Generating User-understandable Privacy Preferences

Jan Kolter and Günther Pernul

Department of Information Systems
University of Regensburg
93040 Regensburg, Germany
Email: {jan.kolter, guenther.pernul}@wiwi.uni-regensburg.de

Abstract—Making use of the World Wide Web’s numerous services increasingly requires the disclosure of personal user data. While these data represent an important value for service providers, users are increasingly concerned about growing privacy threats, as more and more of their personal and private information is released to a rising number of parties.

Privacy-enhancing technologies, like the P3P specification, assist users in protecting their privacy. P3P provides means to express a machine-readable P3P privacy policy of a Web site and allows the interpretation of a dedicated P3P user agent that recommends a certain disclosure behavior. The agent’s recommendation, however, is based on the quality of pre-defined privacy preferences of the user. Accordingly, the creation of these disclosure rules requires tools that accurately record individual privacy preferences in an understandable way.

This paper introduces a novel, user-friendly privacy preference generator that allows the definition of privacy preferences for twelve different Internet service types, allowing for more precise and practical user preferences. Addressing the needs of users with different levels of experience, we present a multi-level user interface. Our solution includes a user-friendly P3P-based wizard as well as a clear and understandable configuration summary. The resulting privacy preferences of this tool will allow more accurate recommendations of future privacy agents.

I. INTRODUCTION

Due to the rising number of Web sites that collect personal user data privacy threats in the World Wide Web are growing. Much of this personal information is used for value-added services like product promotions and personalization, a promising opportunity for service providers to appeal to loyal users [1]. While surveys show that personalized services are also valued by consumers [2], this benefit is outweighed by the simultaneous loss of privacy. As users frequently transmit personal data to a growing number of service providers, users get increasingly concerned about privacy [3].

Motivated by users seeking technical means to protect their privacy, privacy-enhancing technologies emerged. An early representative is the W3C recommendation “Platform for Privacy Preferences Project” (P3P) [4]. The goal of P3P is to assist users in managing the disclosure of personal data.

P3P offers a policy language that enables service providers to define formal privacy policies that state the amount, the intended use and third party recipients of personal user information. This machine-readable privacy policy allows a dedicated privacy agent to indicate any deviation from individual, pre-defined privacy preferences of the user. With this recommendation users receive a quick estimate about a Web site’s privacy policy without the need of reading complex privacy policies. For the definition of privacy preferences P3P offers “A P3P Preference Exchange Language” (APPEL) [5].

P3P and APPEL have both been subject of strong criticism in the past [6], [7]. A frequently mentioned deficit of P3P and its underlying applications is usability. In particular, the process of obtaining users’ individual privacy preferences requires user-friendly solutions for average Internet users, as the result directly influences the user agent’s recommendations. Studies, however, reveal that users are misled by the language and ambiguous expressions during that process [8].

In this paper we introduce a tool that allows a usable definition of privacy preferences. We focus on suitable user interfaces and clear information visualizations. Three consistent complexity modes address the needs of both inexperienced users and privacy experts. Our proposed solution provides a wizard that helps users understand and comprehend all privacy-related options. Privacy preferences are configured separately for twelve categorized Internet service types, enabling users to build more realistic preferences. Finally, we developed a privacy cockpit that clearly summarizes all privacy settings along with a graphical feedback.

The remainder of this paper is structured as follows. In Section II we describe a well known privacy agent and identify its usability deficits. After discussing relevant design challenges in Section III, we present our novel privacy preference generator in Section IV. Finally, Section V concludes the paper and gives an outlook on future work.

II. PRIVACY BIRD

A prominent P3P privacy agent that builds on the potential of P3P is the Privacy Bird [9], an extension of the Microsoft Internet Explorer. As described in the last section, the Privacy Bird automatically retrieves P3P policies of Web sites. After a privacy policy is interpreted and matched with privacy preferences of the user, a bird in the browser header is used to signal the result of the matching process to the user. The result is underscored by a tweet sound.

