Improving Understanding of Website Privacy Policies with Fine-Grained Policy Anchors

Stephen E. Levy
Watson Research Center, IBM
19 Skyline Drive, Hawthorne, NY, 10532
United States
+1 877 754 7529

levysn@us.ibm.com

Carl Gutwin
Computer Science Dept., University of Saskatchewan
110 Science Place, Saskatoon, SK, S7N 5C9
Canada
+1 306 966-8646

carl.gutwin@usask.ca

ABSTRACT

Website privacy policies state the ways that a site will use personal identifiable information (PII) that is collected from fields and forms in web-based transactions. Since these policies can be complex, machine-readable versions have been developed that allow automatic comparison of a site's privacy policy with a user's privacy preferences. However, it is still difficult for users to determine the cause and origin of conformance conflicts, because current standards operate at the page level – they can only say that there is a conflict on the page, not where the conflict occurs or what causes it. In this paper we describe fine-grained policy anchors, an extension to the way a website implements the Platform for Privacy Preferences (P3P), that solves this problem. Fine grained policy anchors enable field-level comparisons of policy and preference, field-specific conformance displays, and faster access to additional conformance information. We built a prototype user agent based on these extensions and tested it with representative users. We found that fine-grained anchors do help users understand how privacy policy relates to their privacy preferences, and where and why conformance conflicts occur.

Categories and Subject Descriptors

K.4.1 [**Computers and Society**]: Public Policy Issues – *privacy*; H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia – *user issues*.

General Terms

Design, Security, Human Factors.

Keywords

Privacy policies, privacy preferences, P3P, APPEL, conformance conflicts, user agents, e-commerce.

1. INTRODUCTION

Web-based transactions are now a common part of life on the Internet: people can search for merchandise, browse catalogues, choose and pay for selected items, and arrange for shipping and delivery. During these activities, websites often require the user to disclose personally identifiable information (PII) in order to establish a customer relationship. As with traditional commerce, user confidence and trust in the use of this information is essential for the transaction to succeed: people must feel confident that the personal information they disclose will be used only for agreed-upon purposes and will not be misused by the vendor.

Human-readable privacy policies are now being displayed on websites to help build user confidence and trust in the process of personal information disclosure. These policies explain how personal information collected by the vendor will be used. However, simply having the policy on the website does not guarantee understanding, since policies can be complex. A user must take additional time and expend additional effort to understand the content of the privacy policy and determine for themselves whether the website conforms to their personal privacy preferences.

Technologies have been developed to help reduce user effort in navigating website privacy policies. First, the Platform for Privacy Preferences (P3P) allows privacy policies to be encoded in machine-readable form [1]. Second, A P3P Preference Exchange Language (APPEL) provides a machine-readable rule set for the user's privacy preferences [2]. P3P and APPEL allow website privacy polices and user privacy preferences to be compared automatically for conformance.

P3P user agents are the mechanisms for this automation. They read the privacy policies implemented by a website and show the conformance of the vendor's privacy policy with the user's privacy preferences. Current P3P user agents (e.g., ATT Privacy Bird [3]) present a visual indication of site conformance in the browser's title bar and give a more detailed conformance report in a separate window (see Figures 4 & 5).

Although these user agents are a large step forward over simple text policies, it is still difficult for users to determine the cause and origin of conformance conflicts. The main reason is that user privacy agents have a coarse view of the vendor privacy policy, operating only at the level of the web page. There is no machine-readable connection between privacy statements and the specific input fields of a web form, and as a result, conformance information presented to the user has no visible link to the field that caused the problem. Since there can be many fields on a page, the user must still do considerable work to understand the conflict. Since users are often unwilling to expend this effort [4], many transactions are cancelled as a result.

In this paper, we show that this problem can be addressed by representing and visualizing privacy conformance at the input field level. Our solution – the Integrated Privacy View – has two parts: first, fine-grained policy anchors that map privacy policy statements to specific input fields, and second, a user agent that displays conformance information with an icon next to each field.

We describe the design and prototype implementation of these two parts below, after a brief review of previous work in the area. We then report on a small user study that confirmed the potential value of our approach and provided feedback on the visual design, and conclude by discussing the steps that must be taken for fine-grained anchors to be used more generally.

2. RELATED WORK

There are four areas of previous research that underlie the problem outlined above. In the following sections we review relevant work on trust and privacy in economic exchange, specifications for privacy policies and preferences, and P3P user agents.

2.1 Privacy and PII

There are several definitions of privacy that relate to different aspects of a person's relationship to the outside world. Privacy is often seen as an issue of control over the inflow and outflow of information: control over one's degree of interaction with others, and control over other's access to information about us [5].

The latter issue is the one of interest in this paper – the right of individuals to determine when, how, and to what extent information about them is communicated to others [6]. In particular, we are concerned with Personal Identifiable Information (PII), which is the set of information that can be stored and associated with an identifiable person [7]. For example, PII includes a person's name, email address, telephone numbers, medical statistics, membership in groups, relationships to other people, financial data, and purchasing history.

Some of this information is required for transactions on the Internet – usually data like the person's name, email address, shipping address, and credit card information. As described below, the storage and use of this information can have a large effect on the trust that a consumer has in the transaction.

2.2 Trust and Privacy in Economic Exchange

Economic exchanges only take place when parties trust one another to fulfill their obligations in a timely and efficient manner [8,9,10]. One of these obligations is that information collected during the transaction will be used in appropriate and agreed-upon ways [9].

