Ajax engine of the Web front-end transforms these SOAP messages to a user-friendly GUI using HTML (Raggett et al., 1999) and CSS (Bos et al., 2009).

The local privacy components – the Privacy Preference Generator (PPG), the Privacy Agent (PA) and the Data

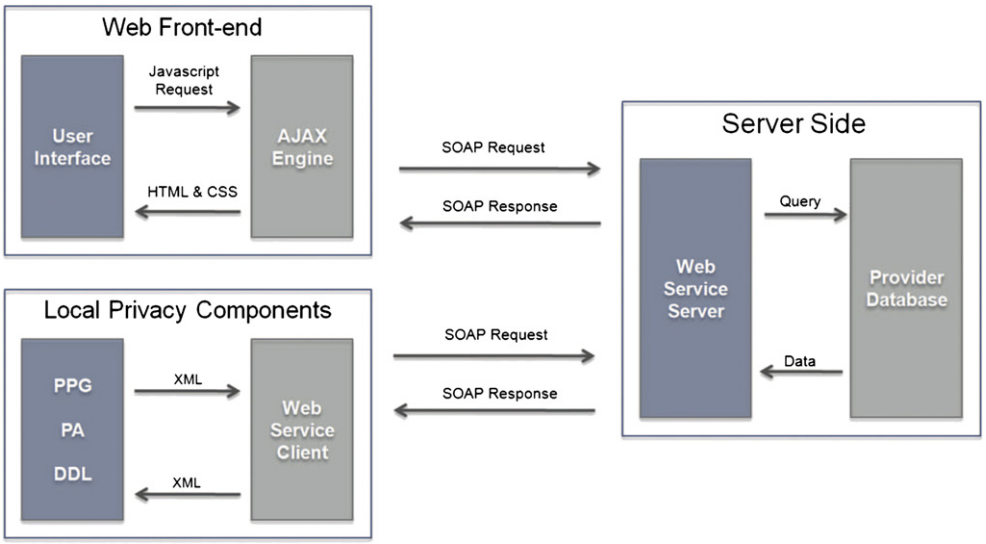

Fig. 2 – Interaction of the privacy community.

Disclosure Log (DDL) components – directly access the Web service server via SOAP messages.

4.3.2. *Implementation details*

The back-end of the privacy community employs NuSOAP,[2] a PHP-based SOAP toolkit that provides the required service functionality for our proposed privacy infrastructure. The Web service interface definitions can be accessed following this link.[3]

The SOAP server accesses a MySQL database,[4] which stores all service provider information as well as the presented data type and service type vocabularies.

For the Web front-end we utilize the Web application framework CodeIgniter.[5] The PHP-based framework facilitates the MVC-compliant development of dynamic Web applications and allows for the smooth integration of the back-end Web services. On the client side, the JavaScript framework jQuery[6] offers AJAX and dynamic HTML technologies, which provide necessary drag & drop and auto-complete functions, overlays, as well as high performance DOM parsing.

4.3.3. *Graphical user interface*

The designed Web front-end aims for a clear layout and a high degree of user-friendliness. With regard to the assigned role users are able to look up, view and edit service provider articles, register, login and share privacy preferences with selected members, and administer users and structural data of the community. In the following, we focus on the design of the service provider catalog and a service provider article. For a complete review of the graphical user interface, the interested reader is referred to the prototypical implementation of the privacy community, which is accessible at the following link.[7]

The welcome page shortly explains the purpose and the content of the privacy community and its related local privacy components. From this starting page, the user has the option to enter the catalog page, which lists all service providers maintained in the privacy community (see Fig. 3). Service providers are represented by tiles that contain a large provider logo, contributing to an easy association of the underlying article. In addition to the provider logo, a calculated average privacy rating is shown at the bottom of each tile. The rating is presented as star rating whose interpretation and usage is familiar to most users. Selections at the left side allow users to filter service providers based on their service type and their average privacy rating. Alternatively, users can type a service provider name in the search field at the top right of the page. Here, an auto-complete function eases the correct article selection. A third page provides a detailed presentation of the local privacy components and offers the download of an installer.

If a service provider is not listed in the privacy community, users can create a new service provider article at any time. The optional registration and login functionality is offered at the header of each page.

