Improving Understanding of Website Privacy Policies

A Thesis Submitted to

The College of Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the

Department of Computer Science

University of Saskatchewan

Saskatoon, Saskatchewan

By

Stephen E. Levy

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A9

Abstract

Machine-readable privacy policies have been developed to help reduce user effort in understanding how websites will use personally identifiable information (PII). The goal of these policies is to enable the user to make informed decisions about the disclosure of personal information in web-based transactions. However, these privacy policies are complex, requiring that a user agent evaluate conformance between the user's privacy preferences and the site's privacy policy, and indicate this conformance information to the user. The problem addressed in this thesis is that even with machine-readable policies and current user agents, it is still difficult for users to determine the cause and origin of a conflict between privacy preferences and privacy policies. The problem arises partly because current standards operate at the page level: they do not allow a fine-grained treatment of conformance down to the level of a specific field in a web form. In this thesis the Platform for Privacy Preferences (P3P) is extended to enable field-level comparisons, field-specific conformance displays, and faster access to additional field-specific conformance information. An evaluation of a prototype agent based on these extensions showed that they allow users to more easily understand how the website privacy policy relates to the user's privacy preferences, and where conformance conflicts occur.

Acknowledgements

I would like to thank my thesis advisor, Dr. Carl Gutwin, for guiding me through my graduate work. Without his understanding and patience in dealing with the work habits of an adult learner, this thesis would never have been completed.

I would like to thank my beloved Carolyn for everything good in my life.

I would like to thank IBM for providing the financial assistance and leave from work to attend the University of Saskatchewan and to work on this thesis.

Finally, I would like to thank Dr. Stephen Boies for bringing me into the field of human computer interaction research over 20 years ago. I have never looked back.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABA	American Bankers Association
APPEL	A P3P Preference Exchange Language is an XML application to define a rule set to express a user's privacy preference.
cookie	A message given to a Web browser by a Web Server. The browser stores the message in a text file. This message is then sent back to the server each time the browser requests a page from the server
CSS	Cascading Style Sheets are attached to an HTML document to influence its layout when accessed via browser.
DOM	Document Object Model is a platform independent interface for accessing HTML and XML documents usable from within programming and scripting languages.
E-P3P	The Platform for Enterprise Privacy Practices
GBDE	Global Business Dialog on Electronic Commerce
HIPAA	Health Industry Portability and Accountability Act
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IE	Internet Explorer
IPV	The Integrated Privacy View is a privacy agent that utilizes the p3pdataelement to display privacy policy conformance on a web page.
MIME	Multipurpose Internet Mail Extension
OECD	Organization for Economic Cooperation and Development

P3P	The Platform for Privacy Preferences is an XML application to define the privacy policy for a website
P3PDataElement	The P3PdataElement is an attribute specified on an input field to associate a specific privacy policy statement with the input field.
PII	Personally identifiable information consists of attributes that may uniquely identify an individual. Examples of PII are a person's name, address, or social security number.
W3C	World Wide Web Consortium
WYSIWYG	What you see is what you get
XHTML	Extensible Hypertext Markup Language is a reformulation of HTML 4.0 in XML 1.0.
XML	Extensible Markup Language
XPATH	XML Path Language defines a search language for locating nodes in a document.
XSLT	XSL Transformations is a language for transforming XML documents into other XML documents
XSL	Extensible Stylesheet Language is a language for creating a style sheet that describes how data sent over the Web using XML is to be presented to the user.

Chapter 1 Introduction

Electronic commerce is a rapidly expanding market segment for consumer retail sales, and web-based transactions are now a common part of life on the Internet. There are now thousands of commercial websites for purchasing goods and services: these sites allow customers to search for merchandise, browse catalogues, choose and pay for selected items, and arrange for shipping and delivery. During these activities, e-commerce 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 the disclosed information is essential to e-commerce on the Internet. The user must feel confident that the personal information disclosed will be used only for agreed-upon purposes and will not be misused by the vendor. For example, users may agree to give their e-mail address in order to complete a transaction, but may still wish to prevent unsolicited e-mail or phone calls from telemarketers.

Human-readable privacy policies are now being installed 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 policies on the website does not guarantee understanding, since 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 complies with their personal privacy preferences.

To help reduce user effort in navigating website privacy policies, a new technology called the Platform for Privacy Preferences (P3P) was developed to provide machine-readable privacy policies. A P3P Preference Exchange Language (APPEL) was developed to provide a

machine-readable rule set to express the user's privacy preferences. P3P and APPEL allow vendor privacy polices and user privacy preferences to be compared automatically for conformance. This approach reduces the effort needed by the users to determine if their desired privacy preferences are matched by the website.

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 present a visual indication of site conformance in the browser's title bar and provide detailed conformance information in a separate window that is reached by opening a menu on the indicator icon. However, both the general and detailed information is separate from the web page itself, requiring the user to interpret the conformance information appearing in one window separately from the input display window. Although current user agents do provide the information, the additional effort required in interpretation reduces the user's understanding of conformance and lessens their understanding of how their personal data will be used.

1.1 Problem

The problem addressed in this thesis is: **It is difficult for a user to determine the cause and origin of a conformance conflict between the user's privacy preferences and a website's privacy policy**.

Conformance conflicts arise when the vendor's privacy policy does not match the user's privacy preferences. For example, a vendor's policy may state that the user's e-mail address can be provided to third parties, but, if the user's privacy preference states that their e-mail address should only be used for the purpose of completing the current transaction, then there is a conformance conflict. In this situation, a privacy agent would display an indication of non-conformance when the user views any page from the website. The cause of the

conformance conflict is the release of the user's e-mail address to third parties, and the origin of the conflict is the input field associated with the e-mail address. However, the cause and the origin are difficult to determine with current agents.

The main issue in this problem is that user privacy agents have a coarse view of the vendor privacy policy. There is no machine-readable connection between the privacy policy statement and the specific input field of the applicable privacy statement. The implication of this coarse view is that conformance information presented to the user has no visible link to the specific context of the input field presented on the web page. Current user agents only present general statements related to the current policy for a web page. For each conformance conflict associated with a web page, the user will have to interpret which conflict statement applies to which input field. This interpretation requires additional effort by the user to understand the conflict, and makes it more difficult for them to decide whether to finish the transaction.

1.2 Motivation

The main motivation for improving understanding of privacy conformance is to improve user confidence and trust in the disclosure of personal information. E-commerce requires the disclosure of personal information in order to complete a purchase or transaction. User confidence and trust in personal data disclosure is partly based on understanding how personal data will be used and protected by a website. Increasing confidence and trust leads to increased completion of transactions since fewer users will abandon the transaction due to uncertainty about how their personal data will be used. Increasing the percentage of completed transactions leads to increased customer satisfaction and increased business.

1.3 Solution

The solution explored in this thesis is that **user understanding can be improved by visualizing privacy preference conformance at the input field level.** This solution has two main parts: first, refining the mapping of privacy policy to input fields, and second, providing a contextual display of conformance on the web page.

Finer-grained mapping of privacy policy to input fields makes it possible for vendors to indicate the relationship between an input field on a web page and the specific privacy policy statement that is related to that field. A web page developer is thus able to specify attributes on the input field whose values link the input field to the associated privacy statement in the vendor privacy policy. A user agent can then use the value of these attributes to select the appropriate policy statement for the conformance evaluation.

Fine-grained mapping also permits a contextual display of conformance. Since the web page developer is able to identify exactly which input field a conformance result should be associated with, a user agent is able to insert a conformance artifact into the web page such that it is easily associated with the source field. For example, the user could be able to view a conformance-indicating icon next to each input field.

This solution is embodied in an Integrated Privacy View (IPV) user agent. IPV is an enhanced privacy user agent that maps privacy preferences to site privacy policies and augments the vendor web page with an integrated visual privacy conformance display. Integrating privacy conformance and web pages at a fine-grained level aids understanding in three ways. First, visual indication of a conflict can be physically located adjacent to the input field. Second, detailed information about the conflict can be made accessible with a single button click. Third, detailed information about the conflict can be made specific to the input field.

The goal of this solution is to reduce user interpretation effort. Placing the conformance indicator adjacent to the input field allows the user to continue to focus on the work area without being concerned about peripheral indications in the browser title bar. The physical presence of the indicator is an immediate visual clue to the user that a particular input field, (for example, a field asking for an e-mail address), has a privacy conformance issue. Field-specific detail is provided directly (for example, by clicking on the conformance indicator associated with the field), allowing the user to inspect the privacy preference specifically associated with this PII.

1.4 Steps in the solution

The four steps in this solution are listed as follows:

1. *Design a machine-readable link between HTML input forms and privacy policy statements.* The implementation of P3P is extended by expanding the HTML schema to add the P3PdataElement attribute to the input element. The P3PdataElement attribute is used as a key-value pair in which the value is taken from the P3P Data Schema. When inserted into the HTML of a web form, the P3PdataElement attribute can uniquely identify the relevant privacy policy statement and the location of the input field in the HTML.
2. *Create an example conformance visualization technique.* The P3PdataElement defined in Step 1 enables an adaptive presentation of conformance by a privacy agent. A specific visualization technique that places a conformance indicator next to each input field was chosen through discussion with representative users and was tested for basic usability.
3. *Build an IPV user agent.* The IPV user agent examines each web page presented by the server to the client. Web pages implementing the P3PDataElement attribute on input

fields are checked for conformance with the user's privacy preferences. The conformance result is then inserted into the web page next to the specified input field. Finally, the browser renders the page with the conformance indication placed next to the target input field.

4. *Develop a website that implements the P3P extension*. An e-commerce website was simulated which utilizes the extended P3P implementation to be able to compare an IPV user agent with a current P3P user agent. The website implemented a privacy policy reference file and privacy policies typical of current best practices in this area.

1.5 Evaluation

The idea of fine-grained integration was evaluated through a comparison study between IPV and a user agent that uses the current P3P implementation standard. The goal of the study was to see if IPV leads to better understanding of privacy policy conformance. The study measured user understanding of privacy conformance during e-commerce transactions, and measured the user's perceived effort and efficiency in finding the privacy conformance information for conflict situations. Measures of user understanding of privacy conformance were based on whether the user could determine when a conformance conflict had occurred, and whether they could determine which input field generated the conflict. Measures of perceived effort were based on recording the steps the user took to understand and resolve the conflict.

The evaluation showed that locating a privacy conformance indicator with the input field and providing context-sensitive privacy information in the web page itself does improve the understanding of website privacy policies. The study also showed that there are several visual design issues that must be taken into consideration when implementing the conformance visualization, such as a positive global conformance indicator for web pages that

do not have conformance conflicts and the need to show explicitly the conformance of every input field.

1.6 Contributions

The major contribution of this research is the design and demonstration of a scheme that allows fine-grained comparison and display of privacy conformance information for web transactions. The scheme extends the machine-readable privacy policy standard (P3P) to permit the integration of policy statements at a fine-grained level with the vendor's input form. This allows web form designers to link specific HTML input fields with specific privacy policy statements. This link improves users' understanding of website privacy policies by permitting an adaptive presentation of conformance through a better visualization of privacy policy and related conformance information.

There are also several secondary contributions:

- A reference implementation of the IPV prototype, which shows how a user agent can take advantage of the improved P3P implementation specification, and shows that the fine-grained approach can be carried out without compromising browser performance and without adding undue complexity to design of the user agent.
- A better understanding of the visual issues involved in presenting conformance visualizations. This thesis documents user requirements of when to show conformance indication, where to show conformance indication, and how to present additional conformance information.
- A better way for the privacy policy designer to collaborate with the web page designer. Implementation of the P3P attribute allows the designers to ensure the privacy policy is

complete and reflects all personal data gathered. Missing or inaccurate privacy policy statements will be detected during web page implementation.

- A better understanding of how P3P agents can be characterized as adaptive hypermedia agents. The IPV agent is the first adaptive hypermedia P3P user agent to adapt privacy meta-data into the HTML content presented to the user.

1.7 Thesis Outline

The thesis is organized as follows:

Chapter 2 reviews previous work in several areas that underlie this research. These include privacy, trust and privacy in economic exchange, privacy policies, privacy policy specifications, P3P user agents, user adaptation, and conformance visualization in privacy agents.

Chapter 3 describes the design and implementation of the P3PdataElement attribute, the HTTP Proxy that permits the intercept and modification of HTML destined for the user's browser and the IPV privacy agent. The chapter also describes the way in which the web page is adapted to present the conformance indication.

Chapter 4 describes the evaluation of IPV. Six typical Internet users were asked to perform a web-site registration task using both IPV and the AT&T Privacy Bird. The evaluation looks at the utility of IPV, the design of the conformance visualization, and the ways that the overall scheme may be used by the privacy policy designer, web page designer, and privacy agent developer.

Chapter 5 presents a summary of the research and contributions of this thesis. Future directions are discussed for improving IPV and for changing IPV from a user adaptable system to a user adaptive system.

Chapter 2 Review of Literature

This chapter reviews previous work in several areas that underlie the proposed research. These areas include privacy, trust and privacy in economic exchange, privacy policies, privacy policy specifications, P3P user agents, conformance visualization in privacy agents, and adaptation in hypermedia systems.

2.1 Privacy

There are many definitions and dimensions to privacy (Westin, 1967; Clarke, 1997; Boyle and Greenberg, 2003; Shneiderman, 2000). In general, however, privacy is the right of an individual to maintain a personal space, free from interference by other people and organizations (Boyle and Greenberg, 2003). The definitions explored in this section refer particularly to information privacy: the right to understand and control the conditions under which information about an individual is made available (Boyle and Greenberg, 2003).

One popular definition of information privacy is the right of individuals to determine for themselves when, how, and to what extent information about them is communicated to others (Westin, 1967). In particular, this thesis will focus on Personally identifiable information (PII). PII is defined by the European Union (2000) as data that is associated with an identifiable person. For example, personal information includes name, addresses, phone numbers, membership in groups, relationships to other people, financial data, buying history, and web transaction logs.

Altman (1975) describes privacy as a boundary control process regulating access to the self. Altman describes three genres of control; solitude, confidentiality, and autonomy. Originally applied to interactions between people and groups, these control processes need to be present when applied to the online exchange of information (Boyle and Greenberg, 2003).

Solitude is control over one's interactions with others. Individuals need to be able to define when and to what extent other people have access to them. On the Internet, control over solitude is present when a web form is completed, but the user may worry that their solitude may be violated as a result of personal information being provided. For example, an unsolicited e-mail is one example of the violation of solitude.

Confidentiality is control over access to information about the self; that is, the ability to control what information is disclosed and under what circumstances. People are concerned about the sensitivity of the information divulged and to whom the information will be revealed. For example, a person divulges information to another person based on that person's reputation and on an agreement to respect the confidentiality and use of that information.

Autonomy is control over one's own actions, the freedom to choose how one interacts with the world. Autonomy is limited by economic, political and social constraints. Understanding the environment that an individual interacts with and the response of that environment to an individual's actions, control how much autonomy an individual has.

To have privacy, a person must be able to predict how his actions will drive interactions, information access, and behaviour in chosen ways (Boyle and Greenberg, 2003). This prediction is based on the reputation and promises made by the party to which the information is disclosed.

The definition of privacy for Internet transactions is the positive expectation a person has for another person or an organization based on past performance and truthful guarantees (Shneiderman, 2000). Individuals maintain relations with companies that have delivered goods and services that have met their expectations in the past. From this experience the individual expects that the quality of this relationship will continue in the future. This implies that successful Internet privacy policies rely on truthful and accurate guarantees by a vendor and

third party enforcement of those guarantees to create a positive expectation of privacy for the user. Privacy policies implemented in P3P are designed to provide the accurate and truthful guarantees by the vendor and provide a mechanism to indicate the third party responsible for enforcing those policies.

2.2 Trust and Privacy in Economic Exchange

Trust in e-commerce is dependent on many complex factors, including consumer rights, freedom of expression, and social equity (Clarke, 1999). Central to each of these factors is the element of privacy. Without the proper expectation of privacy for a given exchange, the exchange will not take place.

An economic exchange only takes place between two parties that trust one another to fulfill their obligations in a timely and efficient manner (Aberdeen Group, 2002; Shneiderman, 2000; Siau and Shen, 2003). That trust is usually based on personal knowledge, references, or past experience. Before the Internet, this trust was built up through interactions with vendors and their representatives in markets in the consumer's local area. The probability of having personal knowledge of trustworthy partners was high.

The Internet changed the local market into a global one where the likelihood of personal experience with a vendor is limited. Knowledge of the vendor is often only available through the commitments and guarantees made on the vendor's website. These commitments and guarantees express not only how the order or service will be fulfilled but also how the information required to complete the transaction will be used. The vendor's reputation will be based on how well the obligation is fulfilled, but also on how well the vendor protects and respects user privacy.

The results of consumer surveys on privacy and security on the Internet show that individuals are very concerned about disclosing personal information.