In pre-Internet commerce, two factors contributed to consumer trust in a vendor: personal or word-of-mouth knowledge of the vendor's practices, and the fact that personal information was relatively difficult to collect, store, and analyze. With Internet transactions, however, these factors are less likely to occur. Personal experience with vendors is limited, and it has become much easier to gather personal information, share or sell it, and mine it for further business purposes.

The results of consumer surveys on privacy and security on the Internet show that individuals are very concerned about disclosing personal information. Several studies show that a large majority of Internet users (70-85%) are concerned about the security of personal information [10,11,12,13], and about the possibility that businesses will use their data for undesired purposes such as telemarketing or spam [12,14]. Approximately two-thirds of users polled in two different studies were unwilling to shop online because of privacy concerns [15,16], and 27% of consumers had abandoned online shopping carts because of privacy reasons [10].

Therefore, the public assurances made by vendors and organizations are often the only means that a consumer has to establish trust in the relationship. Knowledge of the vendor is often only available through the commitments and guarantees made on the vendor's website. These assurances express not only how the order or service will be fulfilled but also how the customer's information will be used. The vendor's reputation will be based both on how well the transaction is fulfilled, and on how well the vendor protects and respects user privacy.

There are real benefits to be gained from being able to generate trust in the way that privacy will be handled. One survey [17] found that more than 72% of web users said they would give websites their personal information if the sites would provide a statement regarding how the information would be used, how long the information would be maintained, and to whom the information would be disclosed.

2.3 Privacy Policies and Preferences

Privacy policies are a way to provide information and guarantees that will increase a consumer's trust. Privacy policies are now available at most commercial web sites. However, the effectiveness of making a policy available is limited by the amount of effort consumers are willing to invest in reading and understanding that document. Privacy policies are often long and complex, and it is difficult for users to find the issues of interest to them [18]; perhaps as a result, a recent survey found that only 54% of respondents said that they would read a site's privacy policy on the first visit [4].

The problem of effort is one that can be assisted through technological means. By electronically capturing both the user's privacy preferences and the website's privacy policies, a conformance evaluation may be done automatically for the user. Machine-readable privacy standards have been developed to represent both an organization's policy and an individual's preferences.

The Platform for Privacy Preferences (P3P) defines a common way for websites to publish a privacy policy stating what the website does with data it collects [1]. P3P is an XML language designed such that browsers or other user agents can easily match a user's privacy preferences with a website's privacy policy before the user provides personal data to the website. For the individual, A P3P Preference Exchange Language (APPEL) allows users to specify what uses are acceptable and what actions to take to inform the user of conflicts [2]. Figures 1 and 2 show example P3P and APPEL statement related to a user's email address.

```
<STATEMENT>
  <PURPOSE>
    <contact/>        (the PII will be used to contact the user)
    <current/>        (the PII will be used for the current transaction)
  </PURPOSE>
  <RECIPIENT>
    <ours/>           (the PII will be used by the company)
    <delivery/>       (the PII will be given to the delivery company)
  </RECIPIENT>
  <RETENTION>
    <indefinitely/>   (the company will retain the PII indefinitely)
  </RETENTION>
  <DATA-GROUP>
    <DATA ref="#user.home-info.online.email"/>
  </DATA-GROUP>
</STATEMENT>
```

Figure 1. Example P3P statement stating how a user's home email address will be used.

```
<appel:RULE
    behavior="limited"
    description="This site intends to share information that personally
        identifies you with other companies and telemarketers">
    <p3p:POLICY>
      <p3p:STATEMENT appel:connective="and">
        <p3p:PURPOSE appel:connective="or">
          <p3p:contact />
          <p3p:telemarketing/>
        </p3p:PURPOSE>
        <p3p:DATA-GROUP>
          <p3p:DATA ref="#user.home-info.online.email"/>
        </p3p:DATA-GROUP>
      </p3p:STATEMENT>
    </p3p:POLICY>
</appel:RULE>
```

Figure 2. Example APPEL statement, indicating that if a site's purpose is to contact the user regarding other products or services, a conflict will be reported.

2.4 P3P User Agents

A P3P user agent is a personal assistant designed to help users understand a website's privacy policy in relation to their privacy preferences. The job of the agent is to accurately and simply present policy conformance to the user.

P3P user agents differ in the amount of the P3P specification that they implement, and how conformance indication is conveyed to the user. The P3P specification provides policy definition for PII and cookies; an agent may decide to address only cookies, or all user-provided data. Conformance indication may be displayed in a separate window, in the frame of the browser, or in the page itself. Currently the two most widely used P3P user agents, IE6 and AT&T Privacy Bird, use the frame of the browser.

2.4.1 The IE6 User Agent

The IE6 agent uses the P3P 'compact policy' that can be transmitted in HTTP headers when cookies are set. IE6 uses this information to make cookie-blocking decisions. When the cookie policy of the vendor's site does not match the user's preferences, IE6 displays an eye covered by a do-not-enter sign (Figure 3). Additional information about the conflict is available by mouse click on the eye icon.

Figure 3. IE6 privacy icon displayed in browser frame.

Without a full set of user preferences, however, IE6 can only warn the user about personal information stored in cookies, and not about the collection of PII in web forms.