After the initial installation the user is asked to enter his/her individual privacy preferences, which represents the foundation for all future matching processes. The simplicity of the interface (see Fig. 1) can be explained by the design goals of the Privacy Bird, which aimed for reducing the complexity of the privacy topic and, as a consequence, gaining.

1http://www.privacybird.org
For the design of our proposed privacy application we faced three design challenges identified by Cranor et al. [10], which are discussed in the following:

- **Complexity of Users’ Privacy Preferences:**
  In theory, the diversity of possible privacy preferences is infinite. The development of an application should find the balance between providing simple interfaces that allow the description of understandable preferences and the possibility to fine-tune preferences to individual personal needs. In this context, the observed distortion between predefined preferences and the actual behavior of users marks a special challenge [12], [13]. In many cases users’ behavior varies to predefined preferences, because an advantage is expected for releasing additional personal data.

- **Inexperienced Users:**
  Users’ lack of experience in the area of privacy represents a further design challenge. Even though users are increasingly concerned about their privacy, their knowledge about privacy threats and technologies that could help protect their privacy is considerably low [12]. Furthermore, users are not familiar with technical and legal terms related to privacy [14]. As a consequence, basic contents and dialogues should explain all privacy terms sufficiently. A simple and intuitive tool allows the overwhelmingly inexperienced amount of Internet users to use and understand the privacy application.

- **Complexity of the P3P Specification:**
  The first version of P3P provides eight element types, the latest version 1.1 introduces additional types. Disregarding optional attributes, 36,000 different policies can potentially be modeled with P3P. A user interface trying to capture user preferences at that fine-grained level is obviously not usable. This is especially evident, if one considers the previously discussed level of user experience.

**IV. PRIVACY PREFERENCE GENERATOR**

In Section II we identified prevalent usability deficits of the Privacy Bird. Considering the presented design challenges, we introduce a novel, user-friendly, P3P-based privacy preference generator, including a configuration wizard and a preference summary. In the following we present the main features and capabilities of our solution. We exemplify our ideas with examples and screenshots.

**A. Multi-level User Interface**

In the last section we pointed out the potential complexity of individual privacy preferences and service providers’ privacy policies. A tool focusing on the generation of fine-grained privacy preferences inevitably requires intense user interaction and fine-tuning as well as a considerable amount of time to complete the definition process. This is contrary to the prevalent inexperience of most users. Meeting the needs of this majority of Internet users calls for a tool that is understandable
and logically structured. It should be restricted to a small number of configuration options, in order to facilitate a quick generation of privacy preferences. Such an interface enables inexperienced users to fully use the application and, as a consequence, encourage them to understand and get informed about privacy-related topics and threats [15], [16]. As their level of experience gets more advanced, users will be able to define more detailed privacy preferences.

Our proposed privacy preference generator acknowledges this evolution of privacy novices. For this reason, we present a multi-level user interface, splitting the user interface into three user modes. Each mode provides a different amount of configuration options. Offering modes with differing complexity, our application targets both inexperienced privacy novices and privacy experts. As inexperienced users advance over time, modes can be switched at any time during the privacy preference generation process.

The beginner mode is designed for the majority of Internet users with little knowledge about privacy threats and data handling practices of service providers. This mode accounts for most users’ low level of experience in this area. Only basic and necessary settings are available. All settings are described in an understandable way using clear examples. Choosing the beginner mode, the user generates a complete set of individual privacy preferences in about five minutes. More experienced users can choose the advanced mode for building their privacy preferences. In this mode a more detailed definition of privacy preferences is available, allowing privacy agents to make more individualized recommendations. The expert mode allows the maximum amount of configuration settings. This mode takes advantage of most elements of the P3P vocabulary, giving privacy experts the chance to fine-tune their settings at a high degree.