The privacy-related content of service provider articles is divided into five tabs. In the following, we present the community article of the eCommerce provider Amazon.[8]

The initial tab offers a quick overview of the service provider (see Fig. 4), including the service provider's name, its average rating, a dynamically generated screenshot of its current Web site as well as privacy contact information. In addition, a general textual description of the service provider is presented. Moving the mouse over the average rating in the tab header triggers an overlay that lists the individual star ratings of each subcategory. At the bottom of the page three

[2] http://sourceforge.net/projects/nusoap/.
[3] http://www-ifs.uni-regensburg.de/Privacy/soap_ws/.
[4] http://www.mysql.com/.
[5] http://codeigniter.com/.
[6] http://jquery.com/.
[7] http://www-ifs.uni-regensburg.de/Privacy/.
[8] http://www.amazon.com/.

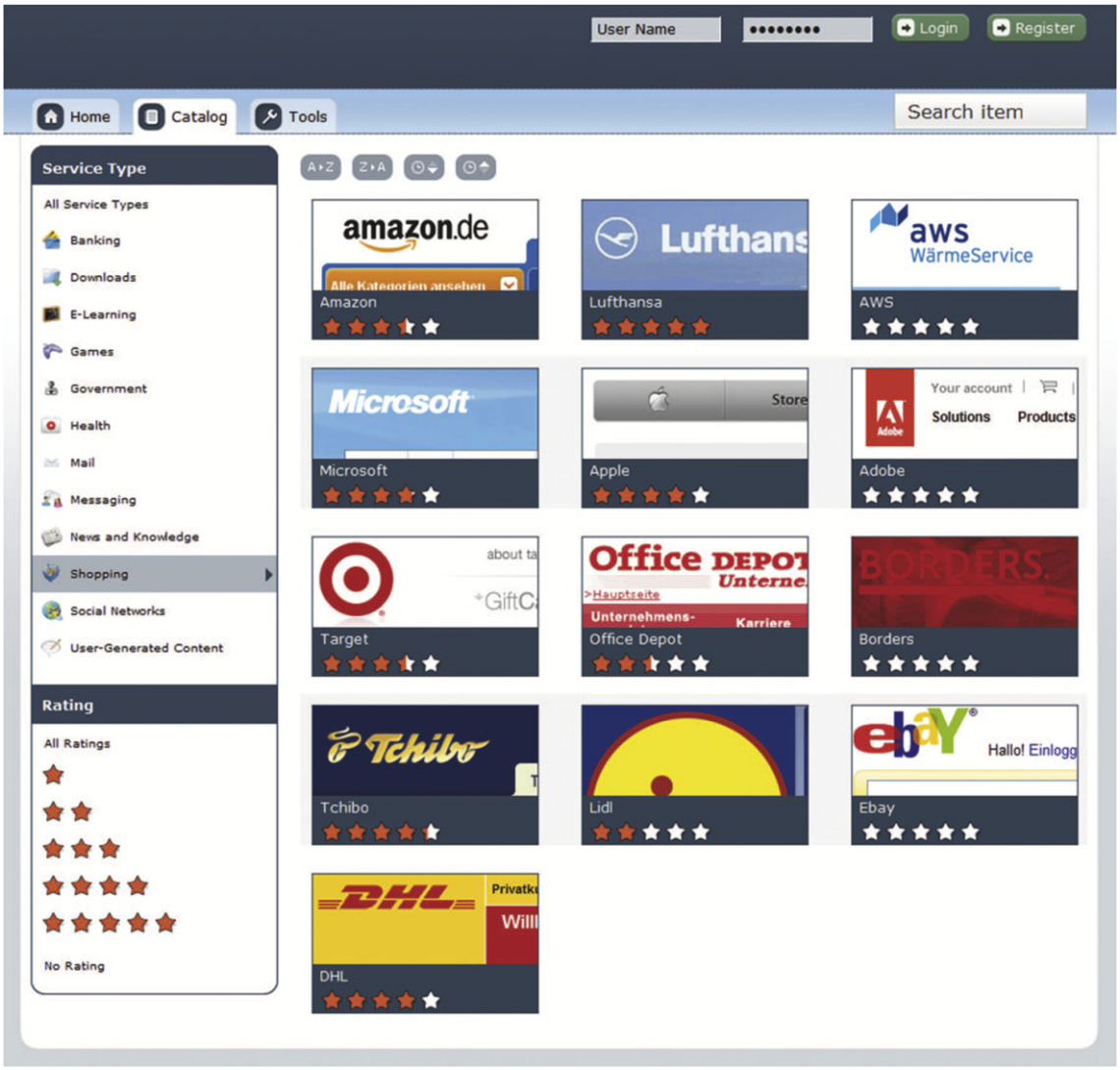

Fig. 3 – Service provider catalog.

lists contain the most recently visited service providers, randomly selected service providers of the same service type and service providers with the most similar average privacy ratings. These article links allow for a quick comparison of similar service providers, facilitating an effective evaluation of service providers' privacy practices. The tab header and the three lists at the bottom of the page are presented in all remaining article tabs.