For example:

- Several studies show that a large majority of Internet users are concerned about the security of personal information; proportions that have been reported include 83% (Cyber Dialogue, 2001), 70% (Behrens, 2001), 72% (UMR, 2001), and 84% (Fox and Rainie, 2000).
- Two studies showed that two-thirds of individuals are unwilling to shop online because of privacy concerns: 66% (Ipso Reid, 2001) and 64% (Culnan and Milne, 2001).
- 27% of consumers polled had abandoned online shopping carts because of privacy reasons (Cyber Dialogue, 2001).
- Two studies report that individuals are concerned that a business will use their data for other purposes, such as telemarketing and spam: 91% (UMR, 2001), 90% (Roy Morgan Research, 2001).

Privacy laws and regulations are designed to protect the privacy of large groups of individuals but do not reflect a specific individual's privacy preferences. This requires the user to be responsible for maintaining their privacy. If users are to be successful in this, vendors need to provide information about the ways that collected PII will be used, how long the information will be maintained, and to whom the information will be disclosed. This information becomes a promise to the user that the user will evaluate against their need for privacy. The potential for this type of informal contract has been demonstrated in an empirical survey (Hoffman, Novak and Peralta, 1999) where over 72% of web users said they would give websites their personal information if the sites would only provide a statement regarding how the information collected would be used.

2.3 Privacy Policies

There are three ways to address privacy issues in an electronic transaction: through law, self-regulation, or technology (Spiekermann, Grosslags, and Berendt, 2001). In each model, a privacy policy can be used to embody the commitments and guarantees that a vendor makes to protect personal data provided by the consumer. In the current worldwide web, governments, industry, and independent global consortiums have encouraged companies to define their practices for handling and sharing personal information, including reasonable communication of these policies to individuals. A recent survey of the most popular websites found that 77% of those websites posted a privacy policy (Adkinson, 2002).

Governments regulate privacy through law. For example, the Health Industry Portability and Accountability Act (HIPAA) in the United States requires privacy policies for collection, storing and disseminating of individually identifiable health information (Congress, 1996). Laws can protect the privacy of a group, but may not meet the needs of the individual; in addition, law often lags behind the needs of current systems and situations.

Industries and consortia also regulate privacy concerns through voluntary practices, and, to date, most of the guidelines and rules that have been produced have come from these sources. For example, the American Bankers Association (ABA) has established privacy principles to be followed by its members (ABA, 2000). Global consortiums establish privacy policy guidelines that span both industries and countries. The two most generally accepted guidelines for dealing with PII come from the Organization for Economic Co-operation and Development (OECD, 1980) and the Global Business Dialog on Electronic Commerce (GBDE, 2001).

The OECD guidelines shown in Figure 2.1 provide a basic set of fair information practice principles, a set of principles that is independent of industry and government. These

guidelines have been utilized as the framework for industry, government and international privacy policies. P3P, discussed in the next section, was designed to capture the spirit of the OECD guidelines as they apply to the use of PII.

There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.
Personal data should be relevant to the purposes for which it is to be used, and, to the extent necessary for those purpose, should be accurate, complete, and up-to-date.
The purpose for which personal data is collected should be specified no later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes.
Personal data should not be disclosed, made, available, or otherwise used for additional purposes, except with the consent of the data subject or by the authority of law.
Personal data should be protected by reasonable security safeguards.
There should be a general policy of openness about developments, practices, and policies with respect to personal data.
An individual should have the right to obtain data about himself, and the right to challenge data about himself.
A data controller should be accountable for complying with measures that give effect to the principles

Figure 2.1: The OECD Principles (OECD, 1980*)*

Two decades after the development of these initial principles, the GBDE personal data privacy protection guidelines were defined to embrace the OECD guidelines and extend them to the online environment of e-commerce. The GBDE guidelines deal specifically with online data collection in the context of the Internet and provide enough detail to allow implementation on websites. For example, this guideline specifies what online information gathering may be done, how to provide notice of data gathering to consumers, request consumer consent to gather data and use of data by third parties. A key aspect of the above guidelines as applied to online transactions is that privacy policies need to be easy to locate and read (Shneiderman, 2000). The GBDE guidelines specify that the privacy policies, like the

one shown in Figure 2.2, should be available by a link from all web pages that collect data, and should be written in plain language.

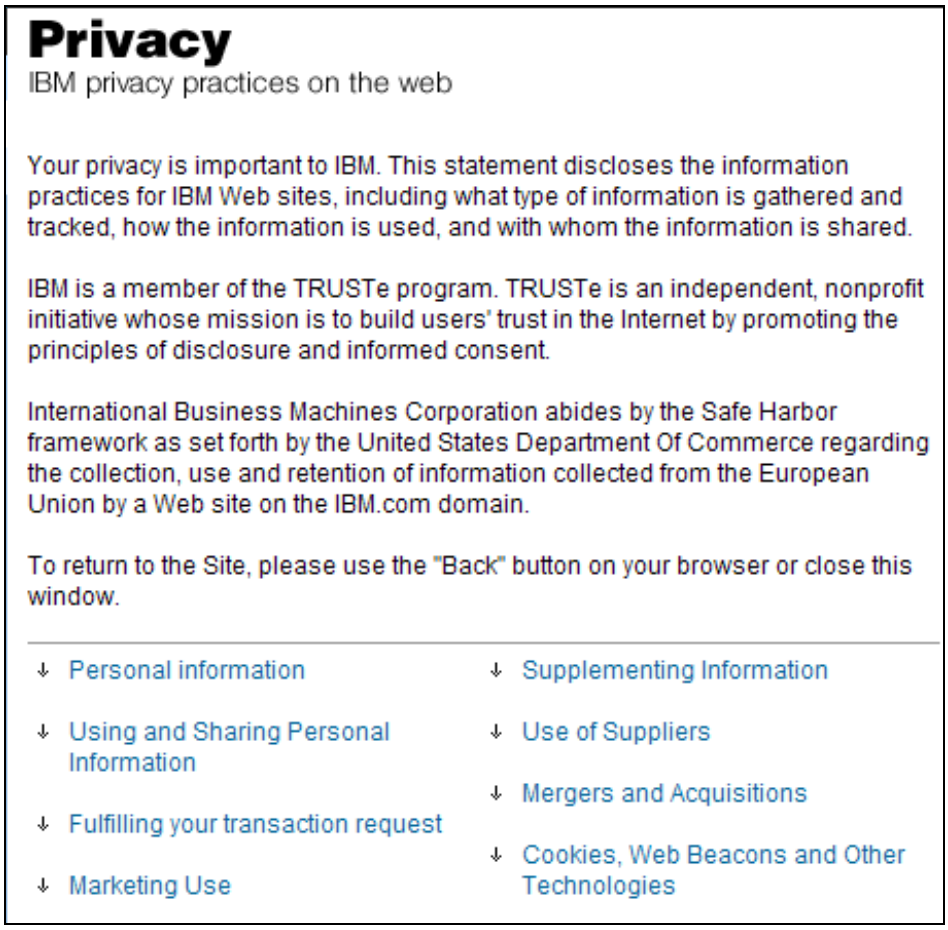


Figure 2.2: First page of privacy policy for IBM DeveloperWorks

These guidelines are important because surveys indicate that most consumers expect to see detailed information regarding the privacy policy when visiting commercial websites (Furnell and Karweni, 1999; Hoffman, Novak, and Peralta, 1999). However, for the privacy policy to be effective in increasing trust and confidence, it must be read, and these same surveys indicate that only 54% of respondents indicated they would read the privacy policy on first visiting the website (Earp and Baumer, 2003). For example, the privacy policy shown in Figure 2.2 is over eight pages in length, requiring considerable effort by a user to evaluate.

The privacy policy must also be written in a language that is readily understood by the user, and must address the privacy issues that the user is interested in (Jensen, 2004).

Therefore, attempts to increase consumer trust in e-commerce are limited by the amount of effort consumers are willing to invest in understanding how their personal data will be used. The third type of control – regulation by technology – has the capability of tailoring privacy policy understanding to the needs of the individual. One way that technology has been used to reduce effort is in machine-readable privacy policies that can be interpreted by a user agent. Machine-readable policies will be discussed in the next section.

2.4 Privacy Policy Specifications

Technology can reduce the effort required for a user to understand a website's privacy policy. 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 and each aims to support the individual, the corporation, and the communication between the two.

There are two main types of policies: those that deal with the corporate end of the relationship, and those that deal with the individual's concerns. Standards for *corporate* policy definition vary in complexity. For example, 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 (W3C, 2002a). P3P is designed to define privacy policies for an organization's clients and consumers. The privacy policy in P3P is a contract between an individual and the website owner. As an example of how P3P can be extended to focus on particular business-to-business concerns, IBM's Platform for Enterprise Privacy Practices (E-P3P) extends P3P to provide a privacy policy language for enterprises to document and enforce their internal practices for handling personal data of customers and employees

(Karjoth, Schunter, and Waidner, 2002; Ashley et al., 2002). E-P3P allows for the definition of roles for accessing data and for providing authentication tokens to secure access. For example, E-P3P would permit disclosure of a consumers address to the order fulfillment department, but not the consumer's credit card number.

Standards for *individual* privacy preference also vary in complexity. A P3P Preference Exchange Language (APPEL) allow users to specify what website privacy policies are acceptable and what actions to take to inform the user whereas IBM's *Individual Privacy Based Access Control* (Bohrer et al., 2003) describes how user preference information can be used to authorize actions on personal data, replacing traditional permission or role based access control.

This project will utilize the W3C P3P recommendation and the W3C APPEL working draft as the basis of the Integrated Privacy View (IPV), as these two types of policy are the current de-facto standard on the WWW. Details of the two specifications specific to this project are provided in the next two sections.

2.4.1 Platform for Privacy Preferences (P3P)

P3P is an XML language designed to describe the privacy policy of a website, so 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 (W3C, 2002a; Cranor, 2002; Byers, Cranor and Korman, 2003; Agrawal et al., 2003). P3P is now a widely known standard, and defines many of the basic concepts of private data usage, including purpose, retention and recipient.

P3P is a description language composed of several elements that describe different aspects of a site's privacy policy. An overview of the elements in a P3P privacy policy is

shown in Figure 2.3. The ENTITY element provides contact information for the business, organization, or person who owns the website. The ACCESS element describes what access an individual has to personal data maintained by the website. The DISPUTE element identifies how the individual resolves privacy related disputes. A key part of the DISPUTE element is the identification of third parties responsible for enforcing the website's privacy policies. PII is completely described by the DATA, PURPOSE, RECIPIENT and RETENTION elements. These elements are combined to create a privacy STATEMENT such as the example shown in Figure 2.4. Each policy may contain one or more privacy statements, each relating to different PII data elements.

ENTITY – contact information for the business, organization, or person who owns the site

ACCESS – whether individuals can find out what personal data a site keeps about them in its databases (6 types of access policies are specified)

DISPUTE – how to resolve privacy-related disputes with the site (customer-service desk, privacy seals, relevant privacy laws, etc.); also includes REMEDIES sub-element

DATA – the kinds of data collected (17 data CATEGORY elements and dozens of specific data elements are specified)

PURPOSE –how collected data is used (11 types of purposes and "other-purpose" are specified), and whether individuals can opt-in or opt-out of these uses

RECIPIENT – whether and under what conditions data may be shared and whether there is an opt-in or opt-out (6 types of recipient policies are specified)

RETENTION – policies for periodic purging of collected data (5 types of retention policies are specified)

CONSEQUENCE – human-readable explanation of site's data practices

Figure 2.3: P3P policy elements overview (Reagle and Cranor, 1999)

The privacy STATEMENT (line 1) in Figure 2.4 states that the vendor makes the following statements about the PII consisting of the user's name (line 19) and home address (line 20). The PURPOSE of collecting this PII (line 2) is to do website and system administration (line 3), to contact you through means other than telephone to interest you in other products or services (line 4), to complete the activity for which the data is provided (line

5), and to contact you by telephone to interest you in other services and products (line 6). The RECIPIENT of this PII (line 9) is this website and the companies that help the site provide services to you (line 10) and delivery companies who may use your information for other purposes (line 11). The RETENTION period for this PII (line 14) is indefinitely (line 15).

```
1. <STATEMENT>
2.   <PURPOSE>
3.      <admin/>
4.      <contact/>
5.      <current/>
6.      <telemarketing />
7.   </PURPOSE>
8.
9.    <RECIPIENT>
10.     <ours/>
11.     <delivery/>
12.   </RECIPIENT>
13.
14.   <RETENTION>
15.     <indefinitely/>
16.   </RETENTION>
17.
18.   <DATA-GROUP>
19.     <DATA ref="#user.name"/>
20.     <DATA ref="#user.home-info"/>
21   </DATA-GROUP>
22. </STATEMENT>
```

Figure 2.4: Example privacy policy statement

The P3P 1.0 recommendation is a base XML application that is designed to be extended to enable the best possible interpretation of privacy policy by the user (Kaufman et al., 2002; Spiekerman, Grossklags, and Berendt, 2001). For example, a vendor may extend the specific data elements to cover PII not defined in the base specification.

2.4.2 A P3P Preference Exchange Language (APPEL)

APPEL is an XML language for specifying a user's privacy preferences as the set of web privacy policies that are acceptable to users, which can subsequently be matched against a P3P privacy policy to determine whether the website policy is acceptable, and how or whether to inform the user of the decision (W3C, 2002b; Byers, Cranor and Korman, 2003; Agarwal et al., 2003). The RULE element shown in Figure 2.5 encapsulates a P3P privacy statement to be

used for the matching against a website privacy policy statement. The behaviour attribute on the rule element indicates the action to take on a match and the description attribute provides an explanation for that action. The elements used in the matching process are from the P3P schema: STATEMENT, DATA, PURPOSE, RECIPIENT, and RETENTION. The elements may be combined using logical operators to create privacy preference rules.

The privacy RULE (line 1) in Figure 2.5 states the user's preferences about the use of their home address (line 5). If the PURPOSE (line 6) of collecting the PII is to contact the user through means other than telephone to interest them in other products or services (line 7) "or" (line 6) to contact the user by telephone to interest them in other products or services (line 8) then the behaviour of this rule is set to "limited" (line 1). A P3P user agent evaluating the privacy policy in Figure 2.4 with the privacy preferences in Figure 2.5 would determine that a conformance conflict exists. The user agent may then provide the warning to the user that "This site may share information that personally identifies you with other companies and telemarketers" (Figure 2.5, line 2).

```
1.  <appel:RULE behavior="limited"
2.      description="This site may share information that personally identifies you with
3.                   other companies and telemarketers" >
4.      <p3p:POLICY>
5.          <p3p:STATEMENT appel:connective="and" >
6.              <p3p:PURPOSE appel:connective="or" >
7.                  <p3p:contact />
8.                  <p3p:telemarketing  />
9.              </p3p:PURPOSE>
10.
11.             <p3p:DATA-GROUP >
12.                <p3p:DATA ref="#user.home-info"/>
13.             </p3p:DATA-GROUP>
14.         </p3p:STATEMENT>
15.     </p3p:POLICY>
16. </appel:RULE>
```

Figure 2.5: Example privacy preference rule

P3P and APPEL allow a user's privacy preferences to be compared with a website's privacy policy. The results of the comparison need to be presented to the user in a manner that

allows the user to make an informed decision about any conformance conflict. P3P user agents are the mechanism for this presentation.

2.5 P3P User Agents

A P3P user agent is a personal assistant (Maes, 1994; Ackerman and Cranor, 1999) designed to reduce the complexity of privacy policy information presented to the user. The goal of a P3P user agent is to help a user understand a website's privacy policy in relation to a user's privacy preferences, permitting the user to control whether or not to disclose their personal information. The user agent needs to accurately and simply present policy conformance to the user. However, the P3P specification does not provide any guidelines on how the policy and policy conformance should be displayed (Coyle, 2001) or how much of the P3P specification to implement. Therefore, agent implementers are responsible for determining when and how to show privacy policy conformance and related policy information. Agents may also be implemented with different levels of complexity (Kaufman et al., 2002).

Agent developers may decide to implement stand-alone agents utilizing only P3P and APPEL, or may decide to use multi-agent systems to utilize peer and expert opinion to shape user privacy preferences (Ackerman and Cranor, 1999). User agents can be implemented in a HTTP proxy (Internet Society, 1999), as a browser helper object (Esposito, 1999), or as an integral part of the browser.

P3P user agents may be further differentiated by how much of the P3P specification 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 (IE6) or all user provided data (AT&T, 2002). Conformance indication may be displayed in a separate window, in the frame of the browser, or in the web page itself.

Currently the two most widely used P3P user agents, IE6 and AT&T Privacy Bird, use the frame of the browser. These implementations are discussed in the next section. In contrast to these existing agents, the IPV agent adapts the actual web page to present the conformance indication in context.

2.5.1 The Internet Explorer 6 User Agent

The P3P specification defines an abbreviated version of a P3P policy, called a "compact policy," that can be transmitted in HTTP headers when cookies are set. IE6 uses the information in P3P compact policies to make cookie-blocking decisions (Cranor, 2002). In Figure 2.6, IE6 graphically displays an eye covered by a do not enter sign to indicate that the cookie policy of the vendor's site does not match the user's preferences. Additional information about the conflict is available by mouse click on the eye icon (as shown in Figure 2.7).