2.4.2 The AT&T Privacy Bird User Agent

The AT&T Privacy Bird implements the complete P3P and APPEL specification [19]. The system displays a bird icon in the browser's title bar that changes color (and can also make sounds) to indicate whether or not a website's P3P policy matches a user's privacy preferences (Figure 4). The icon is also used to access the privacy policy information and the conformance information. This additional text is presented in a separate window (Figure 5).
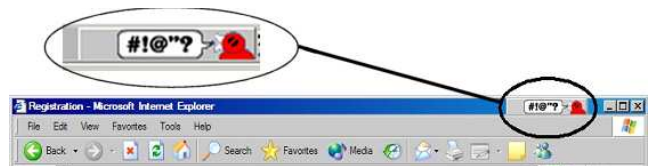

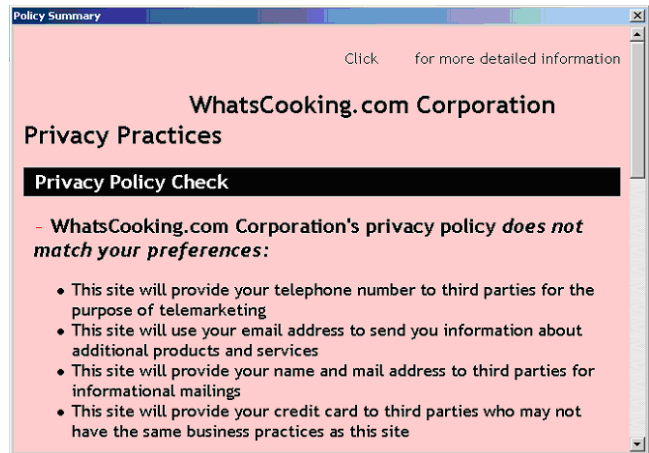Figure 4. Privacy Bird indication of a conformance conflict.

Figure 5. Privacy Bird detailed conformance report.

P3P user agents must work within the current specifications of P3P and APPEL. However, the P3P specification limits the granularity of policy information to a single web page, which limits any visualization of conformance to the page level (and usually in practice, only one P3P policy is created for the entire website). This coarse granularity leads to several problems for user agents (including Privacy Bird). Agents are unable to indicate the particular field that causes a conformance conflict, and web pages that have no input fields will show the conformance result for the entire site. The visualization is the same whether there is one conformance problem or several, and there is no way for the user to know when fields are *not* in conflict; they must read through the subsequent information to see the result for each field. This also means that when a user wishes to get further information about a particular element on the form, all of the privacy policy conflicts must be presented in a separate window, which can occlude the web page. Finally, a global conformance indicator in the title bar may not be noticed, particularly if the user is focused on the form-filling task. The sound alert can grab the user's attention, but at the risk of being annoying, since the alert will appear on all pages where there is a conflict.

In the next section we describe our solution to these problems, called the Integrated Privacy View (IPV). A demonstration version of IPV can be explored at: hci.usask.ca/IPV/.

3. THE INTEGRATED PRIVACY VIEW

IPV is a system for checking and displaying privacy conformance information at the input field level. IPV has three main parts: an extension to P3P that allows fine-grained linking of policy statements to HTML elements, a mechanism for intercepting web pages that contain the extensions, and a new user agent to perform the conformance check and insert the visual indicators into the web page. These parts are described below, after we outline the user's view of the web when using the IPV agent.

3.1 A User's View of the Web with IPV

When a user arrives at a web page that contains input fields, IPV inserts icons beside each field to indicate whether the PII required for that field will be used in a way that conforms to the user's privacy preferences (preferences are set up elsewhere using a separate system). As shown in Figure 6, a green smiley-face icon indicates conformance, and a red frowning face shows a conflict.
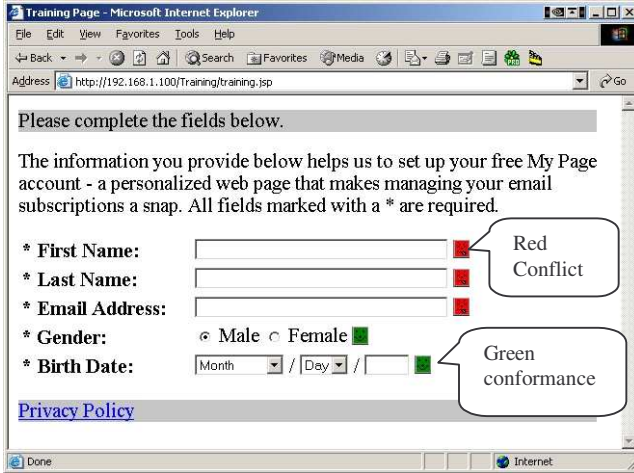

Figure 6. IPV showing conformance indicators next to fields.

To get more information about a field, the user mouses over the conformance icon, resulting in a popup window beneath the input field. The popup displays a short description of the conformance result, and in the case of a conflict, shows additional information from the privacy policy describing the problem (see Figure 7).
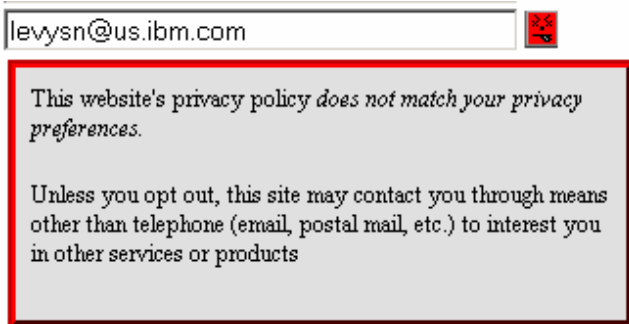

Figure 7. Popup window resulting from mousing over the conformance icon.