After the privacy policy generator is installed and started the first time, the user is asked to choose his/her preferred mode (see Fig. 2). As mentioned above, the selected mode can be changed at any time enabling users to switch and reassess their level of experience. The contents of each mode are described in the following sections.

### B. Service-based Approach

Apart from the inexperience of most Internet users, one of the main privacy challenges we pointed out in Section III is the irrational and unpredictable disclosing behavior of users. Even though the compensation users receive for disclosing comprehensive personal data is generally small, the perceived value for users is estimated much higher [17]. As a consequence users tend to disclose more personal data than required to a service provider, not asking for the purpose of the additional data requests. This poses a further challenge for the development of a user interface that captures privacy preferences.

Considering these circumstances leads us to conclude that users transmit personal data goal-oriented, i.e. personal data is disclosed focusing a specific goal, such as purchasing a product at an online store. Furthermore, trying to achieve this goal, the decision of users what personal data to disclose is a subjective matter. E.g. some users might also disclose their phone number to an online store, even if this personal data is not necessary for the fulfillment of the service.

Addressing this goal-oriented behavior of users, our solution provides the definition of privacy preferences for each of twelve pre-defined service types. Each service type represents a World Wide Web service category and corresponds to a certain user goal. With users defining privacy preferences for a specific service type our solution is able to assist users more accurately through the generation process by providing a set of required attributes for each service type (see Section IV-C). Additionally, defining privacy preferences for each service type allows for more practical and realistic results, as it allows e.g. to define the user’s willingness to disclose his/her credit card information to an online shop, and not to release this information to any other service type, such as an Internet forum or a Web mail service.

For the identification of meaningful service types we first looked at the P3P specification. With the release of version 1.1 the Primary Purpose element (P Purpose element) allows service providers to specify the primary purpose for collecting personal data. The 23 P Purpose elements were designed to group service offers and served as a basis for our service types. Aiming for a more user-friendly and condensed service type set, we collected surveys aiming at the categorization of Internet services. A survey conducted within the PRIME project identifies ten service categories out of 46 most frequently used online services [18]. SevenOne Interactive [19] published a survey about Internet user groups. The survey is based on users’ online interests, Internet usage, eCommerce behavior as well as demographic characteristics of the test persons.

Based on the P3P P Purpose elements and the described surveys, we define the following twelve service types:

- Web Mail
- Online Shopping
- News and Knowledge Portals

![Mode Selection after the Initial Start](image)
After our tool is initially started and a mode has been selected, the user chooses one out of twelve service types (see Fig. 3). Subsequently, a configuration wizard starts that collects user’s privacy preferences for the selected service type. After the completion of the wizard the user can choose to configure the next service type or to proceed to the privacy cockpit (see Section IV-F). A non-configured service type indicates that the user is not willing to interact with service providers of this service type. In order to allow easy recognition, we designed icons, which allow for an easy association with each service type.

C. Data Minimization

A central content of privacy preferences is the amount of data a user is willing to disclose. For this purpose, the P3P specification offers a basic set of data types and data categories focusing primarily on online and demographic data. In order to represent all commonly used personal data transmitted in the World Wide Web, we extended the P3P vocabulary with additional data types.

In our approach we enable users to define a set of appropriate data types users are willing to transfer to a particular service type.

Using a given service in the World Wide Web should not require more than a certain, adequate amount of personal data from users. This set of appropriate personal data generally depends on the service type a certain service belongs to. For example, an instant messaging service should not require more than a nickname and, possibly, an e-mail address, while online shopping services additionally need users’ payment information and delivery address to fulfill their offered task.

In an effort to address the needs of inexperienced users, we provide recommended data sets for each of the twelve service types. A survey helped us identify the data set each service type generally requires. Furthermore, our results also incorporate the outcome of a PRIME survey [18].