The second tab shows the required amount of personal data for all offered processes, which are represented by green arrows (see Fig. 5). If the user clicks an arrow, a text box unfolds containing the required personal data elements. In our example, the process *Purchase* is selected, which requires an *Address*, a *Phone Number*, *Payment Information* and an *E-mail Address*. The arrangement of the process arrows indicates that the *Purchase* process requires the completion of the *Registration* process that may require additional personal data. Based on comparable *Purchase* processes, the community evaluates the amount of required personal data as negative.

Concentrating on the data handling practices of the service provider, the fourth tab lists a privacy policy history – including the currently valid policy – along with a star rating of each policy (see Fig. 6). The history is capable of storing textual and P3P versions of privacy policies. Below the list of privacy policies, a short explanation of the effective privacy policy is offered to inexperienced users. At the bottom of the tab, users can rate the policy adherence of the service provider. Upcoming releases of the privacy community will offer multilangual storage of policy explanations, enabling a broader group of users to benefit from that provided information.

Focusing on personal data sharing, the fourth tab employs a directed graph to visualize third parties the service provider shares personal user data with (see Fig. 7). Originating from the examined service provider, arrows point to additional data recipients, which are represented by white boxes that contain their names and Favicons. If the user clicks a data recipient, the article of the respective provider is loaded.

Finally, a fifth tab lists posted user comments about privacy-related experiences with the service provider.

Edit buttons in the tab headers facilitate the addition and the revision of the collaboratively maintained information. If