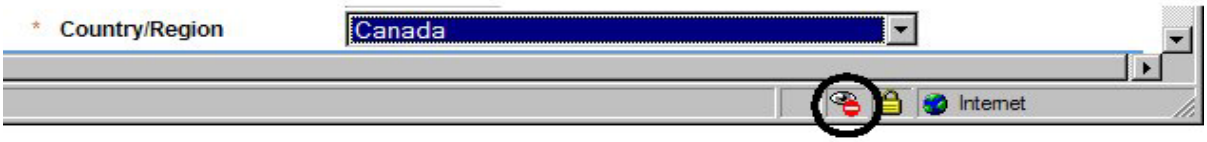


Figure 2.6: IE6 privacy icon displayed in browser frame

IE6 will also display the human-readable privacy policy and generate a privacy policy description from the P3P policy located on the website. Figure 2.8 shows the rendering of the P3P privacy policy for Doubleclick. Each element described in Section 2.4.1 is rendered. However, IE6 currently does not implement a full interface to gather user privacy preferences and compare them with the full P3P policy; only a brief dialog for preferences about cookies is available (see Figure 2.9).

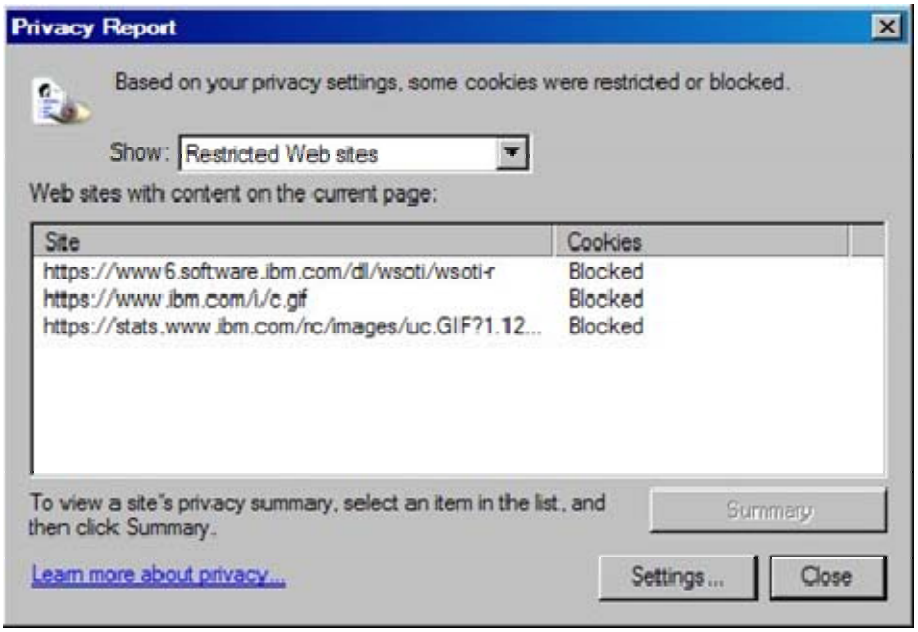


Figure 2.7: IE6 privacy report.

Figure 2.8: P3P Privacy Policy rendered by IE6.

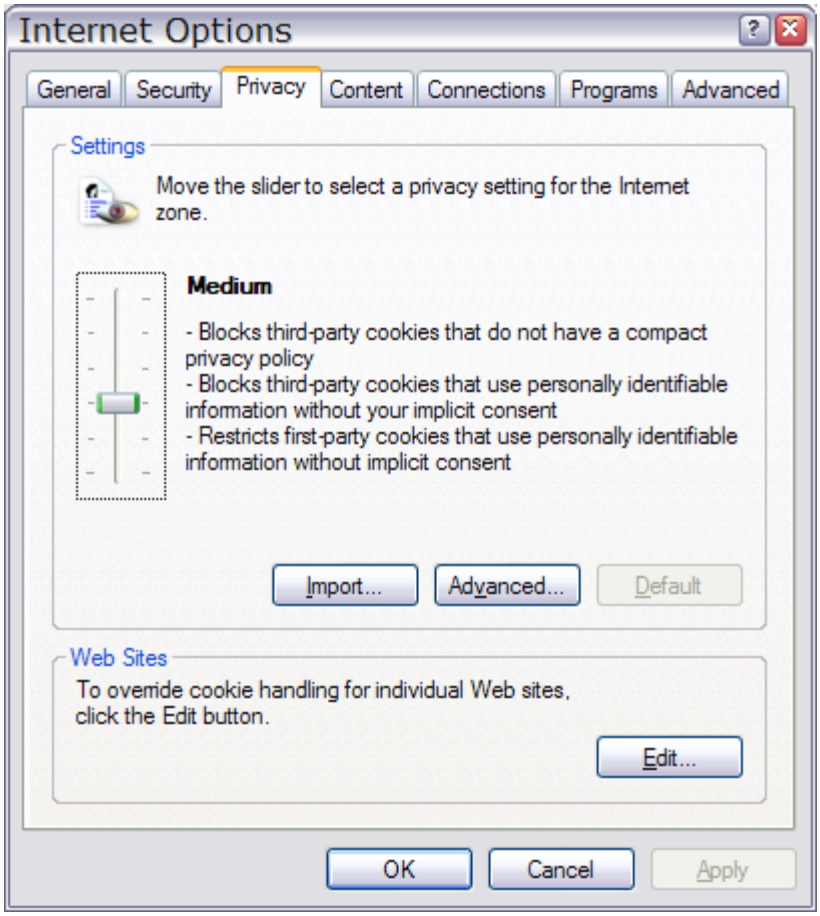


Figure 2.9: User preferences dialog for IE6 user agent.

Without a full set of user preferences, IE6 can warn the user about personal information stored in cookies, but not about the initial collection of PII in web forms. IE6 does not identify to the user the collection of the personal information when it takes place, regardless of whether that information will be used in a cookie or not. This might lead the user to the conclusion that if the site has an acceptable cookie policy, that their complete personal privacy preferences are being agreed to.

2.5.2 The AT&T Privacy Bird User Agent

The AT&T Privacy Bird implements the complete P3P and APPEL specification (Cranor, Arjula, and Guduru, 2002; AT&T, 2002). The AT&T Privacy Bird displays a bird

icon in the browser's title bar that changes color, shape, and optionally generates sound to indicate whether or not a website's P3P policy matches a user's privacy preferences. The icon is also used to access the privacy policy information and the conformance information.

Figure 2.10 shows the bird icon indicating a conformance conflict. The user then may access the reasons for the conflict through the bird icon. The steps to access the reasons for the conflicts are shown in Figure 2.11 The conformance information for the entire page is presented in a separate frame, which may occlude the original web form as shown in Figure 2.12.

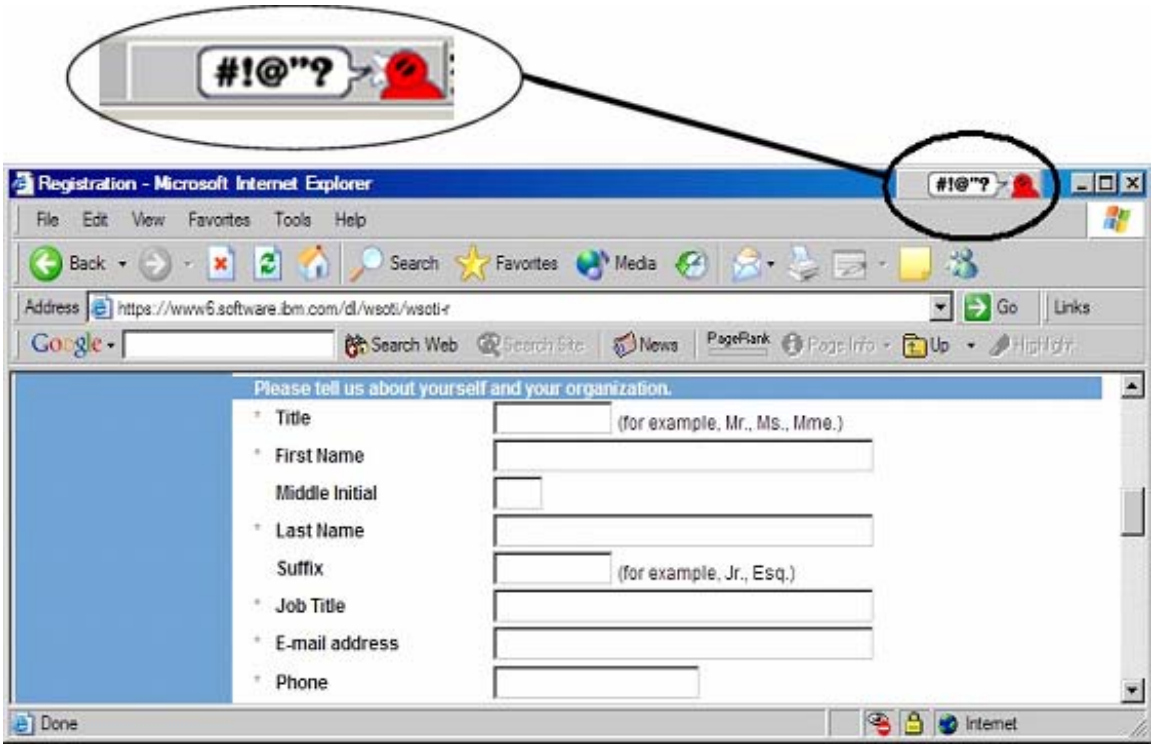


Figure 2.10: AT&T Privacy Bird indication of a conformance conflict

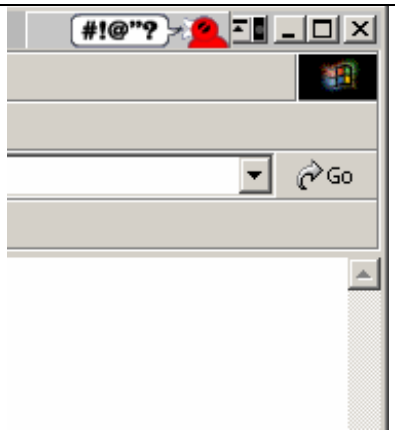
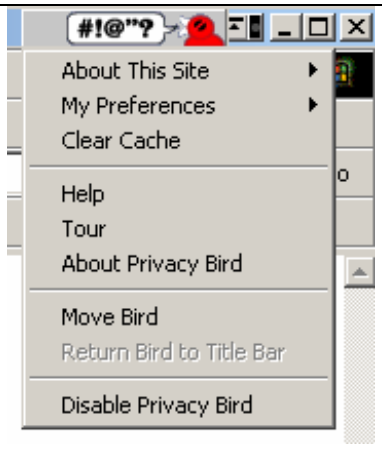
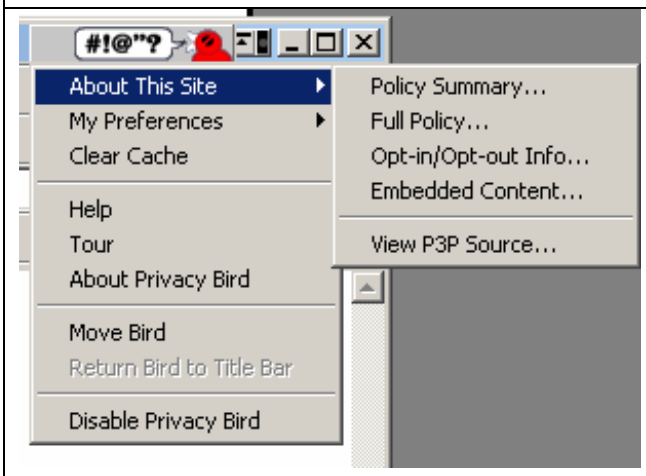
	Step 1. The user must notice that the bird icon has turned red. Step 2, Navigate the mouse pointer from the input field to the bird icon in the frame and click
	Step 3. Slide the mouse pointer over "About this Site"
	Step 4. Slide the mouse pointer over "Policy Summary…" Step 5. Click mouse button 1 Step 6. Examine the Privacy Policy Check as shown in Figure 2.12. Step 7. Dismiss the dialog box to continue with the form.

Figure 2.11: Steps to obtain privacy conformance information with the AT&T Privacy Bird.

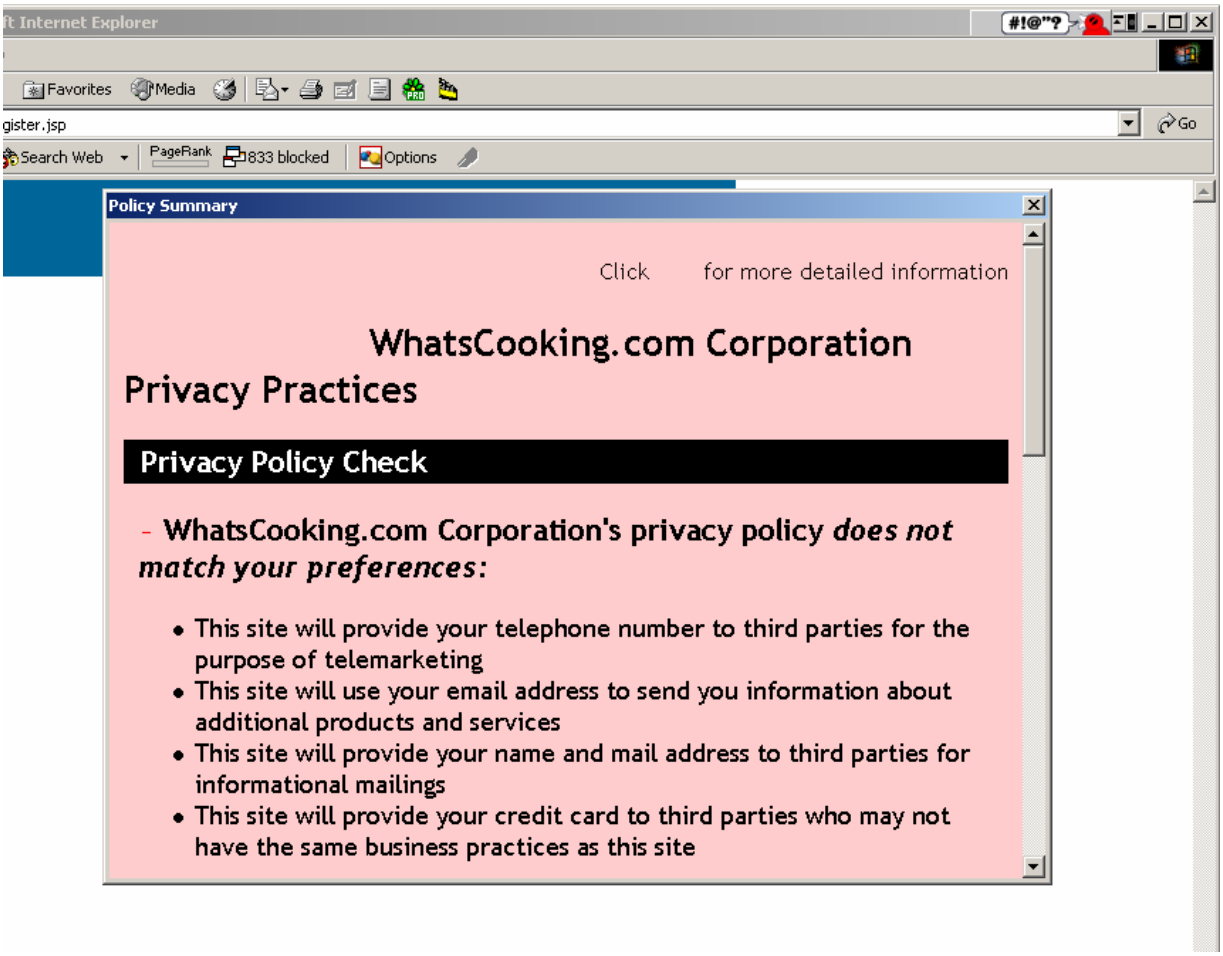


Figure 2.12: Privacy Policy Check display in the AT&T Privacy Bird.

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. In practice only one P3P Privacy policy is created for an entire website. This coarse granularity makes it more difficult for the user to determine the cause and origin of a conformance conflict.

Five user interface problems are present in the AT&T Privacy Bird. First, the conformance indicator in the frame, as shown in Figure 2.10, may not be noticed when filling out the input form. The user is focused on the form-filling task and may not notice the surrounding environment. Second, a multi-step process, outlined in Figure 2.11, is required to

obtain conformance information. This requires additional time and effort that the user may not want to perform.