Knowing the maximum amount of personal data a service provider should request, our application is able to display this information during the privacy preference generation process. In the beginner and advanced mode the application recommends assuming this data set and not releasing additional data to a service provider. In order to offer a well arranged user interface, we grouped personal data into categories and designed an icon for each data category. In our example, Fig. 4 shows that the selected service type "Online Shopping" requires data from the categories login data, online data, name data, postal data and financial data. Moving the mouse over a data category triggers a tool tip that shows the individual data types for a category.

In addition, the user is asked to specify the behavior of a privacy agent, if a Web site requests more data than the displayed recommended data set. The user can choose to

- not accept the release of additional data,
- accept the release of personal data, but to be specifically warned and informed about this practice or
- accept the release of additional personal data without notification.

This configuration is offered to the user in accordance with the Behavior element of APPEL [5]. The Behavior element can take on the values block, limited and request. The values are applied to the three options respectively. Additionally, option two employs the APPEL Prompt element to trigger a user notification.

If the user chooses the expert mode, the wizard allows to directly adjust the data set he/she is willing to release to
a service type (see Fig. 5). The left area lists all available data types a user owns. All data types are grouped into their corresponding categories. The right box shows data types selected for disclosure. By default, the right box contains the recommended data set.

Double-clicking an element moves it from one box to the next. The same functionality is provided by the buttons placed between both boxes. One button resets the elements to the recommended data set.

D. Purpose of Collecting Data

Apart from the amount of personal data, controlling and restricting the usage of these disclosed data is a primary goal of privacy-enhancing technologies. In the context of our proposed application, users can choose, what purposes in addition to the fulfillment of the original service a service provider may use transferred personal data for. The two most frequent additional purposes are personalization and contacting.

With personalization techniques service providers tailor content of their Web sites to users’ needs and interests by using disclosed personal data and observing users’ surf behaviors [20]. A well known example of personalization measures is the individual product offering a returning customer gets at an online shopping Web site. A further, more intrusive example is an individually tailored advertisement.

Our proposed application enables users to limit personalization activities of service providers. The P3P specification defines five personalization elements. Catering usability needs, we aggregated the elements into three distinctive groups:

- One-time Personalization:
  A service provider collecting personal data for the purpose of one-time personalization adjusts content and design of a Web site using transferred personal user information as well as click stream information. Personalization is only conducted for the current visit. Personal data collected for one-time personalization will not be stored by the service provider. Hence, personalization is not applied at any future visit. The P3P element indicating one-time personalization is tailoring.

- Pseudonymous Personalization:
  Unlike one-time personalization, a service provider collecting personal user information for pseudonymous personalization stores personal user data for future personalization measures. Additionally, the service provider uses personal user data to analyze web offers. This purpose only uses non-identifiable user data, such as gender, year-of-birth. The corresponding P3P elements for pseudonymous personalization are pseudo-analysis and pseudo-decision.

- Individual Personalization:
  In addition to pseudonymous personalization, service providers collecting personal user data for individual personalization also use identifiable user information for personalization measures (e.g. date-of-birth, postal address, e-mail address). The P3P elements representing this group are individual-analysis and individual-decision.

A further prominent and valuable purpose of collecting personal data is to contact users directly. In many cases this contact is triggered by a service provider’s marketing department that promotes a product or offers an individualized service. The communication channels users can be contacted are e-mail, postal or telephone.

As most users dislike being directly contacted, our wizard provides options to adjust privacy preferences accordingly. Following the P3P specification we adopt the two contacting categories mail (e-mail and postal) and telephone. The corresponding P3P elements are contact and telemarketing.

Fig. 6 depicts the personalization and contacting page for the beginner mode. As users in this mode tend to be inexperienced and value user-friendliness over rich configuration settings, we limited available options to allowing or disallowing personalization and contacting. We left out further options of fine-tuning personalization and contacting. All options are
explained thoroughly. In our example the user chooses to allow personalization and not to allow contacting for the selected service type "Online Shopping". Serving understandability and user acceptance of the application, we designed an individual purple icon and placed it at the right side of the page underscoring the selection of the user. The icon is grayed out for unselected options. Moving the mouse over a text field or icon triggers a tool tip that presents additional information and examples to the user.