We now turn to the three parts of IPV: fine-grained anchors, a page interception mechanism, and the user agent.

3.2 Fine-Grained Policy Anchors

The first part of IPV provides a way to link a specific P3P policy statement to a particular input field on a web page. This link requires three parts: a way to specify a particular P3P statement, a way to specify a particular input field, and a way to find the IPV attribute in a web page. We describe these parts first, and then describe the fine-grained policy anchor itself.

Specifying a P3P statement. A P3P privacy policy already permits a designer to state the applicability of individual privacy statements through the <DATA_GROUP> element. These contain DATA elements, which are able to specify either general classes or individual items of PII. This data element reference provides us

with the means to identify a specific privacy policy statement (this will be used in the extension described below).

Specifying an input field. Input fields are built in HTML 4.0 or XHTML with the <input/> element, which allows a range of attributes such as size and default value (see example in Figure 8). New attributes are also permitted, and are ignored by the browser. As described below, adding an attribute that specifies the data element reference provides the link to the privacy statement and also provides an anchor point in the HTML to insert a visualization of the conformance result.

Finding the IPV attribute. A privacy agent needs to be able to find the privacy statement link in the target web page. Web pages coded in XHTML permit fast parsing into a tree structured Document Object Model (DOM). A privacy agent may then utilize XPATH to locate all input nodes with the data element attribute. The value of the data element attribute can then be used to find the associated privacy statement.

3.2.1 The P3Pdataelement attribute

IPV defines a new attribute for the HTML element that defines an input field. The *p3pdataelement* attribute specifies a P3P DATA element that indicates a specific policy statement to associate with the input field (see Figure 8).

IPV uses the fact that P3P defines a base data schema naming each data element that a vendor might collect. For example, a user's business email address would be specified as #user.business-info.contact.online.email. Both P3P and APPEL use this base data schema and its associated data categories to define the scope of policy statements. Conformance evaluation uses the data elements as one of the facts to match between a privacy preference and a privacy policy.

```
<input name="registration_email"
size="40"
value="levysn@us.ibm.com"
p3pdataelement="#user.business-info.online.email" />
```

Figure 8: HTML input field element (p3pdataelement in bold).

The p3pdataelement tag is used to locate both the P3P statement and the input field in the HTML DOM. This location is then used as an anchor point to insert the conformance visualization as described below.

The privacy statement link through this new attribute allows the privacy policy designer to work independently of the web page designer. Changes to the content of a privacy statement for a given data element do not require corresponding changes to be made by the web page designer. However, new input fields that are not covered by the current privacy policy do require the privacy policy designer to update the policy.

3.3 Diverting Web Pages to the IPV Agent

Our prototype uses an HTTP proxy to intercept, examine, and potentially modify the target web page. In future, this part of the system will be re-implemented as a browser helper object, but the proxy version is identical in functionality, simpler to implement, and works for all browsers.

The purpose of the IPV proxy is to look for HTTP responses from the server that may contain the p3pdataelement attribute. For all pages with a MIME type of HTML, the IPV agent is invoked. IPV parses the web page looking for the p3pdataelement. If none is found, then the unmodified web page is returned to the client;

otherwise the privacy policy conformance evaluation is carried out (see below) and the page is modified accordingly.

3.4 The IPV User Agent

The IPV user agent compares the user's privacy preferences and the vendor's privacy policy, and then produces a conformance display for the user. There are two parts to the agent: a part that carries out the conformance checking, and a part that inserts the conformance indicators and additional conformance information into the page. The agent is implemented in Java.

3.4.1 IPV Conformance Checking

The IPV agent takes as input a web page, a user privacy preference specification, and a vendor policy, and produces a conformance report that can be visualized in the web page. The agent performs the following steps:

1. Read the source web page and parse the text into an XML DOM (for our prototype, the page must be valid XML).
2. Read the vendor P3P policy from the web site, and the user privacy preferences (in APPEL) from the local machine; these documents are parsed into two XML DOMs.
3. Create a node list of input elements in the web page that have the p3pdataelement attribute.
4. For each node from Step 3:
   4.1. Identify the P3P statement that is associated with the p3pdataelement attribute.
   4.2. Check the P3P statement against the user privacy preferences. This process involves three steps:
      4.2.1. Normalize privacy preferences and privacy statement to expand default or implicit facts and logical operations
      4.2.2. Perform match by looking for the existence of each fact from the preferences in the statement
      4.2.3. Gather text descriptions for each conformance conflict for later display
   4.3. Place conformance result in the web page by updating the page's DOM with additional XML fragments that display the conformance result (see Figure 9). We also add Javascript and CSS to the HTML header to support IPV's display and interaction capabilities.
5. Deliver the modified web page to be rendered by the browser.

3.4.2 IPV Conformance Visualization

A field-specific visualization of conformance information must clearly indicate the conformance of each input field, and must also provide a means for obtaining additional conformance information. The indicator should not interfere with the web page design or with the functioning of the input field.

In IPV, conformance indication is done through a combination of HTML, CSS, and Javascript. The agent manipulates the HTML in the web page, adding new elements to display icons, borders, and popups. The agent knows where to place the new elements because of the location of the p3pdataelement attribute. CSS is used to control the style, location and visibility of the new elements. Javascript permits the capture of user events such as mouse movements, which enables user interaction with the visualization (e.g., popup windows from mouse-over events).