The last three interface problems are directly related to the lack of context in the P3P implementation. Since the entire privacy policy for the website is compared against every web page, the AT&T privacy bird is unable to indicate if the privacy policy applies to the current web page or any particular field on that web page. Web pages that have no input fields will show the conformance for the site, which may be confusing to the user. If the web page does have an input field, there is no way to know, without reading the complete privacy check information, as shown in Figure 2.12, if the conflict applies to that field or not. Since the privacy agent is unable to narrow down which policy statements apply to the current page, all of the privacy policy conflicts must be presented, which occludes the original input form, as shown in Figure 2.12.

2.6 Conformance Visualization in Privacy Agents

Information visualization attempts to change the problem of understanding an information space from a cognitive problem to a visual perception problem (Card, Mackinlay, and Shneiderman, 1999; Shneiderman, 1997). Although the visualization reduces the level of detail, perhaps to a single indictor, the visualization maintains sufficient information to permit the user to make an informed decision about the data.

Design principles for presenting conformance information that have been identified in previous literature include:

- *Hide Complexity of the underlying data and schema*. The base P3P data schema covers hundreds of data items, each with dimensions of purpose, recipient, and retention. A complete visualization of a matrix style interface for P3P would likely be overwhelming to the user. The user interface must hide the complexity of the

underlying XML specifications (Reagle and Cranor, 1999; Ackerman, Cranor, and Reagle 1999).

- *Be transparent*. Ackerman, Cranor, and Reagle (1999) suggest the user's interaction with an interface for controlling privacy information must be nearly transparent and minimal during the actual social engagement. A transparent interface requires presentation of conformance in a discreet but obvious manner. Conformance conflicts should be readily observable and not easily ignored. A minimal interface requires that the interface artefacts are small and do not clutter the screen.
- *Permit access to conformance information in context of the user's task*. Conformance information detail needs to be readily accessible and accurate (Cranor and Reidenberg, 2002) to enable the user to understand the true nature of the conflict. Providing information detail in context and specific to the PII in question will aid the user in building an accurate model of the conflict for evaluation.

This thesis lays groundwork that enables a privacy agent to present privacy policy and conformance information in a manner consistent with these principles. Establishing the link between the HTML input field and the associated privacy statement minimizes the amount and complexity of information to be displayed, and will permit a user interface that presents privacy conformance in a transparent and readily accessible manner, permitting the user to better understand the website privacy policy.

2.7 Customizable, User Adaptable, and Adaptive Systems

Research into the customization and adaptation of interactive systems is based on the idea that people have different needs, skills, and goals; therefore, applications and documents cannot be designed to adequately support all users in all situations (Morch, 1997; Trigg and Bodker, 1994). In particular, the delivery of customized web page content has become a major

requirement of on-line advertising, direct web marketing, electronic commerce, and teaching applications (Cannataro and Pugliese, 2001). Systems that can adapt to the user, or that can be tailored by the user, have been developed to address this problem. Three levels of adaptability have been discussed in previous literature: customizable systems, user-adaptable systems, and adaptive systems.

2.7.1 Customizable systems

Customization is a process where a user alters a system or web page to suit their needs or tastes. At this level, the system is not involved at all in the decisions about what to change or when to adapt. There are a number of ways to customize a system, each with different effects on the system, and each with different technical requirements (Morch, 1997). *Parameters and options* are settable variables that provide users with an easy way to modify a system's appearance and functionality (e.g. Trigg, Moran, and Halaaz, 1987). *Integration* involves adding or linking together pre-defined components or modules: for example, building a customized portal page in a site such as Yahoo (Manber, Patel and Robinson 2000). *Extension* is needed in situations where components themselves must be changed or where new ones must be created. This type of customization is both the most powerful and the most difficult, as it requires programming ability (Morch, 1997).

Different types of users undertake different types of customization (Mackay, 1991; Mackay, 1990). These types include *workers* – the majority of users, with little technical interest in the system; *tinkerers* – "a worker who enjoys exploring the system, but may not fully understand it" (Page et al., 1996, p. 176); and *programmers* – people who understand the system and have training or experience with computing and coding. The majority of users are

workers without technical knowledge; this implies that the more complex types of customization (integration and extension) will only be carried out by relatively few people.

Two main reasons for customizing are to improve efficiency for a particular task (Mackay, 1991), and to change the appearance of the system (Page et al., 1996) (although there are others, such as to learn about the system (Mackay, 1991)). However, despite the potential benefits of customizing, a trade-off is noted by several researchers between the benefits and costs of creating and using modifications (Mackay, 1991; MacLean et al., 1990; Trigg and Bodker, 1994). For example:

- Not knowing what can be customized, or how changes can be made, is a barrier for workers who are uninterested in system capabilities that are not related to their primary tasks.
- Customization takes time, both to create the actual modification and to deploy the modification in the environment. Amount of effort was found by several researchers to be the largest reason that people do not undertake a customization that could benefit them.
- Customization involves risks – primarily, risk of causing problems to the existing setup. Risks are particular barriers in systems with interdependencies between different settings and components (e.g. login scripts in Unix) (Mackay, 1991).

These problems suggest that customization could in some cases be better handled by the system itself – that is, where a system is able to adapt to the user, based on a model of that user. Although there are applications that adapt interfaces and functionality (e.g., MSWord's frequency-based menus), in the next section we focus primarily on adaptation in hypermedia systems.

2.7.2 Adaptable and Adaptive Hypermedia Systems

The difference between adaptable and adaptive hypermedia systems is the amount of control the user and the agent have over the user model. Teltzrow and Kobsa (2004) define a *user-adaptable* hypermedia system as one where the user is in control of the initiation, proposal, selection, and production of the adaptation. *Adaptive* Hypermedia Systems, in contrast, adapt to users automatically based on their assumptions about them (Fink, Kobsa, Schreck, 1997).

The basic components of Adaptive and Adaptable Hypermedia Systems are the *Application Domain Model*, the *User Model*, and the *techniques to adapt presentations* to the models (Cannataro and Pugliese, 2001).

The Application Domain model

This model describes in an abstract way, the environment in which the content will be used. The environment may encompass a variety of contextual information, such as the user's input-output devices, network bandwidth, and application capabilities. For example, the application domain model for an adaptive advertising system on an e-commerce web site may include information about products that are currently on sale, the graphical capabilities of the web browser the user is accessing, and the user's connection speed.

The User Model

The User Model in the adaptive system can store several different types of information about the user, such as knowledge, preferences, past behaviour, or browsing activity. The User Model in a user-adaptable system often reflects only the user's preferences, although some systems allow the user to control the utilization of activity and other information (Findlater and McGrenere, 2004). In a fully adaptive system, all information in the user model serves as input to decisions about changes to the hypermedia documents.

Adaptation techniques

Techniques to adapt presentations fall into two categories – adaptive presentation and adaptive navigation (Brusilovsky 1994, Cannataro and Pugliese, 2001). Adaptive presentation adapts the content of the web page accessed by a particular user to current knowledge, goals and characteristics of the user. An example is adaptation of online advertisements known as *banner ads*. (Langheinrich et al, 1999). In this system a particular banner ad is selected by considering the keywords presented to a search query service or the URL of the web page that a user has selected. A link to the selected banner advertisement is inserted into the web page by the web server before delivery to the web browser. The web-browser then loads the selected advertisement.

Adaptive navigation supports techniques to help users to find their paths in hyperspace by adapting the presentation of links in order to match user goals, knowledge, and other characteristics. Adaptive navigation can support users in their navigation by limiting the browsing space, providing adaptive comments to visible links or by suggesting the most relevant links to follow (Brusilovsky, 1994). For example, the ADAPTS system (Brusilovsky, 1999) plots a course for a technician's maintenance operation based on what the technician is doing and the skill level and experience of the technician. Inexperienced technicians are provided with additional links to subtasks which are hidden from the experienced technician. Another example is MSOffice 2003 (Microsoft Office 2003, 2003), which offers adaptive menus and toolbars based on frequency of use. When a menu is displayed, only a subset of the menu contents is visible. The contents of the menu subset are based on the frequency of selection of a menu choice. If a menu choice is not visible, the user may make a special selection to display the entire menu. Once a menu item is selected, it will be displayed next time in the menu subset.

The privacy agent implemented in this thesis is an example of a user-adaptable hypermedia system that uses the techniques of adaptive presentation and adaptive navigation to present conformance information. Conformance information is essentially a content modification, but it is not like any of the content modifications that are traditionally done in adaptive systems. Conformance is a different level of information – meta information about the content on the page and how the vendor actions implied by the content correspond to the user's privacy preferences.

In IPV, the Application Domain Model is defined by P3P and the policy link encoded in the HTML web page. The Application Domain Model is static, in that once the privacy policy designer and web page designer have published their work, the Application Domain is not influenced by the network, user hardware, or software. The IPV User Model is defined by APPEL and is user-adaptable: that is, the user explicitly controls the record of their privacy preferences, and there is no adaptation by the agent based on user experience (although future work may consider collaborative user modelling techniques, where peer and expert opinion would allow for automatic inference of a user's privacy preferences). Finally, the adaptive presentation used by IPV is the insertion of a conformance indicator next to the input field. The specific model defined by P3P, APPEL, and the HTML link permit a range of visualizations that could be further adapted by both the user and web page designer. The adaptive navigation involves the linking of the conformance indicator to the specific conformance information created by the P3P and APPEL model.

Chapter 3 Integrated Privacy View

The Integrated Privacy View (IPV) is a scheme for testing and displaying conformance information at the input field level. IPV is made up of three main parts – an extension to P3P that can be inserted into HTML, a mechanism for intercepting web pages that contain the P3P extensions, and an agent to perform the conformance check and insert the indicators. These three parts correspond to three stages in the adaptation process:

- A link must be established between a specific privacy policy statement created by the privacy policy designer and an input field created by the web page designer
- The system must be able to intercept and inspect a web page to find the privacy link before that page is rendered by the browser
- The privacy agent must be able to modify the web page to be rendered by the browser by inserting a visualization of the privacy conformance.

The next three sections provide technological background and implementation details for each of the three parts of IPV.

3.1 Privacy Statement Link

Establishing the link between a specific privacy policy statement and an input field involves extending the XHTML input element to identify a specific privacy statement element in the P3P privacy policy. P3P provides a way to uniquely identify the applicability of a privacy statement as introduced in Section 2.4.1.

3.1.1 Technology foundation: Specifying P3P policies

The link between P3P policy statement and web page input field involves different technologies at each end. This section sets out the basic technologies used by the privacy policy designer and by the web page designer.

At the policy end, the privacy policy designer builds the P3P specification. Currently a P3P privacy policy can apply as broadly as an entire website or as narrowly as an individual web page. The range of web pages is specified by the privacy policy designer at a well-known URL (`http://<host>/w3c/p3p.xml`). The privacy policy designer uses the `<policy-ref/>` element to control the scope through identification of a range of web pages specified in the `<include/>` element. The example shown in Figure 3.1 states that the P3P privacy policy defined in `policy.xml` applies to all pages on the website. All privacy policy statements found in that policy will be evaluated for that web page, regardless of the actual web page content.

```
<POLICY-REFERENCES>
   <EXPIRY max-age="172800"/>
       <POLICY-REF about="/p3p/policy.xml">
          <INCLUDE>/*</INCLUDE>
       </POLICY-REF>
 </POLICY-REFERENCES>
```

Figure 3.1: Policy reference describes which web pages are covered by the specified policy

A P3P privacy policy does permit a designer to state the applicability of individual privacy statements through the `<DATA_GROUP>` element. This element contains DATA Elements, which are able to define a range of PII from a broad category, such as address, down to an individual data element. This data element reference provides a means to identify a specific privacy policy statement (this will be used in the IPV extension). The example shown

in Figure 3.2 states that the given privacy policy statement applies only to the user's email address.

```
<STATEMENT>
   <PURPOSE>
   <RECIPIENT>
   <RETENTION>
   <!--  Base data schema elements.   -->
   <DATA-GROUP>
      <DATA ref="#user.business-info.online.email" />
   </DATA-GROUP>
</STATEMENT>
```

Figure 3.2: P3P privacy statement

At the web page end, the web page designer encodes a web page input form in HTML 4.0 or XHTML, specifying the `<input/>` element with a range of attributes. New attributes are permitted and are ignored in the rendering process by web page browsers. As described below, adding an attribute that specifies the data element reference is what provides the link to the privacy statement and also provides an anchor point in the HTML to do a context sensitive visualization.

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. Introduction of new input fields not covered by the current privacy policy does require the privacy policy designer to update the privacy policy. Automated verification of input field data elements with the privacy policy would ensure completeness and accuracy.

A privacy agent needs to be able to find the privacy statement link in the target web page (see Section 3.1.2). Web pages coded in XHTML permit fast parsing into a tree structured Document Object Model (DOM) (W3C, 1998). A privacy agent may then utilize XPATH (XML Path Language, 1999) 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. The evaluation logic in the privacy agent is simplified in that only one specific privacy statement needs to be compared with the user's privacy preferences. After evaluation, the agent has the result of the conformance evaluation, the text to explain the conformance result, and the location of the input field within the HTML document. The actual implemented link is described below.

3.1.2 Extension of P3P: P3Pdataelement tag

IPV takes advantage of the fact that P3P defines a base data schema that sets out by name the data elements that a vendor might collect. For example, a user's business email address would be specified as `#user.business-info.contact.online.email`. Both the vendor privacy policy and user privacy preference use the P3P base data schema and 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 privacy policy. Use of the data element then permits a direct association between a privacy policy statement and a user privacy preference.

IPV defines the *p3pdataelement* attribute that the web page designer adds to the input field of the web form. An example is shown in Figure 3.3. This attribute specifies a data element that then indicates a specific policy statement to associate with the input field. This permits the conformance evaluator to limit the scope of the evaluation to the specific privacy statement associated with the input field.

The value of the p3pdataelement attribute is taken either from the P3P data element base schema, or may be a value provided by an extended schema from the website developer.

In either case, the value uniquely identifies the privacy policy statement applicable to the input field.

levysn@sasktel.net

```
<input name="r_reg_email"
       size="40"
       value="levysn@sasktel.net"
       p3pdataelement="#user.business-info.online.email" />
```

Figure 3.3: The input field element extended with the p3pdataelement.

The HTML to display the input field in Figure 3.3 has been extended to add the p3pdata element. HTML allows the extension of attributes not defined by the HTML schema. During the page rendering the browser simply ignores attributes that are not understood by the HTML schema.

The p3pdataelement is also used to locate the HTML input element in the HTML DOM. This location is then used as an anchor point to insert the conformance result. The conformance result is expressed in HTML using the span element as the 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. An example of the HTML span warning of non-conformance is shown in Figure 3.4.

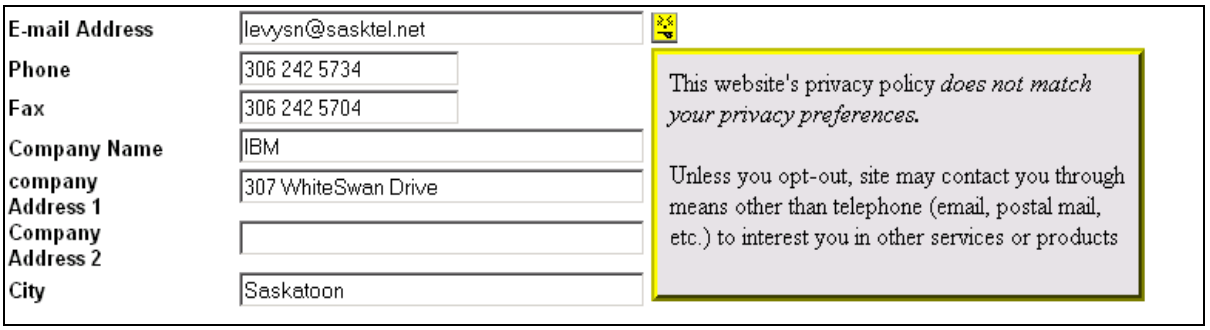


Figure 3.4: IPV conformance display with additional information display.

3.2 Diverting web pages to the IPV agent

The IPV privacy agent could have been implemented as an HTTP proxy or a browser helper object (also called a plug-in) in order to intercept, examine, and potentially modify the target web page. Each technology has certain advantages and disadvantages.

The browser plug-in has the advantages of speed and not utilizing the browser's single proxy port. The disadvantages are that a plug-in is browser specific and must be written in a programming language specific to the browser. The HTTP proxy has the advantage of being able to be written in Java and may be utilized with any web browser. The disadvantages are that the web page must be parsed twice, making the process slower to the user, and utilizes the single proxy port (i.e., there can only be one proxy at a time).

This thesis focuses on exploring the utility of contextual conformance information and not on a specific agent implementation. Therefore, the HTTP Proxy was chosen for its flexibility and use of Java. This also permitted standard toolkits to be used for network communication and XML manipulation.

3.2.1 The IPV HTTP Proxy

Communication between a web browser and a web server takes place over HTTP. The browser issues an HTTP request and the server responds synchronously with an HTTP response. The HTTP response has a header called a MIME type, which indicates whether the response is HTML, an image, or some other data type.