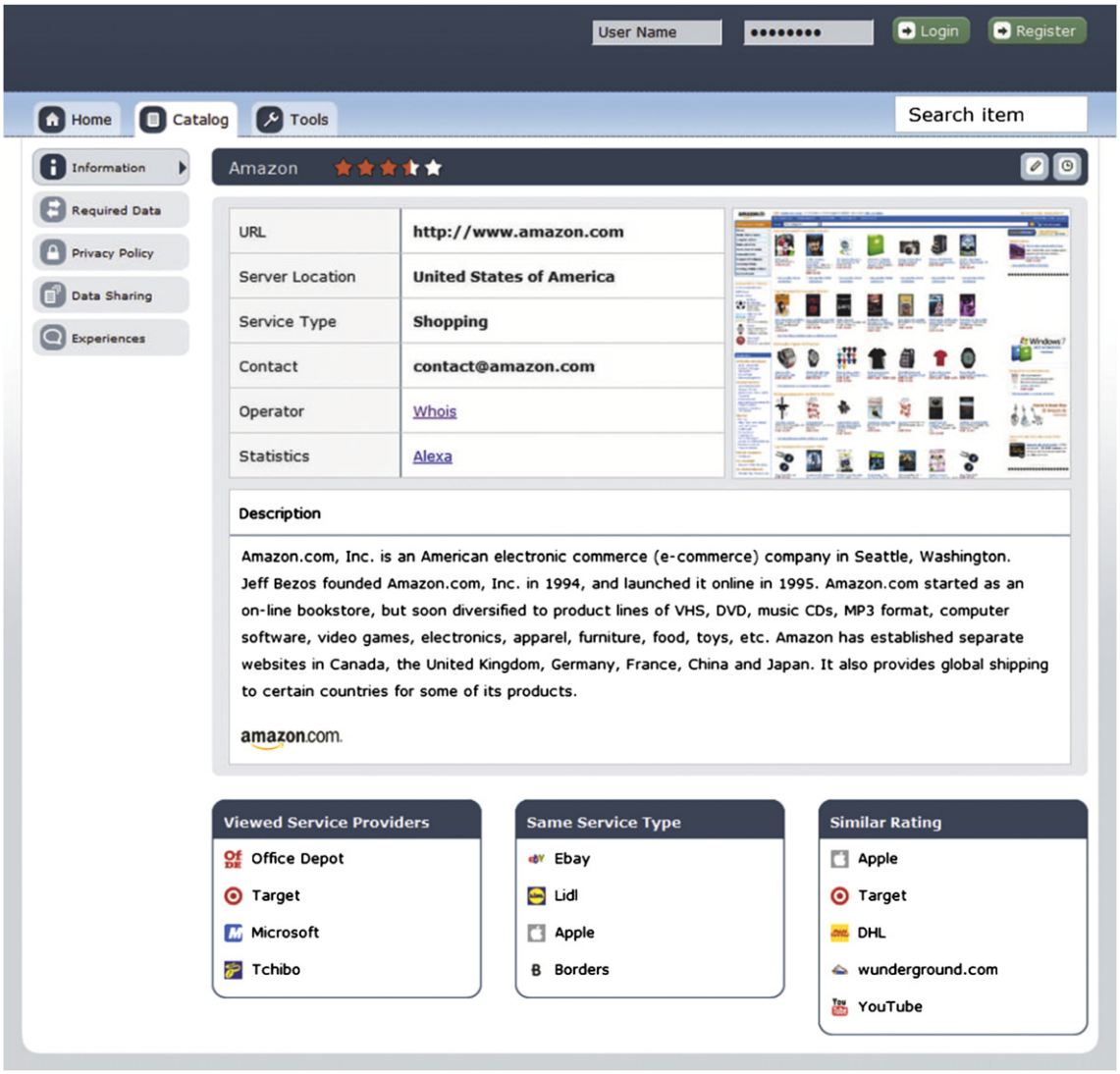

Fig. 4 – Article – static provider information.

clicked, overlays capture the revised user input, using text boxes and drag and drop selections. Also placed in the tab header, a Versions button allows the recovery of older revisions.

4.4. *Evaluation*

In order to evaluate the design and the structure of the community Web front-end, we conducted a user test that assessed usability, user acceptance and the potential user participation of the collaborative privacy community.

For the user experiment we recruited 26 test persons, acknowledging frequent recommendations that a single-digit sample is insufficient for a user test (Faulkner, 2003; Perfetti and Landesmann, 2001; Spool and Schroeder, 2001). Aiming at a heterogeneous test sample, the invited test persons showed a diverse academic and professional background. However, basic knowledge of Microsoft Windows as well as the occasional use of the World Wide Web were prerequisites for participating candidates. In order to avoid biased results, persons with close relationships to the interviewers were not considered.

In particular, the test sample included 17 university students, while nine test persons were graduated professionals. Hence, 15 out of the 26 test persons were 25 years old or younger, seven between 26 and 30, and four between 30 and 45. 22 of all test persons were male. Out of the 17 students nine were enrolled in a technical program and five in a business program. From the remaining students two were pursuing a teaching degree and one a diploma in mathematics.

In order to measure the unbiased understandability of the page layout, test persons were only informed about the general purpose of the community. No detailed explanation of the Web page was provided. Before the first assignment, test persons had the opportunity to get familiar with the structure and the tabs of the Web page.

The first task targeted the submission of a provided privacy posting about the eCommerce shop Tchibo.[9] This task required the search for the proper service provider in the community using the offered navigation elements. Doing that, more than half of the test persons used the search field, which

[9] http://www.tchibo.com.

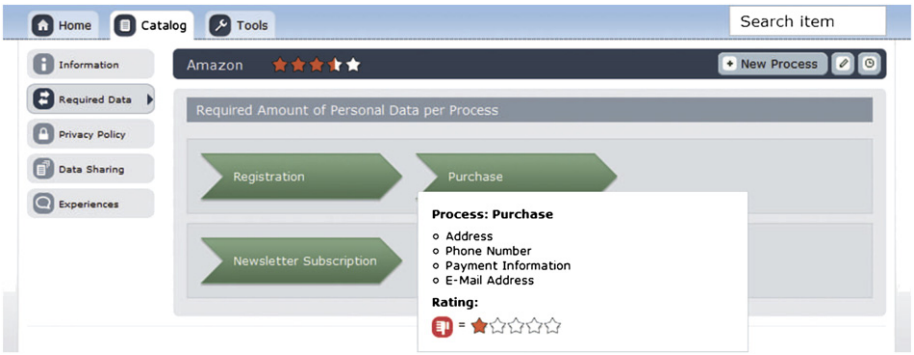

Fig. 5 – Article – required amount of personal data.

is placed in the page header, while the remaining test persons clicked on the index tab, an early version of the catalog page, which presented a textual listing of all available service providers. The outcome and the user feedback led us to offer a more meaningful catalog page, which contributes to more intuitive page navigation. Nevertheless, the test results underscored the necessity of offering both a service provider catalog and a search field. Once the proper article was loaded, all test persons clicked the correct tab and posted a comment with no difficulties.