There are several possible designs for the visualization. One simple approach would highlight the input field background or add a colored border. Additional conformance information could be available on mouse-over. The advantage of this design is that

no additional screen real estate is used, and the form layout is exactly as the web page designer intended. The disadvantages of this approach are that the web page designer may already use color for other purposes, and that mousing over the field is sometimes reserved for context-sensitive help. This suggests that a new object could be added to the page, both to show the conformance result and to give the user a visual target for obtaining further information.

This is the approach followed in the current IPV prototype. We place an icon in the page beside the input field; the icon uses both color and pattern to indicate conformance, and provides a mouse-over target for obtaining additional information. The same color scheme as the AT&T Privacy Bird was used: green to indicate conformance and red to indicate conflict. A smiley face was chosen to indicate conformance and a sad face to indicate conflict (see Figure 10). The icon is attached to the HTML input element by adding an HTML *span* element as an outer container. By embedding the span element as a child of the input element, the browser renders the conformance icon next to the correct input field (see Figure 9).

```
<input
  name="registration_email" size="40" value="levysn@us.ibm.com"
  p3pdataelement="#user.business-info.online.email">
  <span class="IPVframe">
  <span class="IPVconflict"
    onmouseover="IPVMouseOver(this)" />
  <span class="IPVmoreInfo"
    onmouseover="window.status='info'">
  <span>
    This website's privacy policy does not match your privacy
    preferences. Unless you opt out, this site may contact you
    through  means other than telephone (email, postal mail, etc.)
    to interest you in other services or products.
  </span></span></span>
</input>
```

Figure 9: Abbreviated XML fragment for conformance display.

Green happy face indicates conformance	Red frowning face indicates conflict

Figure 10. Conformance icons (3X normal size).

4. EVALUATION OF IPV

We tested the IPV prototype in a small usability study. We had two goals in this evaluation:

- to determine whether IPV helps the user understand the cause and origin of conformance conflicts when compared with the current state of the art (Privacy Bird);
- to explore visual design issues in the presentation of conformance information, such as how multiple fields should be grouped, whether conforming fields should have an indicator, what to do when there are no fields on a page, and the problem of clutter and distraction.

This study looked at only a small number of participants, and we plan larger evaluations in future. Nevertheless, the subjects were realistic users, and the study was able to gather considerable data about the usability of the different agents and the design of the conformance visualization.

4.1 Methods

4.1.1 Participants

Six participants (three male and three female) were recruited from a local company. All individuals were experienced users of the Internet, and all had actually made purchases on the web. Participants were all aware of PII privacy issues, and had taken varying steps to protect their privacy during web-based transactions (e.g., inspected privacy policies, checked that the browser was in secure mode); however, none had ever used a privacy agent before.

4.1.2 Experiment Setup

The following paragraphs describe the websites created for the study, the two user agents, and the tasks given to the participants.

Websites

The experiment recreated a typical experience of requesting a product or service through a web transaction. Two similar websites were created (*WhatsCooking* and *AllThatJazz*), following typical best practices guidelines for creating commercial sites [20].

Each site had three pages (see Figures 11-13).
- The first page was a login or signup form, containing only a few input fields.
- The second page was a longer form where the main transaction takes place, and was designed to gather the complete personal profile of the participant and permit the participant to fully exercise each privacy agent. This page contained some data entry fields that conflicted with the privacy policy, and some that did not. This page was also designed to be longer than the typical monitor could display without scrolling, putting some fields (and some indicators in the case of IPV) below the fold of the page.
- The third page of the website had no input fields, and was designed to explore what an agent should display when no input is required.
- The fourth page was the site's human-readable privacy policy, which was linked to each of the other three pages.

User Agents

The experiment looked at two user agents: the AT&T Privacy Bird and the IPV agent. IPV was set up as described earlier (see Section 3.1). Privacy Bird was set at its default configuration, but with sound turned off.

The pages were set up with the extended IPV attributes as described above. Only one privacy policy was set up for the three pages, following common web practice. This meant that Privacy Bird showed results for the entire site rather than at the page level. For example, for page three of the site where there were no input fields, the AT&T Privacy Bird showed the conformance result for the entire website; the IPV agent showed no indicators at all. This decision was made to more accurately compare IPV against the current state of the art, since no commercial websites currently specify a different policy for each page of the site.

Tasks

The task given to the participants was to carry out the web transaction, supplying personal information as requested by the site. Participants were asked to watch for privacy conflicts during the transaction, and whenever they noticed one, to identify the PII and the privacy policy statement that caused the conflict with their privacy preferences.

The personal information requested by each website during the task was the same. Participants were provided with fictional personal information to use for the transaction. Each participant saw both websites and both user agents, with order balanced so that both sites and both agents were seen the same number of times in each position.