An HTTP proxy intercepts all HTTP requests and responses that flow between a client web browser and a website. The proxy may be a process located on the client machine or at another port on the network. The proxy listens at a specified port for a request from the client browser and forwards that request to the host specified in the HTTP request. The proxy

maintains a session with the client so that the response from the host will be returned to the right client and port. HTTP Proxies are implemented to support web page caching strategies, firewalls, or for any purpose requiring a "filtering" of the HTTP data flow.

The purpose of the IPV proxy is to look for HTTP responses from the server that may contain the p3pdataelement attribute. Candidate web pages are then forwarded to the IPV agent for processing.

3.2.2 The Interception Process

The communication flow for the IPV proxy is shown in Figure 3.5. The process begins when the web browser makes a request for a web page. The proxy notes the client port that the request came from, determines the host port to send the request to, and forwards the request to the server. The server processes the request and sends a response to the proxy.

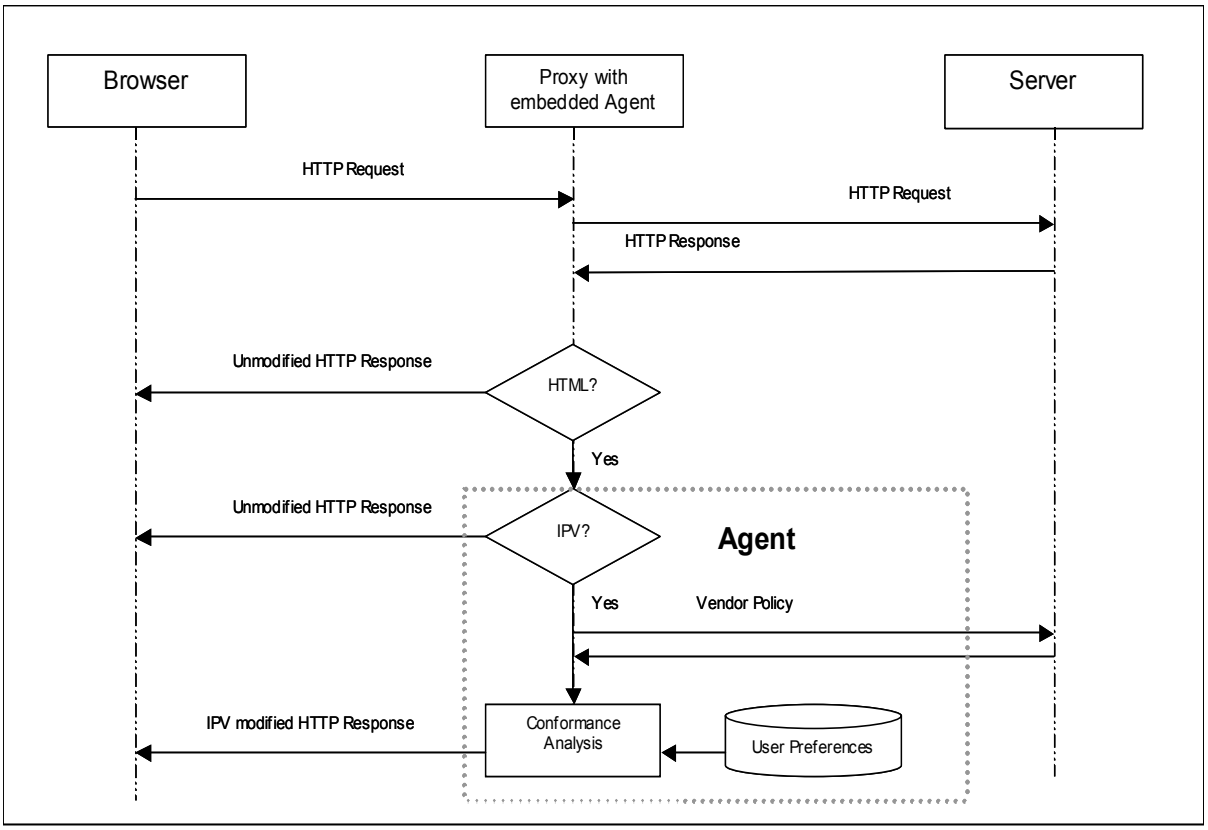


Figure 3.5: Communication flow for IPV proxy

The proxy then looks for a response with a MIME type of HTML. This indicates that the response is a web page and therefore may contain input fields with the p3pdataelement. All other responses are immediately returned to the client browser unmodified. A complete response from the host may consist of one or more segments, so the proxy continues communication with the host until the complete web page is received.

The IPV Agent is then invoked with the retrieved web page. 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 done (described below). The result of the evaluation is inserted as new HTML in the web page and the page is returned to the client browser. The browser can then render the web page. An example input field with conformance information display is shown in Figure 3.6.

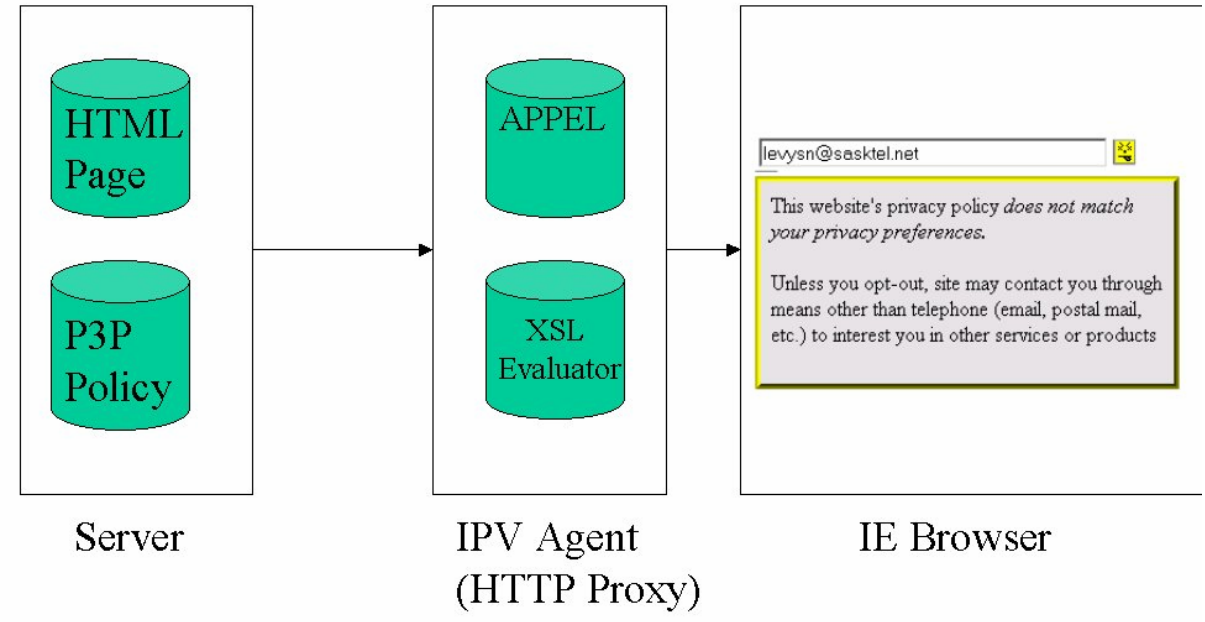


Figure 3.6: Data flow for IPV agent.

3.3 The IPV User Agent

IPV implements a user agent that takes advantage of the p3pdataelement. The user agent is responsible for evaluating the user privacy preferences and the vendor privacy policy to produce a conformance display for the user. The IPV agent utilizes the p3pdataelement attribute to produce conformance information specific to an input field and displays that information in the context of the input field. There are two parts to the agent: a part that does the conformance checking, and a part that adapts the web page by inserting the conformance indicators and additional conformance information.

3.3.1 IPV Conformance checking

The IPV agent is embedded in the proxy and is given all the pages that are IPV candidates. For this project, the agent is implemented in Java, Xerces (Xerces Java Parser, 2002) and Xalan (Xalan-Java, 2002). The IPV agent takes as input a web page, a user privacy preference specification, and a vendor policy, and produces a web page with embedded conformance information.

The IPV user agent performs the following seven steps:

1. *Read in the source web page*. The proxy provides the target web page to the agent. The web page may be written in HTML or XHTML but must be valid XML. The agent performs a parse of the XML text into an XML DOM. This input DOM will later be updated with the conformance result in step 6 and written out as the HTTP response to the client browser.
2. *Create node list of input elements*. Utilizing XPATH, a node list of input elements with the p3pdataelement attribute is created. If this list is empty, then there is no data

collection performed by this page and the agent writes the unmodified web page back to the client.

3. *Read in the vendor policy and user privacy preferences.* The privacy policy specified in P3P is then retrieved from the web server. The communication path is shown in Figure 3.5. The vendor privacy policy is published at a well-known public location determined by the web page URL. The user privacy preferences, expressed in APPEL, are stored locally with the agent. For this project, only one set of user privacy preferences are stored to be used with all web sites. These XML documents are read in and parsed into two XML DOMs (Cranor et al., 2002b). The two XML DOMs will then be used as input to the evaluation process described in step 5. Steps 4 through 6 are repeated for each node found in step 2.
4. *Identify statement in privacy policy associated with p3pdataelement.* The value of p3pdataelement is used to select the privacy statement in the privacy policy that specifies that data element. Again XPATH is used to select the correct statement for the evaluator. For example, for the input field shown in Figure 3.3, the statement shown in Figure 3.7 is selected based upon the DATA element with the ref attribute equal to `#user.business-info.online.email`.

```
<STATEMENT>
   <PURPOSE>
   <RECIPIENT>
   <RETENTION>
   <!--  Base data schema elements.   -->
   <DATA-GROUP>
      <DATA ref="#user.home-info.postal.country" />
      <DATA ref="#user.business-info.online.email" />
   </DATA-GROUP>
</STATEMENT>
```

Figure 3.7: P3P privacy statement selected by the DATA element

5. *Evaluate.* The evaluation of the selected privacy statement against the user privacy preferences is a four-step process all performed by XSLT style sheets (XSL Transformation Version 1.0 Specification, 1999). This process of evaluation is described in the APPEL 1.0 working draft. Each step defines an XML document as its output. The result of each step becomes the input into the next step.

5.1. *Normalize privacy preferences.* Evaluation of the privacy policy statement with the privacy preferences is done by comparing facts and applying logical operators. Normalization expands default or implicit facts and logical operations into an explicit form. For example, the data element in the data group may specify either data element names or data categories. To simplify comparison, the IPV agent converts all data elements to their categories. For example, the statement shown in step 4 is shown normalized in Figure 3.8. In this case, `#user.business-info.online.email` is associated with the category 'online'.

Document normalization is achieved by expanding the P3P data schema and collecting the data categories for each data element. The comments in Figure 3.8 reflect the process of following the data schema path until a terminal node is found. The 'categories' element is then created from the found terminal node.

```
<DATA-GROUP connective="or">
<DATA ref="#user.business-info.online.email">
   <!-- Search path: user.business-info   -->
   <!-- Search path: contact.online   -->
   <!-- Search path: online.email   -->
   <CATEGORIES connective="or">
     <online />
   </CATEGORIES>
 </DATA>
</DATA-GROUP>
```

Figure 3.8. Normalized APPEL document.

5.2. *Normalize privacy statement.* The same process explained in step 5.1 is applied. The result is that both policies may now be matched comparing for the existence of one fact in the other.

5.3. *Perform match.* An XSL Style sheet then compares the two normalized policies, generating a true or false as a result of the comparison of each fact. A match attribute is placed on each element to indicate whether the facts are equal or not. Notice in the rule in Figure 3.9 that all elements of the privacy preference statement have been found in the vendor privacy policy. This implies that this rule will be selected by the evaluator for inclusion in the conformance result in the next step.

```
<RULE behavior="limited"
description="Unless you opt-out, site may contact you
through means other than telephone (email, postal mail, etc.)
to interest you in other services or products">
  <STATEMENT connective="or" match="true">
    <PURPOSE connective="or" match="true">
      <contact connective="or"
         match="true" required="opt-out" />
    </PURPOSE>
  </STATEMENT>
</RULE>
```

Figure 3.9: RULE element after the match.

5.4. *Evaluate match.* The logical operations specified by the connective attributes are performed on the match output from the previous step. Any statement that evaluates to true is expressed in a result tree. Notice in the rule in Figure 3.10 that the statement has evaluated to true. This implies that there is a conflict in the policy that needs to be reflected to the user. The description will be added to the additional information for this conformance result. One or more RULE statements may evaluate to true. In the multiple case, the additional description for each conformance result will be concatenated into a paragraph for display to the user.

```
<RULE behavior="limited"

description="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">

  <STATEMENT result="true" />

</RULE>
```

Figure 3.10: RULE statement after evaluation indicating a positive match.

3.3.2 The conformance visualization

P3P does not mandate how the conformance indication should be displayed. With the additional information provided by the link described in section 3.1 there are several possibilities. What needs to be done is to provide an indication of conformance associated with an input field and some method of obtaining additional conformance information. The indicator should not interfere with the web page design or with the functioning of the input field.

Implementation of the conformance indicator can be done through a combination of HTML, CSS, and Javascript. The agent is able to manipulate the HTML in the web page to add additional elements, enabling it to display icons, borders, and text. It knows where to place the new elements based on the location of the data element attribute. CSS is used to control the style, location and visibility of HTML elements. Javascript permits the capture of user events, such as mouse movements, which allow procedural code to change the visibility of HTML elements through CSS (e.g. a popup window from a mouse rollover event).

One simple design would highlight the input field background or border using color. For example, a red background could indicate a conformance conflict and a green background could indicate agreement. Additional conformance indication would 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 and that mousing-over the field is sometimes reserved for context-sensitive help. This suggests that a new object needs to be provided for the user to target with the mouse to obtain additional information.

This approach places an indicator in the page, located either before or after the input field; the indicator may use both color and pattern to indicate conformance, and provides a target for a mouse-over to obtain additional information. The option chosen in this thesis was to use an inserted icon that is approximately the same height as the input field and approximately two characters wide. 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 3.11). The specifics of how IPV carries out the insertion are given below.

green indicates conformance

red indicates conflict

Figure 3.11. Conformance icons (3X normal size)

6. *Place conformance result in page.* Each statement in the evaluation result is checked to see if it is true. Text describing the conformance is gathered from each true statement to place in the additional information display for the conformance result. The original web page, still represented as an XML DOM, is updated with an XML fragment to represent the display of the conformance result. The HTML header information is also updated with additional CSS and javascript files to support the IPV display style and interaction. An example of the HTML fragment for Figure 3.4 is shown in Figure 3.12. Note that the HTML fragment is completely contained in the input field as specified in Figure

3.3. This ensures that the icon representing the conformance result (in this case a conflict warning) will appear adjacent to the input field. Each fragment is customized based on the conformance result. Using DOM manipulation, the attributes and span contents are updated to reflect the desired conformance result. Note that the style for the *ipvinfored* class is initially hidden. This span is made visible to the user by a mouse-over. The result as rendered by the browser appears in Figure 3.4.