Within the following task, we asked test persons to post a particular rating for the policy adherence of the same service provider. Again, most test persons found the relevant tab and succeeded smoothly. Four out of 26 test persons, however, were not familiar with the handling of star ratings and needed guidance of the interviewers. Being asked after the completion of the task, 17 test persons were able to explain the difference between the ranking of the privacy policy itself and the policy adherence ranking, while nine persons could not differentiate both ratings. This fact highlighted the need of explaining labels, which were added in the latest revision of the community.

Subsequently, test persons were asked to publish a set of data types required by a *Purchase* process at Tchibo. Only nine test persons solved that task problem-free. The majority of test persons did not intuitively find the position of the Edit and Save buttons, which led us to reallocate their positions.

In the interview section 22 out of 26 test persons agreed that the information of a service provider article was structured and designed in an understandable way. 24 out of 26 test persons would consult and use the privacy community in real-life scenarios. Being asked about the reliance of the community data, 18 test persons stated they would trust the privacy community, once the community gained enough members. The remaining eight test persons voiced their concern about the openness of the community, which – in theory – allows service providers to manipulate community data to their favor. One of these test persons admitted that a rising level of popularity and submitted data would alleviate that threat, as proven by well-known open reviewing systems of Ebay and Amazon.

17 test persons agreed that they would actively participate in editing articles of the privacy community. The remaining nine test persons would not add or change community data, hinting at their general reluctance to post content in public forums and other open content management systems.

A few test persons suggested the incorporation of other non-privacy-related information like delivery time or shipping costs. In our opinion, however, these data would contradict to the goals of the privacy community and would not contribute to its reputation of a provider-independent information source.

Finally, the user test revealed that 15 out of 26 test persons would upload and share their individual privacy preferences with selected users of the privacy community.

5. Agenda

This section outlines criteria for the successful launch as well as the long-term financing of an online privacy community.

As mentioned earlier, the idea of a collaborative privacy community is primarily based on the successful Wikipedia concept (Leuf and Cunningham, 2001). Consequently, the launch of a privacy community can be related to success factors for the launch of a general ''Wiki''.

In order to achieve the critical number of participating users, the successful launch of a privacy community requires the full exploitation of the Wiki effect. The Wiki effect is defined as a large number of Internet users visiting a Web site on a regular basis and voluntarily contributing to the structure, shape and quality of its content (Ebersbach and Glaser, 2007). In order to benefit from the Wiki effect, Davies (2004) recommends selectively seeding initial content that introduces the Wiki's goal to new participants. The character of this initial content should not be final or complete, as this could prevent users from editing that content.

Hence, before the launch of a privacy community, articles of the best-known service providers should be created and filled with sparse information. In particular, we recommend entering the offered service type, the amount of required personal data and the release of personal data to third parties, as this information represents essential input for the local privacy components.

Addressing the promotion of a privacy community, comments in public forums and other public communication channels contribute to a Wiki's level of popularity (Parry, 2006). In addition, a press release is recommended, which should be forwarded to authors of blogs and other topic-related Web sites. Furthermore, the creation of a clear tutorial is suggested that outlines the community's main goals and functionality. A