In the advanced mode we allow more fine-grained configuration of each purpose. For this purpose, we split the page of the beginner mode into a personalization page and a contacting page. The personalization page presents all three identified personalization types, the contacting page the two available contacting channels. Fig. 7 shows the personalization page in the advanced mode. In our example the user opts to allow one-time personalization as well as pseudonymous personalization. The user does not allow individual personalization. Again, the user selection is supported by individually designed purple icons for each option.

The personalization and contacting pages for the expert mode are based on the respective pages of the advanced mode. In addition, users can refine the settings of each option. Using the required option of the P3P specification, we enable users to bind their decisions on conditions. Using a drop-down menu, users can choose to

- Never allow this option (unchecked box in the advanced mode)
- Require an explicit user consent (opt-in)
- Require an explicit user rejection (opt-out)
- Always allow this option (checked box in the advanced mode)

The personalization page of the expert mode in Fig. 8 shows that - unlike in the advanced mode - the preferences of the user concerning one-time personalization and pseudonymous personalization are bound to conditions. The user only accepts Web sites with one-time personalization, if the service provider offers an option that disables this function (opt-out). For pseudonymous personalization, the user even requires explicit consent (opt-in). Like in the advanced mode the user chooses not to allow individual personalization.

E. Disclosure to Third Parties

The last page of the configuration wizard enables users to control the recipients of their personal data. As many service providers pass user data to third parties, our application gives users means to customize their respective preferences.

Like for personalization, we looked at the P3P specification and aggregated third party groups based on the recipient elements. The P3P element ours describes the service provider itself, or agents that act on the behalf of the service provider. It is set by default in users' privacy preferences. The remaining P3P recipient elements represent third party entities, which are grouped as follows:

- Third Parties Following Equable Practices: This group comprises of third parties with equable privacy and data handling practices. Those parties use personal user data on their own behalf. However, those parties are required to follow the same restrictions concerning user data as the service provider. The corresponding P3P element for this group is same.
- Third Parties Following Different Practices: This third party group’s privacy and data handling practices differ from those of the service provider. These parties are not affiliated with the service provider and use personal data on their own behalf. The P3P elements for this group are delivery and other-recipient.
- Public Areas: This group represent all publicly available areas like fora and public directories. It also includes third parties with unknown privacy and data handling practices. Their associated P3P recipient elements are unrelated and public.
The third party page design, including the composition of each mode, resembles the personalization and contacting pages. In the beginner mode, the user can solely choose to accept disclosure to third parties following equitable practices (the first defined third party group) and to accept disclosure to other third parties (combination of the second and third party group). In the advanced mode all three defined groups are available as options. Like for personalization and contacting, users can bind their decision on conditions in the expert mode.

Fig. 9 presents the third party configuration page for the beginner mode. In our example the user only accepts the disclosure of personal data to third parties following the same privacy and data handling practices as the service provider.

F. Cockpit

After the user completed the configuration wizard for a service type, our solution provides a clear overview of users’ privacy preferences. For this purpose, the application switches to the privacy cockpit, the main page of the privacy preference generator (see Fig. 10). This page provides a quick and comprehensible summary of all configured service types, enabling users to oversee and double-check all settings and options. In addition, the cockpit page provides an evaluation of each configured service type with regard to the impact the configuration has on users’ privacy.

The cockpit page is designed as a dynamic table. The columns represent each service type’s configuration settings. As users value consistent interfaces [15], the selected mode - just like in the configuration wizard - does not change the design of the privacy cockpit. As the number of available options in the beginner mode totals six, the selection of these options is visualized by six rows. In the advanced and expert mode the additional options of the configuration wizard are shown as well.