7. *Write out web page.* Finally, the web page, with embedded conformance information, is written to the client to be rendered by the client browser. An example is shown in Figure 3.4; a full page is shown in Figure 3.13.

```
<input name="r_reg_email"
   size="40"
   value="levysn@sasktel.net"
   p3pdataelement="#user.business-info.online.email"
<span class="ipvframe">
    <span class="ipvsadred"
               ipvinfo="document.all.ipvinfo1001"
               onclick="ipvMouseClick(this)"
               onmouseout="ipvMouseOut(this)"
               onmouseover="ipvMouseOver(this)" />
    <span class="ipvinfored"
      id="ipvinfo1001"
      onclick="ipvMouseClick(this)"
      onmouseout="ipvMouseOut(this)"
      onmouseover="window.status='info'">
      <span>
            <span>This website's privacy policy</span>
            <i>does not match your privacy preferences.</i>
            <p/>
      <span>Unless you opt-out, site may contact you    Through  means
      other than telephone (email, postal mail, etc.) to interest you
      in other  services or products</span>
            <p/>
      </span>
      </span>
</span>
</input>
```

Figure 3.12: XML fragment for displaying conformance result

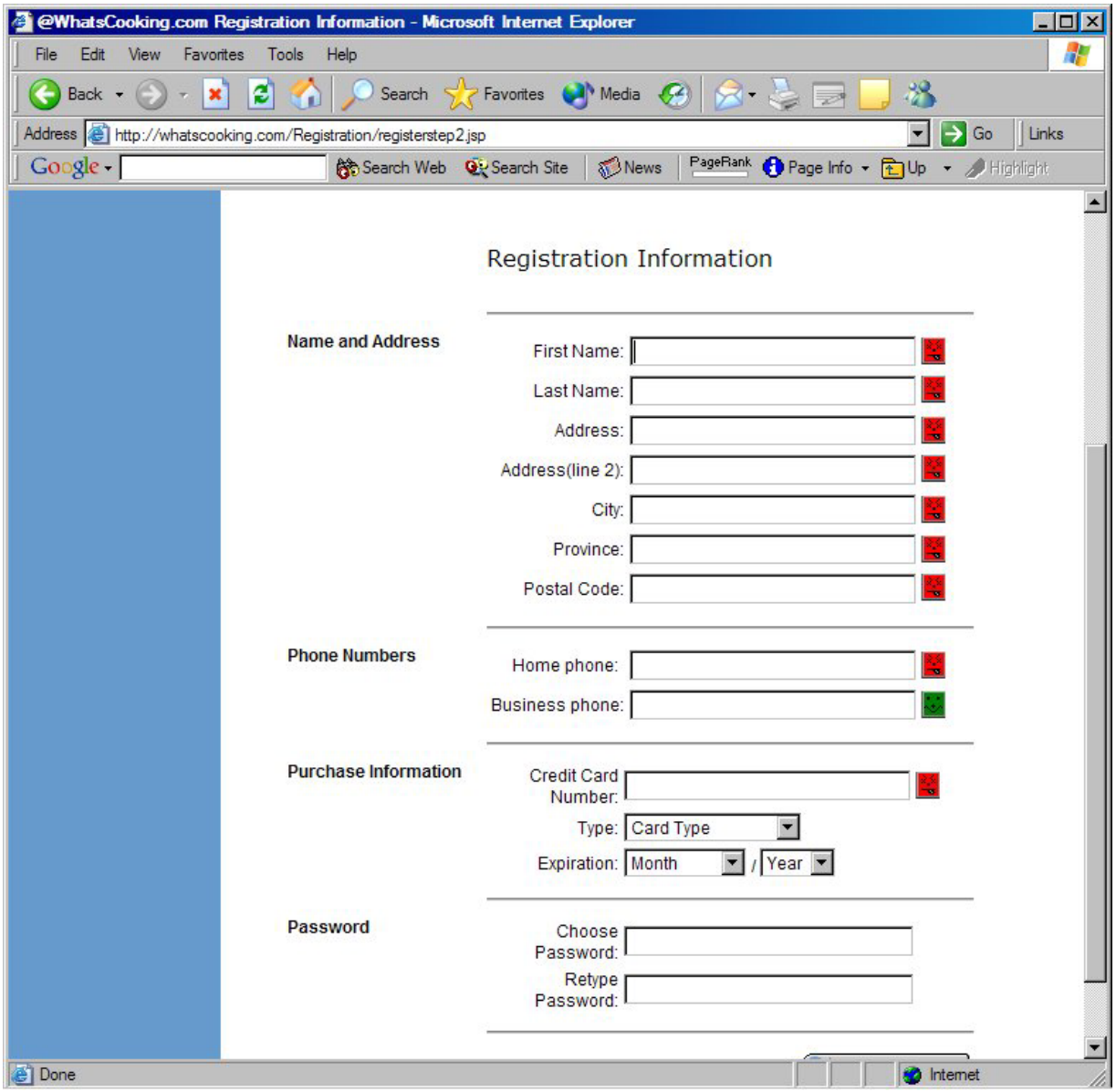


Figure 3.13. Example web page after processing with IPV.

Chapter 4 Evaluation of IPV

The research hypothesis investigated in this study is that visualizing privacy conformance at the input field level will improve the understanding of the source and origin of website privacy conflicts. The evaluation consisted of a usability study to determine if the Integrated Privacy View (IPV) does improve privacy policy understanding when compared to the current state of the art in machine-readable privacy policy agents, the AT&T Privacy Bird. Two typical commercial websites were built, both with pages that gather privacy information of the type that would be required if an individual was going to make a purchase over the Internet. One website used the Privacy Bird user agent, and the other website used the proposed improved product, IPV.

4.1 Purpose of the Evaluation

This evaluation had two main purposes:

1. *To determine if IPV provides an easier way for the user to understand the details of privacy conflicts*. This part of the evaluation considers the fine-grained association of privacy statements to web page input fields. Because the fine-grained approach provides mechanisms for finding information that do not exist in other systems, it is expected that IPV will show clear benefits. The general concept of improving user understanding through contextual display of information has been shown in other situations: for example, advanced help interfaces provide specific context sensitive information based on the user's task and location within the user interface. The first part of the IPV evaluation applies this concept to the presentation of privacy information; therefore, the hypothesis is not tested with a large number of participants (Rubin, 1994).

2. *To explore some of the design issues that must be addressed for this approach to be successful.* This part of the study looks at the way the conformance information is presented, and at how additional contextual information can be given without distracting the user from their main task. The IPV embodiment demonstrates only one method of visualization for the user evaluation. In this part of the study, additional hardcopy displays were shown to the participant to demonstrate different methods of grouping, highlighting, and icon placement. Participants were asked to evaluate these alternate displays and comment on issues of distraction and visibility.

4.2 Methodology

The usability study was based on the work of Rubin (1994). One aspect of Rubin's work justifies the use of a limited number of subjects (3-5) to validate a user interface and discover user-interface difficulties. Rubin also defines criteria for effectively validating a solution.

1. A problem statement or statement of test objectives must be developed
2. The testing phase must contain a representative sample of end users which may or may not be randomly chosen
3. The test must represent the actual work environment
4. During testing, observations of the end users who either use or review a representation of the product must be recorded and captured. This may include a controlled and sometimes extensive interrogation and probing of the participants by the test monitor.
5. Data to be collected include quantitative and qualitative performance and preference measures.
6. Recommendations of improvements to be made to the design of the product are to be included.

For the IPV evaluation, a plan was developed that meets the six criteria. The problem and objectives have been stated above; the remaining five criteria will be described in the following sections.

4.2.1 Participants

The participants in this study included 6 individuals, 3 female and 3 male, who are familiar with using the Internet. Participants were recruited from a local company and were paid for their involvement. All individuals had actually made purchases on the Internet. Participants were aware of PII privacy issues, and had taken varying steps to protect their privacy during web-based transactions; however, none had ever used a privacy agent before. The participant demographics are summarized in Table 4.1.

Table 4.1: Demographic profile of participants

Participant	Sex	Age	Internet Experience	Internet Purchase Frequency
1	F	27-35	Sophisticated	3 to 4 times a month
2	F	27-35	Average	Occasionally
3	M	27-35	Sophisticated	Occasionally
4	M	43-50	Average	Occasionally
5	M	27-35	Average	1 or 2 times a month
6	F	36-42	Average	Occasionally

	For web-based transactions, do you:				
Participant	Inspect privacy policies?	Check that your browser is in 'secure mode'?	Take steps to protect your email address?	Take steps to protect credit card numbers?	Use a privacy agent?
1	Yes	Yes	No	No	No
2	No	No	No	No	No
3	Sometimes	Sometimes	Yes	No	No
4	Sometimes	No	Yes	Yes	No
5	Sometimes	Yes	Yes	No	No
6	No	No	Yes	Yes	No

4.2.2 Experimental Conditions

The experimental set-up for this evaluation recreates a typical website experience for requesting a product or service through a web transaction. Two websites were created, WhatsCooking and AllThatJazz, following the best practices guidelines for creating commercial web sites (Constantine, 2002). Each website consisted of four web pages: an introductory page, a personal information page, a thank you page, and a website privacy policy page (see Figures 4.1 – 4.8). The privacy policy page was available from a link on each

page of the website. The personal information requested by each website was the same, although not in the same order. Figures 4.1, 4.3 and 4.5 show the AllThatJazz website utilizing the AT&T Privacy Bird to display privacy conformance. Figures 4.2, 4.4, and 4.6 show the WhatsCooking website utilizing IPV to display privacy performance.

The participants were provided with fictional personal information to use for the data entry, to prevent any resistance to providing name, address, phone numbers, and credit card information and to protect the participant's privacy. Each participant saw both websites and both user agents; in addition, the agent used with each website was alternated to prevent the visual design of the website from influencing the results. Participants 1, 3, and 5 started with the WhatsCooking website utilizing the AT&T Privacy Bird and then proceeded to the AllThatJazz website utilizing IPV. Participants 2, 4, and 6 started with the AllThatJazz website utilizing the AT&T Privacy Bird and then proceeded to the WhatsCooking website utilizing IPV.

The introductory page for each website, as shown in Figures 4.1 and 4.2, was designed to reinforce the participant training (see Section 4.2.3) and provide a simple task with a high likelihood of success regardless of which privacy agent was used. The user task for each web page (described in detail in Section 4.2.3) was to identify whether a privacy conflict exists, what input field was in conflict and what privacy policy statement was in conflict with the participant's privacy preferences. The introductory page was kept visually simple to permit clear indication of the privacy agent state, and reduce the number of possible conflicting input fields to choose from. Finally, the privacy policy link was kept in plain view in case the participant wanted to user it to determine the cause of the conflict in addition to utilizing the privacy agent.

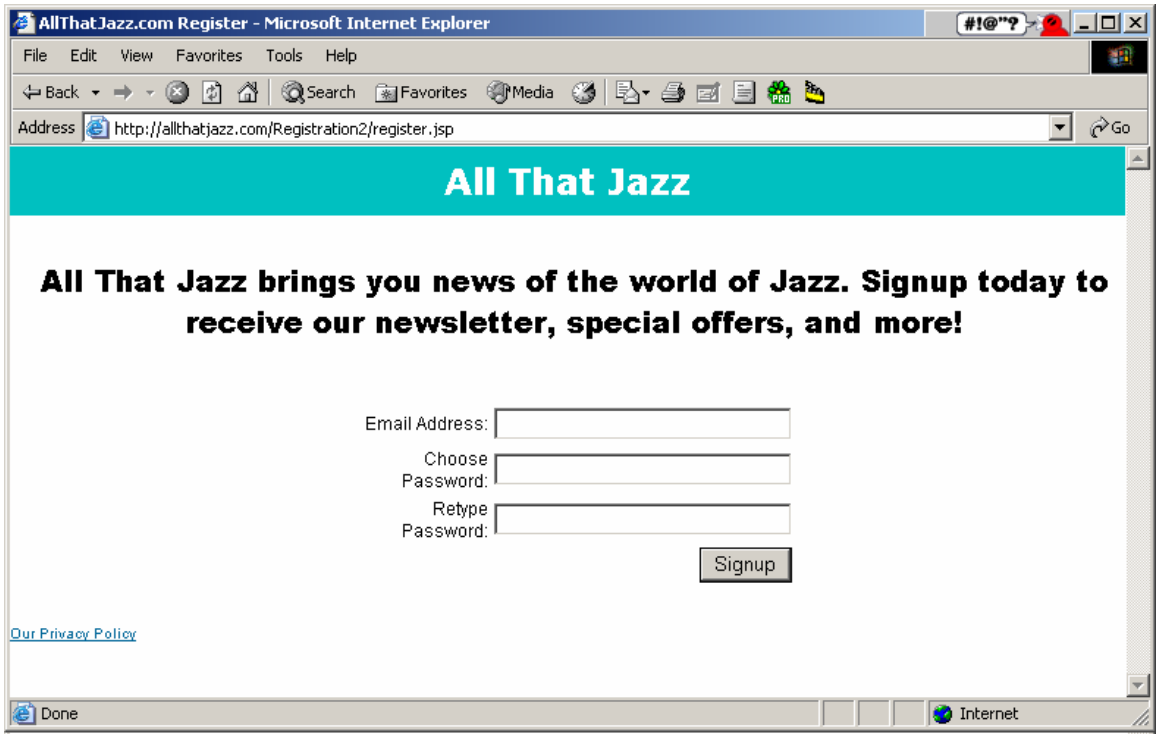


Figure 4.1: Page 1 of the All That Jazz Website utilizing the AT&T Privacy Bird

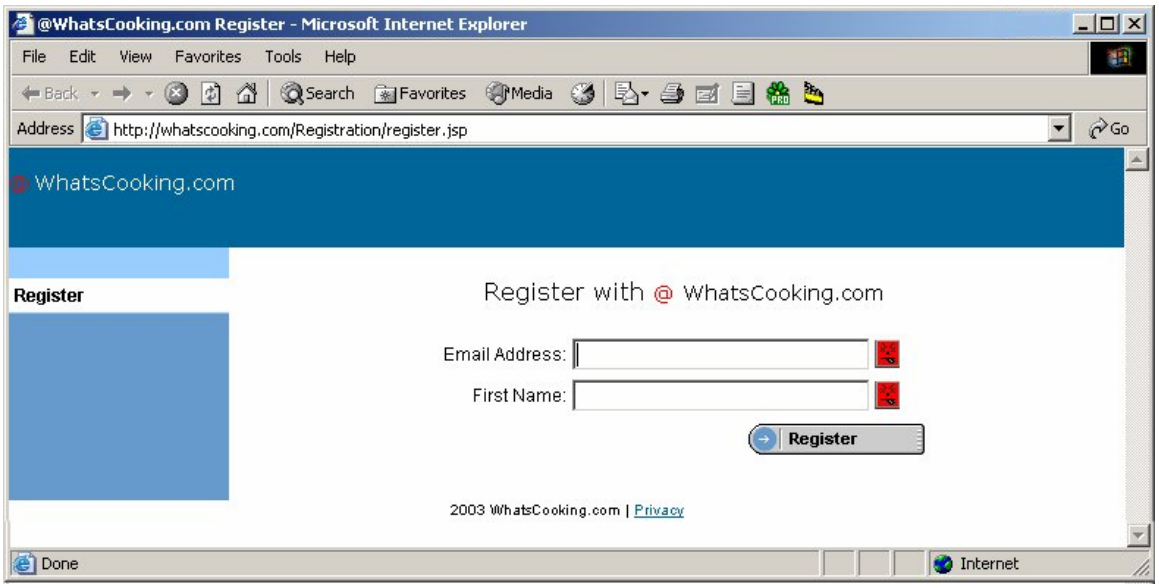


Figure 4.2: Page 1 of the WhatsCooking Website utilizing the Integrated Privacy View

The second page for each website, as shown in Figures 4.3 and 4.4, was designed to gather the complete personal profile of the participant and permit the participant to fully exercise each privacy agent. This page showed data entry fields that both conflicted with the privacy policy, and one specific field, the business phone number, which did not. This page was also designed to be longer than the typical monitor could display without a scroll bar. This would force some of the data inputs fields to be available only by scrolling, and in the case of IPV, having the privacy indicator not immediately visible.

The third page of each website, as shown in Figures 4.5 and 4.6, was designed to understand what the meaning of the privacy agent indicator was in the case where no input fields are required. In particular, the AT&T Privacy Bird typically indicates the privacy policy for an entire website, and doesn't distinguish between pages requiring input and pages which are purely informational. In the case of Integrated Privacy View, no input fields, or no input fields with a specified privacy policy provides no indicator at all. This page explored what the presence or lack of indicators meant to the participant.