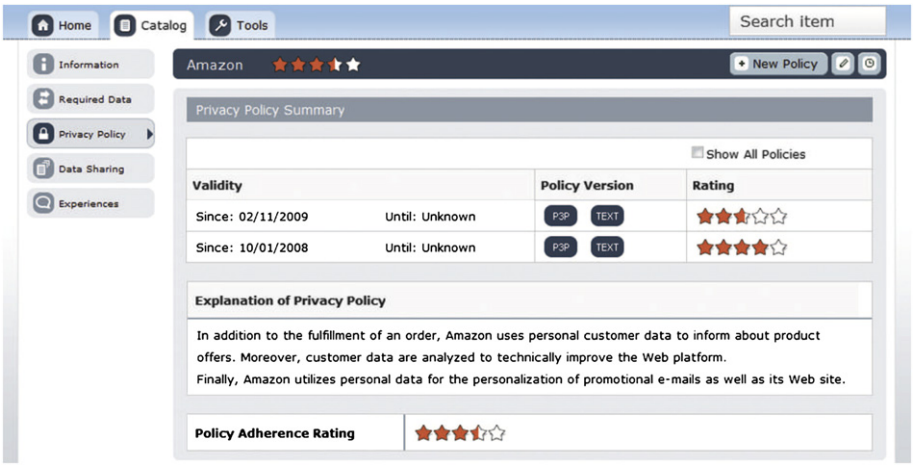

Fig. 6 – Article – privacy policy summary.

tutorial enables new users to gain a quick overview and increases the chances that users contribute actively.

Finally, crucial success factors of the privacy community are security and trust considerations. In theory, the offered platform is capable of tracking the information requests of users. As local privacy components query data from the privacy community on a regular basis, a malicious community provider could create and misuse detailed navigation logs. Our proposed privacy community acknowledges this potential privacy threat and allows the anonymous accessing and editing of all collaboratively maintained data. When users have built the necessary level of trust in the community, they can opt to register, which allows them to connect to friends and to exchange experiences and privacy preferences.

In addition, an attacker could place untruthful information in articles of the privacy community. Addressing this threat, we store the IP address of anonymous postings, which complicates efforts to influence articles. As mentioned in the previous section, a growing number of submitted data for a service provider will further lower the chance of manipulation. As proven in large collaborative platforms, the prevention of misuse will additionally rely on administrators and vigilant users.

After a successful launch and a risen number of users the long-term maintenance of a privacy community inevitably involves expenses, e.g. for the operation of the server infrastructure and, possibly, the entailing maintenance personnel. In the context of Web 2.0 applications, Alby (2007) discusses the potentials of both advertising and a fee-based membership. These sources of financing can also be applied to the maintenance of a privacy community.

Technically there are multiple ways to place advertisements into a privacy community. Considering our proposed solution, advertisement should not be used at the expense of usability and should not affect the clear structure of an article. For the same reason, intrusive advertisements like pop-up windows should be avoided. While dynamic advertisement applications such as Google AdSense provide attractive models for the generation of revenues, we do not recommend the integration of contextual advertisements into a privacy community, as this could result in the placement of a service provider's ad banner in the community article of that provider. Such a behavior does not underscore provider-independence and could weaken users' trust in the privacy community.

A fee-based membership represents a further business model for a privacy community. If this financing source is chosen, the privacy community should provide a basic service offer for free. Special, value-added functions could require a membership involving monthly or annual fees. Specifically,

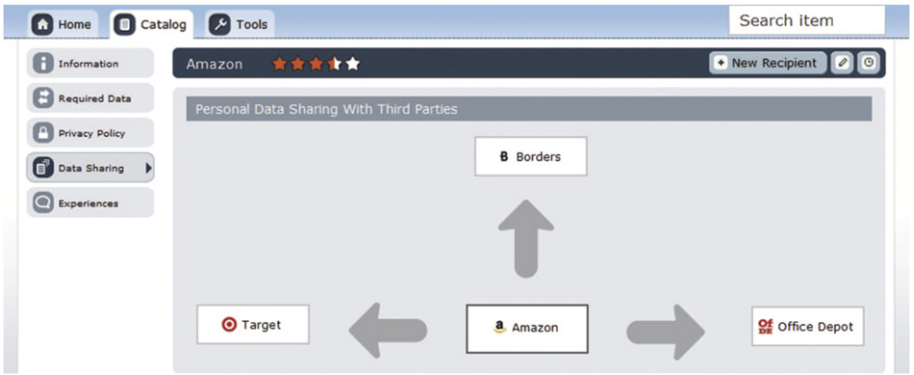

Fig. 7 – Article – third party recipients.

the social networking component that facilitates the direct exchange of privacy preferences and experiences could be defined as a premium function only available for paying members. Premium functions could also involve functionality and certain features of the local privacy components.

Finally, a privacy community could raise donations to cover operating costs, highlighting its service provider-independence. This option would, however, require the abandonment of advertisement. The prominent example of Wikipedia shows that donations can cover operating costs of a large community. Donations are made by satisfied users who are convinced of Wikipedia's goal that knowledge should be accessible to anyone.