In order to give users a meaningful and understandable feedback of their configuration, we sorted service types by their impact on users’ privacy. Doing this, we calculate a privacy level for each service type taking equally into account the configurations for personalization, contacting and disclosures to third parties. Enabling users to switch modes at any time while maintaining privacy level scores for each service types, the aggregated options in the beginner mode have a higher weight in each configuration category than the more fine-grained options in the advanced mode. The resulting score for each service type ranges from 0 (all options selected; lowest privacy level) to 1 (no options selected; highest privacy level).

As mentioned above service types are arranged by their privacy level in ascending order. Fostering user acceptance and providing a quick recognition and evaluation of privacy levels, we align all service types on a dynamic color scale. The color scale underscores the calculated privacy level, starting with red that fades into yellow and green. According to the calculated privacy level score, we allocate service types as follows:

- Privacy Level Score between 0 and 0.33: Red
- Privacy Level Score between 0.34 and 0.66: Yellow
- Privacy Level Score between 0.67 and 1: Green

Fig. 10 shows a screenshot of the cockpit in the advanced mode. The cockpit depicts the exemplary configuration of seven service types. The service type “News and Knowledge Portals” shows the lowest privacy level score and, as a result, is listed first. As its calculated privacy level score is 0.28, it is placed on the red side of the background color scale. The adjacent service types “Banking” and “Social Networking” have a score of 0.44 and 0.5, respectively, showing a yellow background. The remaining service types “Online Shopping”, “Web Mail”, “eLearning” and “Forums and User-generated Content” are placed on the green side of the color scale with privacy level scores of 0.67, 0.72, 0.89 and 1, respectively. We are pointing out that the background color scale dynamically shifts, as the configuration and the privacy level scores change.
Targeting usability and understandability, we designed five icons that emphasize the settings. The LiveHelp area helped us to separate detailed information about each setting. Moving the mouse over an icon triggers an explanation in the LiveHelp area.

A click on the "Configure a Service Type" button on the bottom opens the service type selection page (Fig. 3) and allows the addition or change of service types. If the user wants to directly change the configuration, a click on a type icon restarts the configuration wizard for that particular service type.

G. Output

The privacy cockpit page allows users to save their privacy preferences. Our goal is to describe all settings in an understandable and logical manner. As APPEL, the P3P privacy preference language, shows clear design and application shortcomings [7], we employed a simple XML pattern for saving privacy preferences. As far as possible we adhered to the vocabulary defined in the P3P specification.

H. Implementation

A prototypical implementation of the presented Privacy Preference Generator is available for download2 (see footnote).

V. CONCLUSIONS

The rising number of service providers collecting personal user data as well as their unknown data handling policies requires usable privacy tools that assist users in controlling the transfer of their valuable personal attributes. Privacy agents targeting these user needs retrieve and process machine-readable privacy policies of service providers. However, their results rely on the quality of users’ previously specified privacy preferences. Given the inexperience of the average Internet user in the area of privacy, existing solutions for the definition of privacy preferences do not provide sufficient means to build understandable user preferences.

Focusing on usability this paper introduces a user-friendly privacy preference generator that facilitates more accurate definitions of privacy preferences for twelve Internet service types. Three different complexity modes tailor this tool for both inexperienced users and privacy experts. Building on the vocabulary of the P3P specification, a configuration wizard guides users through a set of privacy-related options, showing clear explanations and examples, and recommending an adequate data set for each service type. The privacy cockpit provides an understandable and comprehensible overview of the resulting configuration as well as a quick evaluation. The output of this application results in more meaningful privacy preferences that have the potential to foster the accuracy and user acceptance of coming privacy agents.

Future work will involve usability experiments as well as an evaluation of the defined service types with their recommended data sets. Furthermore, we plan the development of a corresponding browser agent.

ACKNOWLEDGMENT

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

REFERENCES


2http://www-ifs.uni-regensburg.de/Privacy/PPG.zip