Figure 11. *WhatsCooking* site, page one, IPV version. Privacy Bird version would show the red bird icon as seen in Figure 13.
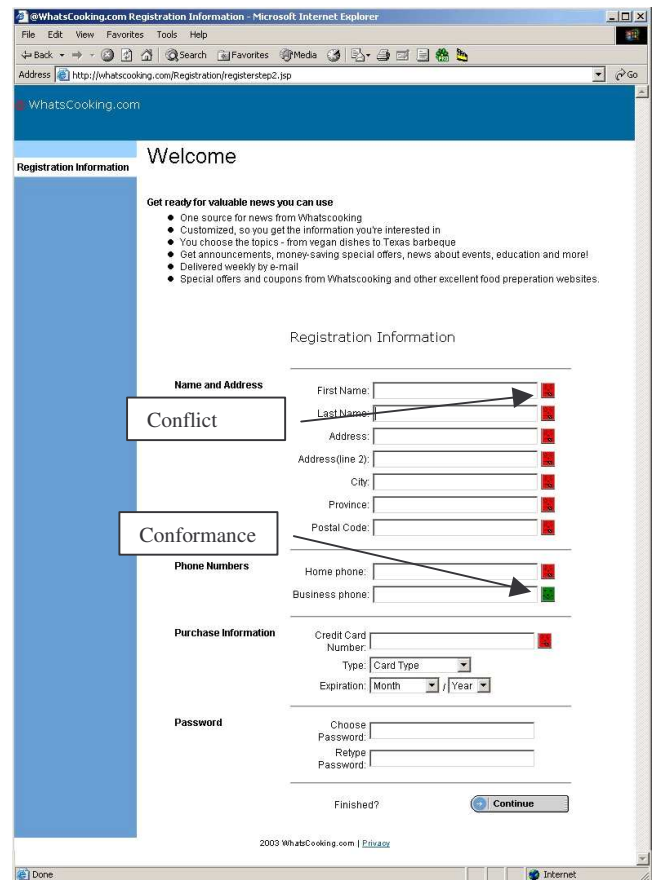

Figure 12. *WhatsCooking* site, page two, IPV version. Privacy Bird version would show red icon as seen in Figure 13.

Figure 13. *AllThatJazz* site, page three, Privacy Bird version. IPV version for this page would show no indicators at all.

4.1.3 Procedure

Participants were given basic training in privacy policies, privacy preferences, user agents in general, and both of the agents used in the study. Participants carried out a training task with both agents; all subjects were able to understand how the agents worked, and how to identify conflicts and request further information about the conflict.

Participants then carried out the main task, first with one site and agent, and then with the other site and agent (order was balanced as described above). People's reports on privacy conflicts were recorded, and a log of their actions on the site was kept. After the tasks were complete, participants filled out a questionnaire asking them about their experiences and preferences.

Finally, participants looked at and discussed a series of paper prototypes that showed different design approaches for grouping, highlighting, and icon placement. Participants were asked to evaluate these alternate displays and comment on issues of distraction and visibility.

4.2 Results

Results from the study are organized below according to our two evaluation goals – testing the effectiveness and usability of IPV in comparison with Privacy Bird, and exploring design issues in the presentation of fine-grained conformance information.

4.2.1 Effectiveness and Usability

We wished to determine whether the finer-grained presentation of conformance information would assist users in determining the source and origin of conflicts, and would make them more confident in deciding whether to disclose personal information to the site. We looked at participants' success in finding conflicts, their success in determining the source of the conflicts, and their opinions and preferences as stated on the questionnaire.

Success in Finding Conflicts

We first looked at whether participants were able to determine that conflicts existed. Table 1 shows the number of participants of the six who correctly reported the presence of conflicts in each of the three pages (there were conflicts in pages one and two).

Table 1: Number correctly determining existence of conflict

	Page 1	Page 2	Page 3
AT&T	3/6	6/6	2/6
IPV	6/6	6/6	2/6

All participants correctly noted the conflict on page two, but there were problems on the other pages. Three participants did not notice that the Privacy Bird was red on the first page of the website. Since these errors occurred after training with the agent,

this result suggests that that indicators in the browser frame may not always be noticed when people are focused on the transaction.

Page three – which had no input fields – was the most difficult for the participants to assess. Only two participants correctly understood that Privacy Bird's red icon was indicative of conflicts elsewhere on the site, and that the lack of fields on page three meant that there could be no conflict here. Similarly, only two participants understood that the lack of any IPV indicator meant that there were no conflicts. This result suggests that a global indicator should always be used to indicate the conformance result for the page as a whole.

Success in Determining Source of Conflicts

We next looked at whether participants were able to determine the source of conformance conflicts. Table 2 shows how many of the six participants correctly determined the element of personal information that was the cause of a conflict. There were five pieces of information that were the source of a conflict (email, name, address, home phone, and credit card number), and one piece not in conflict (business phone).

Table 2: Number correctly determining conflict source

	Email	Name	Addr.	Home Phone	Credit Card	Bus. Phone
AT&T	4/6	6/6	6/6	6/6	6/6	2/6
IPV	6/6	6/6	6/6	6/6	6/6	6/6

Again, participants were mostly correct, but a few problems were observed. Even though the policy information was available through the Privacy Bird, some participants still had difficulty determining where the conflicts arose. For example, two participants thought that both pieces of information requested on the first page of the *AllThatJazz* site were in conflict, when actually only the email field was a problem. On page two, four users of PrivacyBird incorrectly identified business phone as being in the cause of a conflict, although the privacy policy was in conformance for this piece of information. The two participants who (correctly) did not report business phone as a problem had to thoroughly review the detailed privacy policy to make this determination.

Opinions and Preferences

The questionnaire asked participants several questions about which agent they preferred, which (if any) agent distracted them from the task, and which if any helped them be more confident in deciding whether to disclose personal information to the site. Table 3 shows a summary of the responses.