The goal of the introduced privacy community is the provider-independent enhancement of privacy. If the increased level of privacy and the improved privacy awareness of both users and providers are recognized, the privacy community could equally convince users of its higher goals and motivate to make donations.

6. Conclusions

Addressing the need for practical technologies that protect personal data disclosures in the World Wide Web, this paper introduces a user-centric privacy architecture that does not depend on the cooperation of service providers. Marking the central element of the underlying privacy architecture, we present a usable privacy community, which facilitates the collaborative exchange of privacy-relevant information and ratings about service providers. Moreover, our developed solution allows users to know in advance, what personal data are required for a specific service. Benefitting from the knowledge of experienced users, the privacy community enables average Internet users to make a more informed decision about the disclosure and management of personal data.

Provider-independence as well as the collaborative character will help foster the usage and acceptance of privacy-enhancing technologies.

Acknowledgement

The authors would like to thank Alfred Kobsa, University of California, Irvine, for helpful comments and stimulating discussions, as well as Stefan Kendlbacher for his support to this work.

REFERENCES

Alby T. Web 2.0-Konzepte, Anwendungen, Technologien. Munich: Carl Hanser Verlag; 2007.

Arshad F. Privacy Fox – a JavaScript-based P3P agent for Mozilla Firefox, Tech. Rep. Pittsburgh, PA 15213, USA: School of Computer Science, Carnegie Mellon University; December 2004.

Bos B, Celik T, Hickson I, Lie HW. Cascading style sheets level 2 revision 1 (CSS 2.1) specification. W3C Candidate Recommendation, World Wide Web Consortium (W3C); April 2009.

Boyd DM, Ellison NB. Social network sites: definition. History, and Scholarship, Journal of Computer-Mediated Communication 2007;13(1):210–30.

Burkert H. Privacy-enhancing technologies: typology, critique, vision. In: Agre P, Rotenberg M, editors. Technology and privacy: the new landscape. Boston, MA, USA: MIT Press; 1997. p. 126–43.

Chinnici R, Haas H, Lewis AA, Moreau J-J, Orchard D, Weerawarana S. Web services description language (WSDL) version 2.0 part 2: adjuncts. W3C Recommendation, World Wide Web Consortium (W3C); June 2007.

Chinnici R, Moreau J-J, Ryman A, Weerawarana S. Web services description language (WSDL) version 2.0 part 1: core language. W3C Recommendation, World Wide Web Consortium (W3C); June 2007.

Cranor L, Dobbs B, Egelman S, Hogben G, Humphrey J, et al. The platform for privacy preferences 1.1 specification. W3C Working Group Note, World Wide Web Consortium (W3C), http://www.w3.org/TR/P3P11/; November 2006.

Cranor L, Guduru P, Arjula M. User interfaces for privacy agents. ACM Transactions on Computer-Human Interaction (TOCHI) 2006;13(2):135–78.

Davies J. Wiki brainstorming and problems with Wiki based collaboration, Project report submitted for the degree of information processing. Department of Computer Science, University of York; September 2004.

Ebersbach M, Glaser R. Heigel, Gemeinsam weben, c't 2007 (14) 168–171.

Electronic Privacy Information Center. Pretty poor privacy: an assessment of P3P and internet privacy, Tech. Rep. Electronic Privacy Information Center, http://epic.org/reports/prettypoorprivacy.html; 2000.

Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments and Computers 2003;35(3): 379–83.

Flanagan D. JavaScript: the definitive guide. 5th ed. O'Reilly Media, Inc.; 2006.

Fogg BJ, Marshall J, Laraki O, Osipovich A, Varma C, Fang N, et al. What makes web sites credible?: a report on a large quantitative study. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '01). New York, NY, USA: ACM; 2001. p. 61–8.

Garrett JJ. Ajax: a new approach to web applications, adaptive path, http://www.adaptivepath.com/ideas/essays/archives/000385.php; February 2005.

Goldberg D, Wagner E. Brewer, privacy-enhancing technologies for the internet. In: Proceedings of the 42nd IEEE international computer conference (COMPCON '97). Washington, DC, USA: IEEE Computer Society; 1997. p. 103–9.

Gudgin M, Hadley M, Mendelsohn N, Moreau J-J, Nielsen HF, Karmarkar A, et al. SOAP version 1.2 part 1: messaging framework. 2nd ed. W3C Recommendation, World Wide Web Consortium (W3C); April 2007.