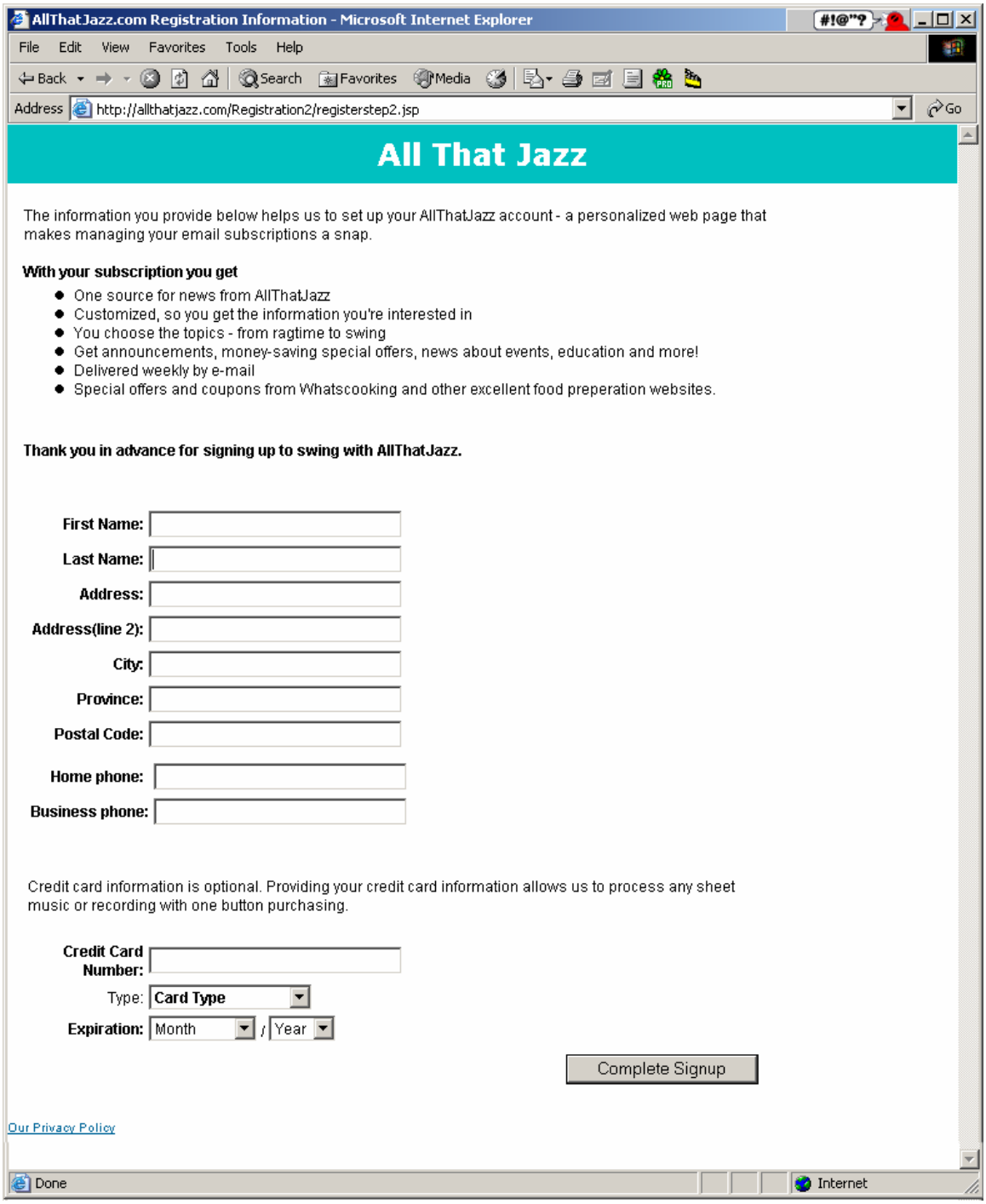


Figure 4.3: Page 2 of the All That Jazz Website utilizing the AT&T Privacy Bird

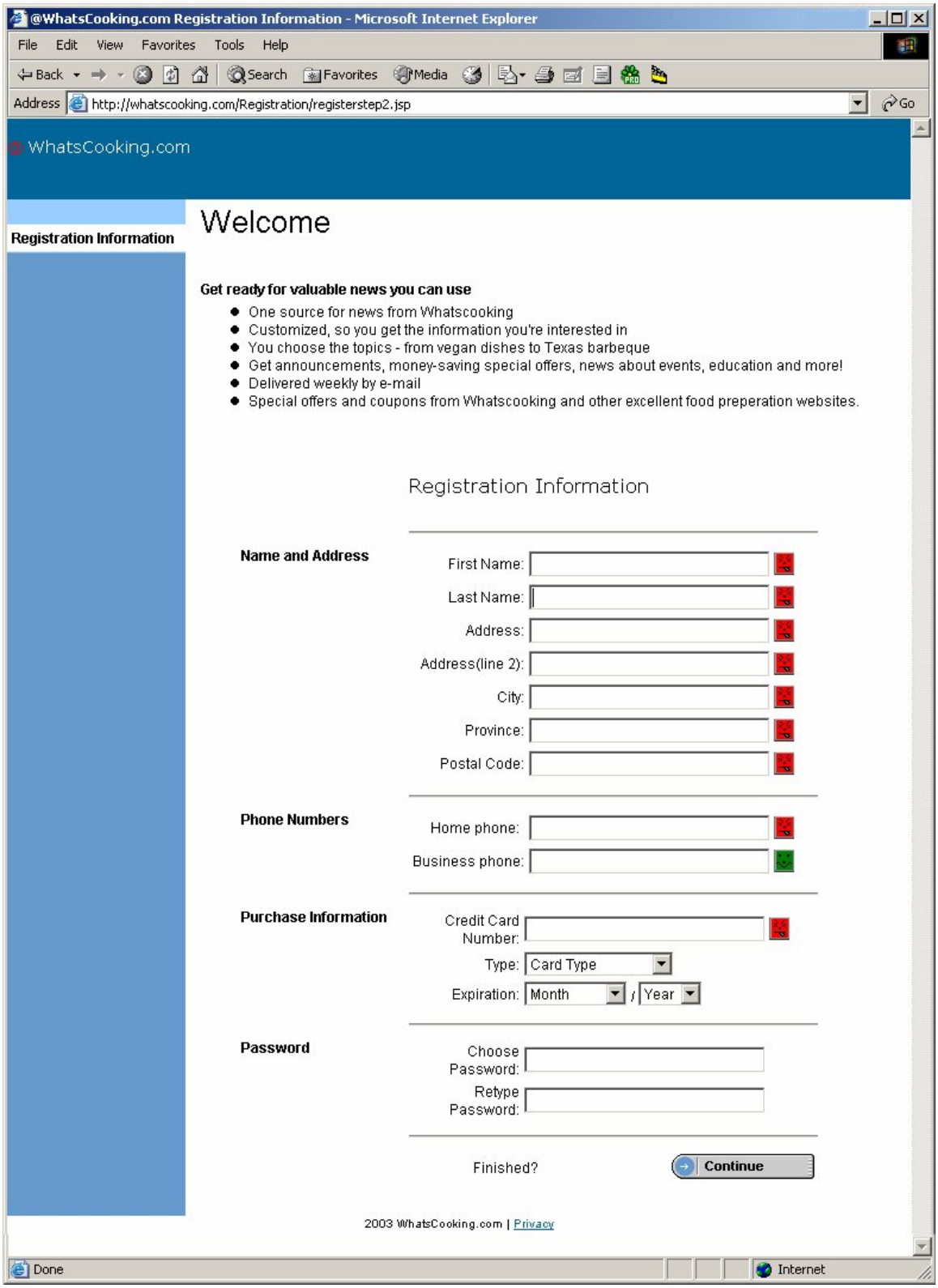


Figure 4.4: Page 2 of the WhatsCooking Website utilizing the Integrated Privacy View

Figure 4.5: Page 3 of the All That Jazz Website utilizing the AT&T Privacy Bird

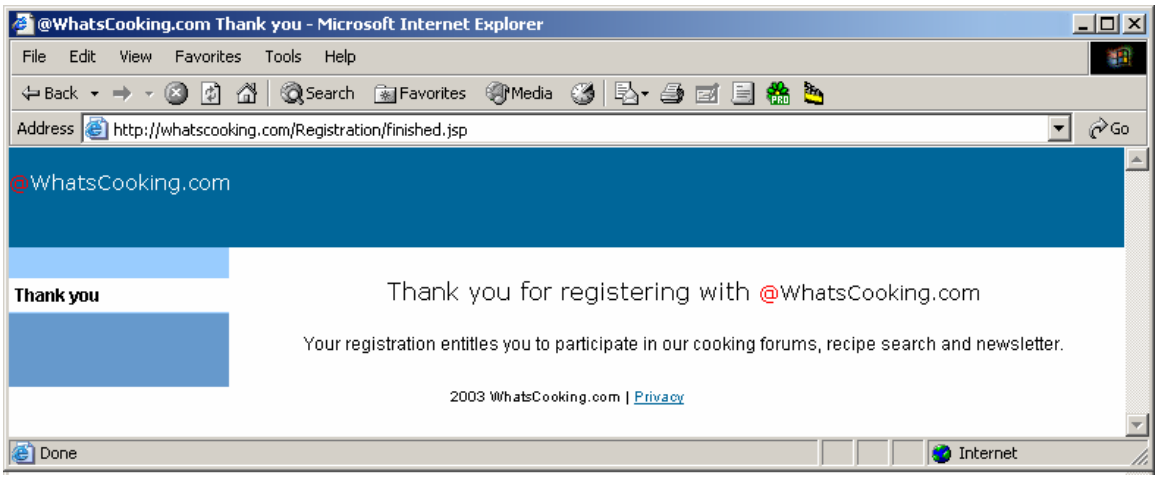


Figure 4.6: Page 3 of the WhatsCooking Website utilizing the Integrated Privacy View

The privacy policies for each website, as shown in Figures 4.7 and 4.8, attempt to describe the privacy policy for a website in simple English. Well-designed websites make the privacy policy readily available on all pages. This is typically done by providing a link to the privacy policy. The participant would be able to completely accomplish the usability task for each web page by utilizing the privacy policy alone if that were their choice.

Each website has the complete privacy policy for the web site available to both the participant and the privacy agents. The privacy policy web page matches the rules encoded in

P3P for use by each privacy agent. This enables the participant to be able to accomplish the usability task for each web page by utilizing the privacy agent alone if that were their choice.

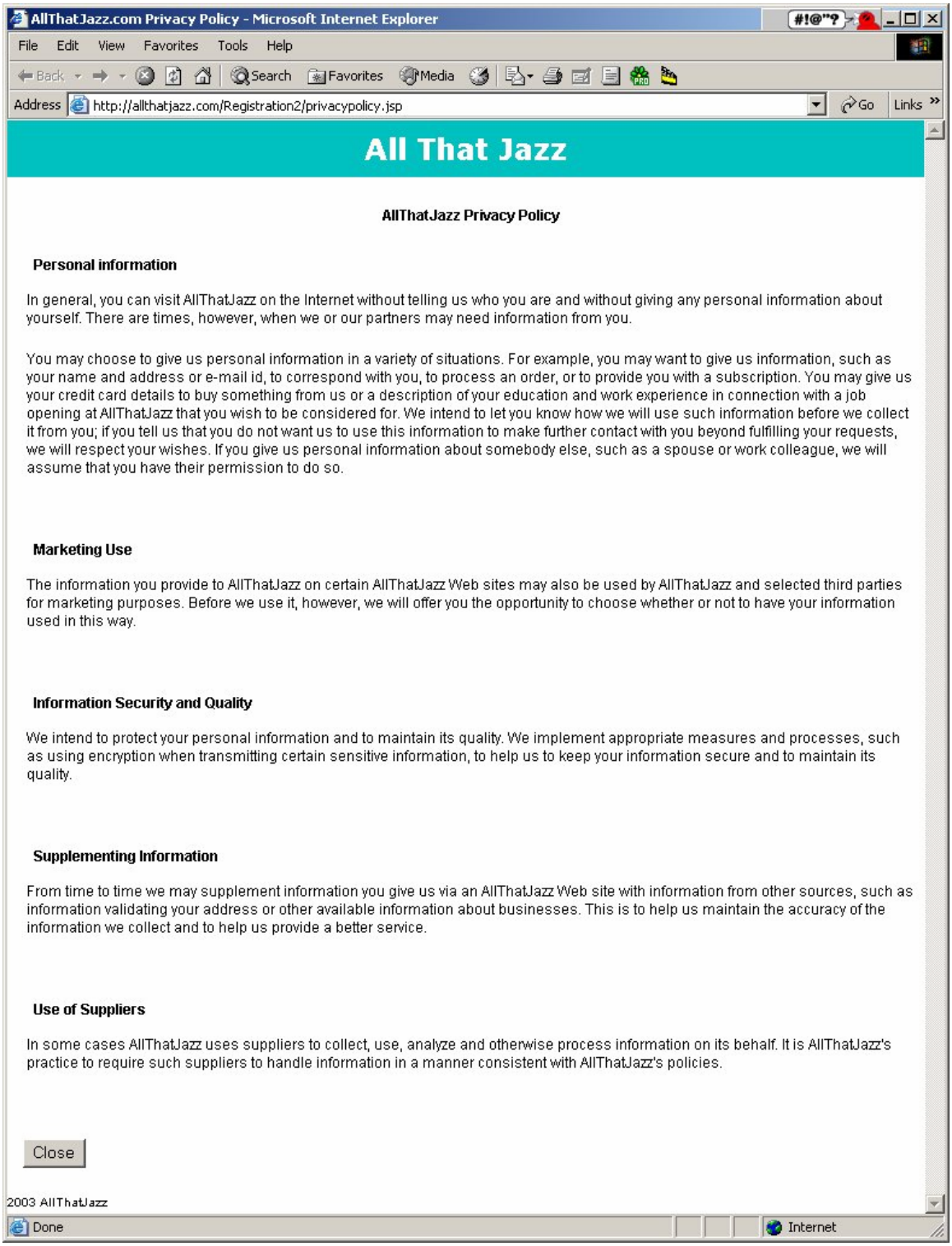


Figure 4.7: Privacy policy for All That Jazz Website utilizing the AT&T Privacy Bird

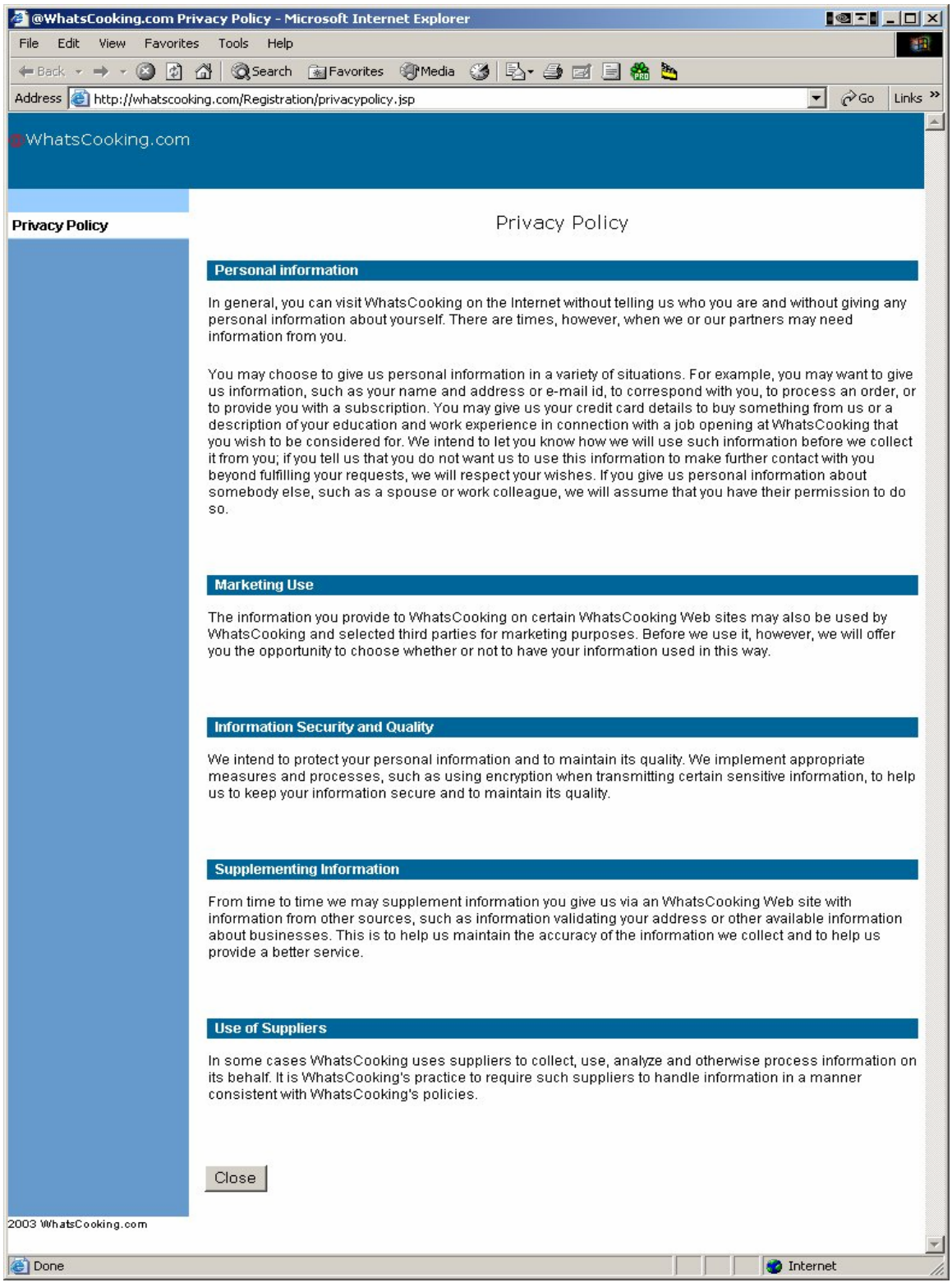


Figure 4.8: Privacy policy for WhatsCooking Website utilizing the Integrated Privacy View

4.2.3 Tasks

The usability test took place in three phases for each agent: training, demonstration of ability, and utilization of the agent to determine privacy conflicts.

Training Tasks

Training was required to illustrate the tasks used in the study. In the training phase, participants were given fictional personal information (name, address, phone numbers, and credit card information) and were asked to perform the following tasks in a training web site (Figures 4.9 and 4.10):

1. Determine the presence of a privacy conflict
2. Determine what field is causing the conflict
3. Determine the cause of the conflict

Training ensures that the participant understands how to accomplish these tasks by utilizing the privacy policy page for a website and the selected privacy agent for that web site. The participant could choose either to read the privacy policy, utilize the privacy agent, or both.

The privacy policy was available on all three pages of the training website through a clearly marked hypertext link located on the bottom of each page. Selection of the link brings up the websites privacy policy in its entirety, which obscures the original web page. To return to the original web page requires scrolling to the bottom of the privacy policy page and utilizing a close button located as the bottom of the page.