Table 3: Participant preferences

	AT&T	IPV
Which system made it easier to find and understand conflicts?	0	6
Which system did you prefer overall?	0	6
Did either system distract you from the task?	0	2
Did either system increase your confidence?	2	6

For the preference questions, all six participants preferred IPV over Privacy Bird. The main reasons given for the preference were the visibility of the conformance icons and the location of the icons next to the input fields.

Two participants stated that they were distracted by the additional icons of IPV, whereas none of the participants were distracted by the Privacy Bird indicator. Upon further discussion, the two participants stated that their distraction was primarily due to their

initial interest in the system, and their interest in exploring its capabilities. We believe that this novelty effect will wear off, but the issue of distraction should be studied further.

Our final question asked whether either system helped the participant to feel more confident about deciding whether to disclose personal information to the website. Both agents increased confidence, although IPV was effective for more people (six as opposed to two). Note that this result does not state that users are more likely to complete a web transaction, only that they can more confidently determine whether or not they would proceed.

4.2.2 Visual Design Feedback

We showed participants several alternate designs for the conformance display, looking at different ways of placing icons and grouping fields (e.g., Figure 14). The current IPV agent puts an icon next to every input field to indicate conformance. We were interested in whether participants would see this as adding clutter to the page, whether they would have difficulty with icons that were below the page fold, and how they would interpret the lack of conformance icons on fields and pages.

Figure 14. Two different ways to group fields that relate to the same piece of PII. IPV currently places an icon next to each field.

From our discussions with the participants, we found that clutter was not a major design issue. Participants liked the fact that each field was identified, even when there was no conformance conflict. Participants also preferred this representation over alternate designs where fields relating to the same piece of PII were visually grouped together. In addition, conformance icons appearing below the fold was not seen to be a problem. The form-filling task does not let the user submit the form until all required fields are complete. This necessitates field inspection, so no conformance icons were missed.

However, interpreting the absence of a conformance icon was more problematic. At the field level, participants were unsure whether the lack of an icon meant that the field was in conformance, that the information was not covered by the privacy policy, or that there was an error in the system. Similarly, people had difficulty determining what it meant when a page had no input fields (and thus no indicators). Several participants stated that a global indicator should be present, even when no information is retained or requested by the page. These participants also felt that the indicator should be in the web page rather than the browser border, in order to improve visibility.

5. DISCUSSION

The study suggests that both user agents are an improvement over none at all, but that the fine-grained approach used in IPV can improve users' understanding of the existence and origin of conformance conflicts during web transactions. Here we look at the reasons why IPV was successful, how our results generalize to real-world use, and what must be done by designers to make use of the approach.

5.1 Explanation of Findings

It was clear that participants liked IPV and the fine-grained approach. The study suggested four main reasons for this success:

- *Visibility of conformance icons*. Web transactions require users to fill in input fields; during this task, their attention is drawn to the input field and not to the web browser frame. Clutter was not seen as a problem, although we plan further work to look more closely at distraction issues.

- *Explicit indication of conformance*. IPV provides a field-by-field indication of privacy policy conformance or conflict; there is no ambiguity about whether a specific piece of personal information will be treated as the user wishes.

- *Fast access to detailed conformance information*. The field-by-field icons provide a quick way (only two steps) to get further information through mouse-over.

- *Visibility of input field*. IPV presents additional conformance information without obscuring the input field; the user can review the information they are entering while at the same time viewing the associated privacy policy statement.

IPV improves the understanding of website privacy policies by reducing user effort. As one user put it, IPV "removed the fine print and made it obvious what information the privacy policy statement was talking about." As a result, IPV gave users more confidence in deciding whether to complete transactions.

5.2 Generalization of Results

Although the study was small, there are reasons to suggest that our results will generalize to web transactions in real-world task situations [21]. The evaluation was similar to a real-world consumer experience: tasks were modeled on real Internet activities, and the participants represented a good sample of typical Internet users, with experience in filling out web forms and making web-based purchases.

However, participants did not carry out the task using their own personal information. We believe that this will increase their interest in IPV rather than decrease it, but this should be tested further in future studies.

In addition, the evaluation did not model long time experience with IPV, and increased experience may change people's attitudes, particularly towards the visual presentation of the conformance icons. As the novelty wore off they might prefer to have fewer icons on the page or different presentations (e.g. smaller icons or highlights) to reduce visual clutter. Permitting customization of visual style and presence would let IPV better serve both the casual and expert user.

5.3 Deployment of a Fine-Grained Approach

The use of a fine-grained privacy system like IPV affects designers of both privacy policies and web pages. First, the fine-grained approach requires policy designers to refine their policies so that each statement is correctly bound to a personal data group or element. This is not a major change, since some P3P policy development tools already support this level of binding (e.g., IBM's P3PPolicyEditor [22]).

Second, web page designers need to use the p3pdataelement attribute to link statements from the P3P policy with input fields. Many HTML editors (such as Eclipse [23]) could easily be extended to read the P3P policy and give the page designer a list of data elements to use as the value of the p3pdataelement attribute. This would minimize the effort needed to implement

IPV in the web page, and would add the benefit of cross-checking the privacy policy with the actual information to be collected.

The web page designer also needs to be aware that the privacy agent will be modifying the displayed web page. For example, placing a conformance indicator in the page may cause some web pages to render incorrectly if the input fields are so close together that they are not visibly separated. The web page designer implementing this solution should take into consideration the range of possible visualizations that privacy agents might use to make sure their page will appear correctly. Again, the ability to customize the location and representation of conformance indicators can provide adequate flexibility for the page designer.