Gudgin M, Hadley M, Mendelsohn N, Moreau J-J, Nielsen HF, Karmarkar A, et al. SOAP version 1.2 part 2: adjuncts. 2nd ed. W3C Recommendation, World Wide Web Consortium (W3C); April 2007.

Hogben G, Jackson T, Wilikens M. A fully compliant research implementation of the P3P standard for privacy protection: experiences and recommendations. In: Proceedings of the 7th european symposium on research in computer security (ESORICS '02). London, UK: Springer; 2002. p. 104–25.

Jensen C, Potts C, Jensen C. Privacy practices of internet users: self-reports versus observed behavior. International Journal of Human-Computer Studies 2005;63(1–2):203–27.

Kolter J, Pernul G. Generating user-understandable privacy preferences. In: Proceedings of the 4th international

conference on availability, reliability and security (ARES 2009). Fukuoka, Japan: IEEE Computer Society; 2009.
Kolter J, Kernchen T, Pernul G. Collaborative privacy – a community-based privacy infrastructure. In: Proceedings of the 24th IFIP TC-11 international information security conference (SEC 2009). Berlin: Springer; 2009. p. 226–36.
Leenes R, Schallaböck J, Hansen M. PRIME white paper version 3, https://www.prime-project.eu/prime_products/whitepaper; 2008.
Leuf B, Cunningham W. The Wiki way: quick collaboration on the web. Amsterdam: Addison-Wesley Longman; 2001.
MacKenzie CM, Laskey K, McCabe F, Brown PF, Metz R. Reference model for service oriented architecture 1.0. OASIS Standard; October 2006.
Parry L. A tale of two wikis: techniques for building, managing and promoting collaborative communities. In: Proceedings of the 2nd international Wikimedia conference (Wikimania 2006). Cambridge, MA, USA; 2006.
Perfetti C, Landesmann L. Eight is not enough, http://www.uie.com/articles/eight_is_not_enough; June 2001.
Pettersson J, Fischer-Hübner S, Bergmann M. Outlining data track: privacy-friendly data maintenance for end-users. In: Proceedings of the 15th international conference on informations systems development (ISD 2006). Springer Scientific Publishers; 2006.
Pollach. What's wrong with online privacy policies? Communications of the ACM 2007;50(9):103–8.
Raggett D, Le Hors A, Jacobs I. HTML 4.01 specification, W3C recommendation. World Wide Web Consortium (W3C); December 1999.
Reay K, Beatty P, Dick S, Miller J. A survey and analysis of the P3P protocol's agents, adoption, maintenance, and future. IEEE Transactions on Dependable Secure Computing 2007;4(2): 151–64.
Sommer D, Casassa Mont M, Pearson S. PRIME architecture version 3, PRIME deliverable D14.2.d, https://www.prime-project.eu/prime_products/reports/arch/pub_del_D14.2.d_ec_WP14.2_v3_Final.pdf; July 2008.
Spool J, Schroeder W. Testing web sites: five users is nowhere near enough. In: Extended abstracts on human factors in computing systems (CHI '01). New York, NY, USA: ACM; 2001. p. 285–6.

Jan Kolter graduated from the University of Regensburg, Germany, in 2005 and holds an honors diploma in Business Information Systems. During his course of study he earned an MBA at Murray State University, USA. Since 2006 he has been working at the Department of Information Systems at the University of Regensburg. In several projects he earned expertise within the areas of usable privacy-enhancing technologies and dynamic access control systems.

Thomas Kernchen graduated from the University of Regensburg, Germany, in 2008 and holds a diploma in Business Information Systems. At the same year he started his professional career as a consultant at Steria Mummert, where he participated in several large projects that allowed him to gain versatile practical IT skills.

Günther Pernul received diploma and doctoral degrees both from the University of Vienna, Austria. Currently he is Professor and Managing Director of the Department of Information Systems at the University of Regensburg, Germany. His research interests are Web-based information systems, information security, and secure applications. Dr. Pernul is a member of ACM, IEEE, GI, OCG, member of the IFIP WG 11.3 and observer of the IFIP WG 11.8 (Security Education). He serves on the steering board of the Communications and Multimedia Security (CMS) and is co-founder of the EC-Web (since 2000) and TrustBus (since 2004) conference series. He has been involved in many research projects on national and international levels.