Finally, it is clear that fine-grained policy anchors will not be added to web pages until user agents are common in current browsers. Therefore, part of our future work in this area is to develop plug-ins for the major browsers, and to build a demonstration site where consumers and organizations can see the idea and how it works. We believe that the fine-grained approach has enough value for consumers that it has a reasonable chance of becoming adopted, in a similar way to RSS or even P3P itself.

6. CONCLUSIONS

Machine-readable privacy policies and preference documents enable user agents to interpret conformance and alert the user to privacy conflicts. However, the coarse granularity of current systems means that it is still difficult for users to determine the cause and origin of conformance conflicts. We developed the Integrated Privacy View, a system that uses fine-grained policy anchors to link fields on web forms to specific statements in P3P privacy policies. The IPV user agent uses the fine-grained anchors to present visual conformance information in the context of the specific input fields. In a user study, we found that participants preferred the fine-grained approach to the current state of the art, and were better able to determine the existence and source of privacy conflicts.

Future work on IPV will include both short-term improvements and longer-term investigations. We will continue to improve the system itself by building browser plug-ins, adding a global indicator, adding capabilities for customizing the visual effects, and developing tools to make it easy to insert the policy anchors needed for the fine grained approach.

In future, we will look at other ways of reducing user effort in the privacy area. In particular, we plan to support the process of building up a user's preferences, something that IPV does not currently address. One approach is to extend the P3P user agent to permit it to join an agent community. The user agent would keep track of privacy policies the user had approved or declined, and report that information anonymously back to the community. A community-based system could then provide assistance to a user when they arrive at an unknown vendor site – for example, by stating whether other people have accepted the policy. Users can thus be given guidance on whether to trust particular vendors and on what constitutes an acceptable privacy policy, allowing them to incrementally build up their own set of privacy preferences.

7. REFERENCES

[1] W3C (2002) *P3P 1.0 Recommendation*, www.w3.org/TR//P3P, accessed May 15, 2004

[2] W3C (2002) *APPEL 1.0 Working Draft*, www.w3.org/TR//P3P-preferences, accessed May 15, 2004.

[3] AT&T Corp. (2002) *Privacy Bird*, www.privacybird.com.

[4] Earp, J., and Baumer, D. (2003) Innovative Web Use To Learn About Consumer Behavior and Online Privacy, *CACM*, Vol. 46, No. 4, 81-83.

[5] Boyle, M., and Greenberg, S. (in press) The Language of Privacy: Learning from Video Media Space Analysis and Design. *ACM TOCHI*, in press.

[6] Westin, A. (1967) *Privacy and Freedom*. New York, NY: Bodley Head, 1967.

[7] European Union (2000) *On the protection of individuals with regard to the processing of personal data*. Council of the European Union Act No. 77, 2000.

[8] Aberdeen Group (2002) *Federated Identity Systems*, Technical Report, Aberdeen Group, Boston, MA, 2002.

[9] Shneiderman, B. (2000), Designing Trust Into Online Experiences, *CACM*, Vol. 43, No. 12, 57-59.

[10] Siau, K., and Shen, Z. (2003) Building Customer Trust in Mobile Commerce, *CACM*, Vol. 46, No. 4, 91-94.

[11] Behrens, L. (2001) Privacy and Security: The Hidden Growth Strategy. In *Gartner G2 Report*, 2001.

[12] Privacy Commisioner of New Zealand (2001) *Privacy Concerns Loom Large*, www.privacy.org/nz/privword//42pr.html, accessed May 15, 2004

[13] Fox, S. and Rainie, L. (2000) Trust and Privacy Online: Why Americans Want to Rewrite the Rules. *Pew Internet & American Life Report*, www.pewinternet.org/reports/, accessed May 15, 2004

[14] Reagle, J., and Cranor, L.F. (1999) The Platform for Privacy Preferences. *CACM*, Vol. 42, No.2, 48-55.

[15] Ipsos-Reid and Columbus Group (2001) Privacy Policies Critical to Online Consumer Trust. *Canadian Inter@active Reid Report*, www.ipsos-na.com/news/pressrelease.cfm?id=1171, accessed May 15, 2004

[16] Culnan, M., and Milne, G. (2001) The Culnan-Milne Survey on Consumers & Online Privacy. In *Get Noticed: Effective Financial Privacy Notices*, 2001.

[17] Hoffman, D., Novak, T., and Peralta, M. (1999) Building consumer trust online. *CACM*, Vol. 42, No. 4, 80-85.

[18] Jensen, C., Potts, C. (2004) Privacy Polices as Descision-Making Tools: An Evaluation of Privacy Notices. *Proc. ACM CHI 2004*, Vienna.

[19] Cranor, L.F. (2002) *Web Privacy with P3P*. Cambridge: O'Reilly and Associates, 2002.

[20] Constantine, L. (2002) Devilish Details: Best Practices in Web Design. *forUse 2002*, www.foruse.com/articles/details.pdf , accessed May 15, 2004.

[21] Barnum, C. (2003) Usability Interface – What's in a Number?, *STC Usability SIG Newsletter*, January 2003, Vol 9, No. 3

[22] IBM Corp. (2003) P3P Policy Editor, www.alphaworks.ibm.com/tech/p3peditor.

[23] Eclipse Editor (2004), www.eclipse.org.