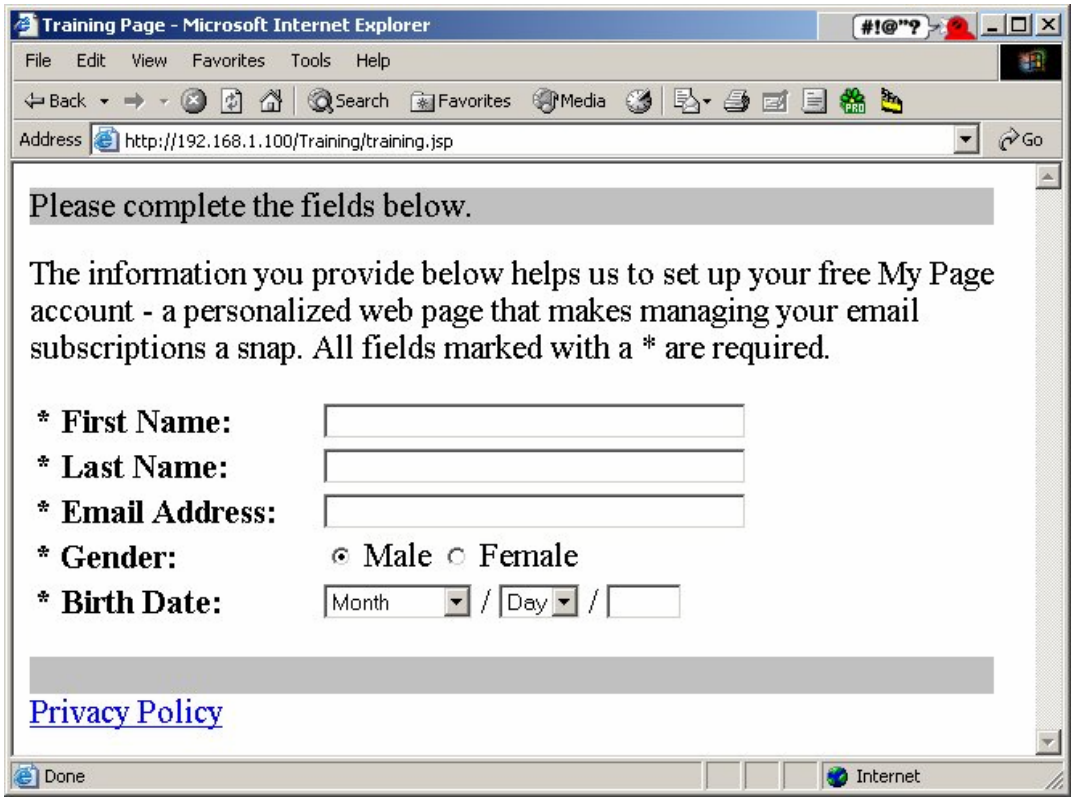


Figure 4.9: Training page utilizing the AT&T Privacy Bird

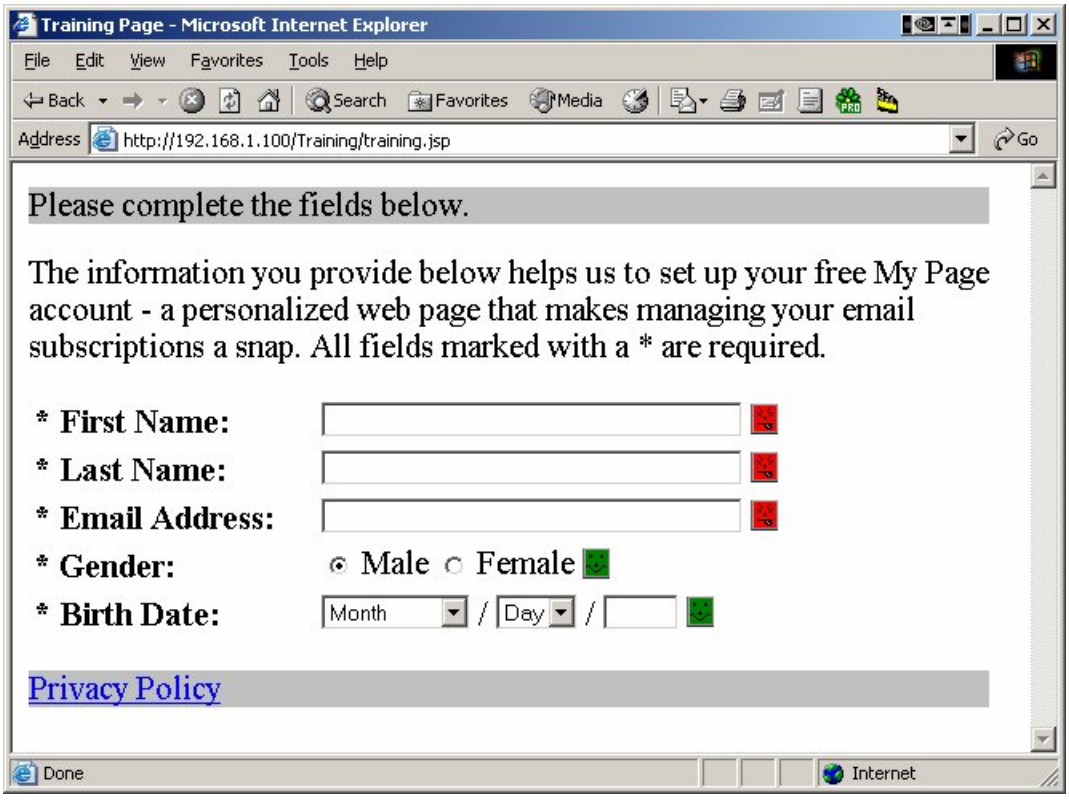


Figure 4.10: Training page utilizing the Integrated Privacy View

Each agent was demonstrated using the same training web page. The training pages for each agent are shown in Figures 4.9 and 4.10. Participants were shown how to accomplish the task for each agent. For the AT&T Privacy Bird a seven-step process is required to examine privacy policy conflicts. These steps are:

1. The bird icon in the upper right frame will turn red for one or more privacy conflicts and green if no privacy conflicts exist. The granularity of this indicator is typically an entire website but may be as fine as an individual web page. This indicator is used to determine the presence of a conflict.
2. Navigate the mouse pointer to the privacy bird and click
3. Slide the mouse pointer over "About this site"
4. Slide the mouse pointer over "Policy Summary …"
5. Click mouse button 1
6. Examine the privacy policy conflicts. This dialog is fairly large and obscures the original document. Based on the text the participant determines the field in conflict and the cause of that conflict
7. Dismiss the privacy conflict dialog.

For the Integrated Privacy View a two-step process is required to examine privacy policy conflicts. These steps are:

1. The red Mr. Yuck or green Mr. Smiley will appear next to the input field. A red indicator indicates a problem, a green indicator that the privacy policy is in agreement. This indicator is used to determine the presence and field in conflict
2. Navigate the mouse pointer to the icon next to the desired field. A popup appears stating the cause of the conflict. The popup does not obscure the desired field.

Demonstration of ability

Users were then asked to demonstrate that they were able to carry out these steps with each agent. Demonstration of ability ensures that the participant will be able to use both the

website's privacy policy and the privacy agent to accomplish the task during testing. The participant was allowed to interact with the training page until they felt comfortable that they understood how to use a privacy policy and each agent.

Testing

The usability task took place utilizing the websites described in section 4.2.1. Users were asked to complete the web transaction normally, but for each web page presented, the user had to determine if there was a privacy conflict, which field caused the conflict, and which privacy statement in the privacy policy constituted the conflict.

4.2.4 Measures of usability

The usability of IPV was measured using two types of assessment: records of user interaction, and a usability questionnaire.

User interaction records

The experimenter recorded specific interactions with the system during the study: identification of conflicts, number of steps taken to investigate a conflict, and correct identification of the cause of the conflict. The form in Figure 4.11 was used for each website and agent trial. In addition, the entire study was captured on audiotape, in order to support later review of participant observations.

Website →			Participant →			
Agent →						
Page 1	Conflict	noticed?	steps	reason?	check policy	Additional Comments
	email address					
	name					
Page 2						
	name					
	address					
	home phone					
	business phone					
	credit card					
Page 3						

Figure 4.11: Usability trial data capture form

Usability Questionnaire

Each participant completed a questionnaire to assess the effectiveness and efficiency of the user agents. The questions are described below.

Q1. Were you able to identify that there were privacy conflicts on a given page?

This question was designed to determine if either privacy agent made privacy conflicts noticeable. For AT&T, there was the concern that indicators placed out of context in the frame

border would not be noticed by a user. In addition, the AT&T Privacy Bird as implemented reports conflicts for the entire site, which might confuse the user on web pages where input is not required or input fields that are actually not in conflict with the user's privacy policy. The expected result is that IPV would make privacy conformance more visible than the privacy policy link or AT&T.

Q2. Were you able to identify the source of the conflicts?

This question was designed to determine if the granularity of conflict identification and indicating input fields that are not in conflict with the user's privacy policy facilitated the ability of the user to complete their task. The expected result is that IPV would more readily allow identification of the field in conflict due to the location of the indicator next to the target field.

Q3. Which system made it easier to identify that there were privacy conflicts?

This question was designed to determine if the IPV system was more useful to participants by providing specific conformance visualization in context of the web page and input form field.

Q4. Which system made it easier to identify the source of the conflicts?

This question was designed to determine if context sensitive conformance messages were accessible and understandable, and how the approach compared to the AT&T system.

Q5. Which system did you prefer overall?

This question was designed to determine if there was a preferred system and to solicit the reasons for that preference.

Q6. Did you notice that a privacy policy link was available on each web page?

This question was designed to determine if putting links to privacy policy on every page is not useful unless they are noticed and referenced by the user.

Q7. Did either system distract you from completing the form-filling task?

This question was designed to determine if adding the visualization of privacy conformance to the website distracts from the form filling task. If the icons excessively cluttered the web page, the user might be distracted.

Q8. Did either system help you feel more confident about completing the transaction on the website?

The interpretation of this question is perhaps better stated as "Did either system help you feel more confident about *deciding whether to complete* the transaction on the website;" (this is the interpretation that was used by all of the study participants). This question determines the utility of the thesis solution. Does the fine grained indication of privacy policy conformance give the user added confidence in deciding whether there are potential privacy problems – and if there are none shown, does this make the user more likely to complete a commercial web transaction?

4.2.5 Data collection for exploration of design issues

After participants had completed the comparison between IPV and Privacy Bird, they were asked to complete two further activities: a questionnaire concerning the design of the IPV agent that they saw in the comparison, and a design exploration with alternate interfaces.

The questionnaire included three questions that looked at issue of whether conformance icons are needed beside all input fields, and whether indicators are needed at the page level. The questions were:

Q9. For IPV, what did it mean to not have an icon next to a field?

Before the evaluation was performed, there was some concern that IPV icons might clutter the web page, distracting the user from filling out the form. The embodiment of IPV used for the evaluation placed an icon next to every field. One possible solution to clutter is not to place an icon next to fields that are in conformance. This question was designed to determine

if the user would understand the correct meaning of having no icon present for fields in conformance.

Q10. For IPV, what did it mean not to have any icons on a page?

Before the evaluation was performed, there was some concern that IPV did not have a global indicator to show conformance on web pages that had no input or all input was in conformance. This question was designed to capture what the lack of icons on a complete page meant to the user.

Q11. For IPV, should an icon indicate fields that have no privacy conflicts?

This question is a follow-up to question 9, to capture the design decision of the participant.

At the end of the session, participants were shown five alternative presentations for the IPV conformance indicators. These presentations showed different ways of displaying grouped conflicts (those originating from the same privacy policy statement) and showed visualizations that involved less visual clutter on the web page. Participants were asked to explore the visual presentations and mark them up with design suggestions to improve the user interfaces. An example of these alternate presentations is shown in Figure 4.12 (the other interfaces are shown in Appendix A). The stars in Figure 4.12 indicate the grouping of the two fields of the name and the grouping of the five fields of the address.

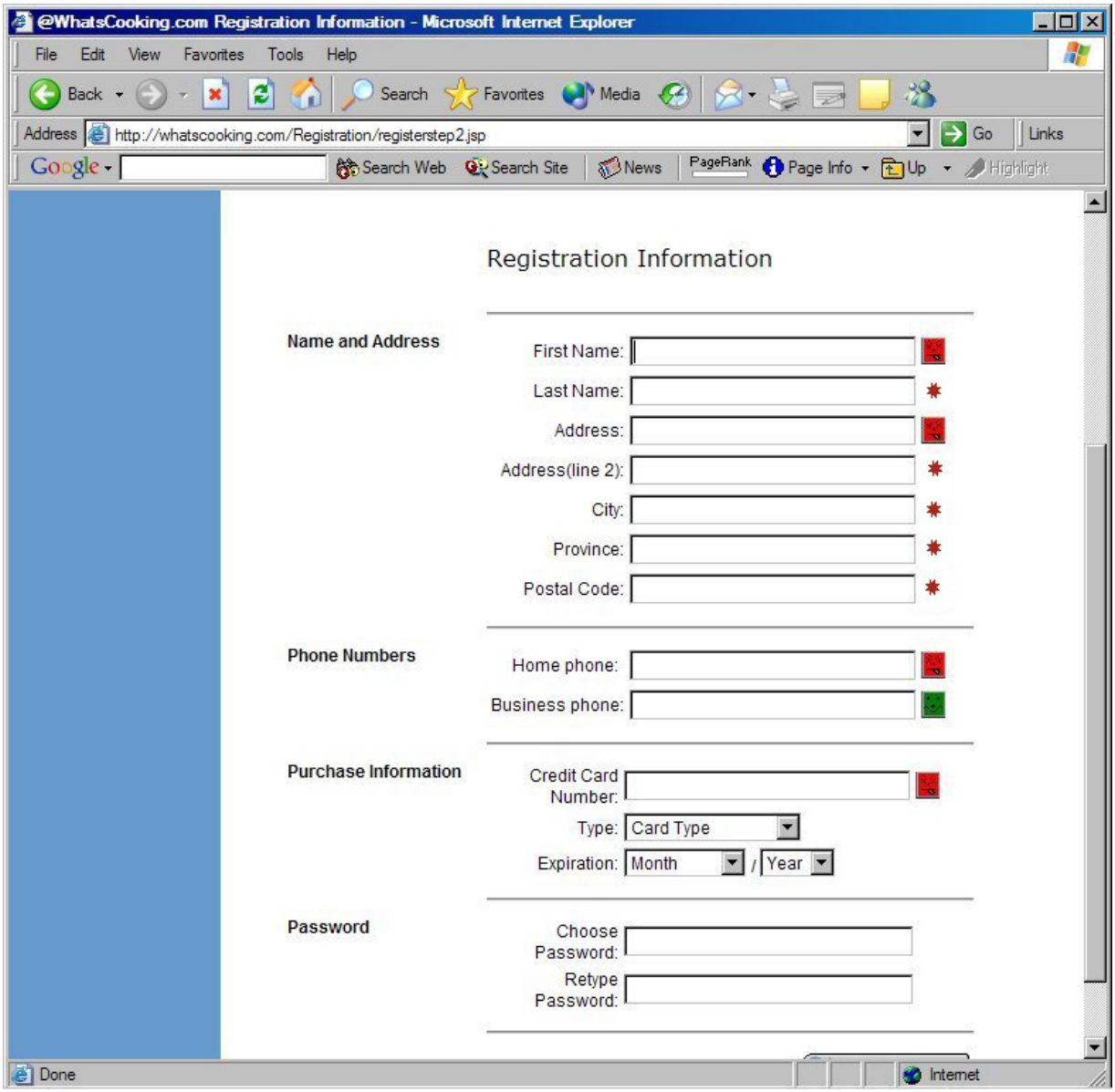


Figure 4.12. Example alternate presentation of conformance indicators

4.2.6 Procedure

The procedure that was used during the study is outlined table 4.2. The experimental set-up is shown in Figure 4.13.

Table 4.2: IPV Evaluation Procedure

Set-up

1. Start server and enable IPV proxy
2. Start Internet Explorer and load the training pages for the AT&T Privacy Bird and IPV

Materials

1. Consent Forms (2)
2. Demographic and Privacy Survey
3. Personal Identity and Information
4. Study recording sheet
5. Post Study questionnaire
6. Alternate presentations (3)

Steps

1. Review and have participant sign consent form. Provide copy of consent form to the participant at this time.
2. Have participant fill out demographic and privacy survey.
3. Introduce study utilizing Personal Identity and Information.
4. Perform training
    a. Show privacy policy
    b. Demonstrate each user agent
5. Have user demonstrate basic knowledge
    a. For each agent let them find the conflict, the field in conflict, and the reason.
6. Ask the user if there are any questions?
7. Begin Study

    Remind participant that they are concerned about their personal information. Specifically their fictional name, address, phone numbers and credit card information. The task on every page is to determine:

    a. Is there a conflict?
    b. What field is causing the conflict?
    c. What is the conflict?
8. Proceed to second site with other agent
9. Post study questionnaire
10. Alternate presentations
    a. Group indicator(s)
    b. Blank screen shot

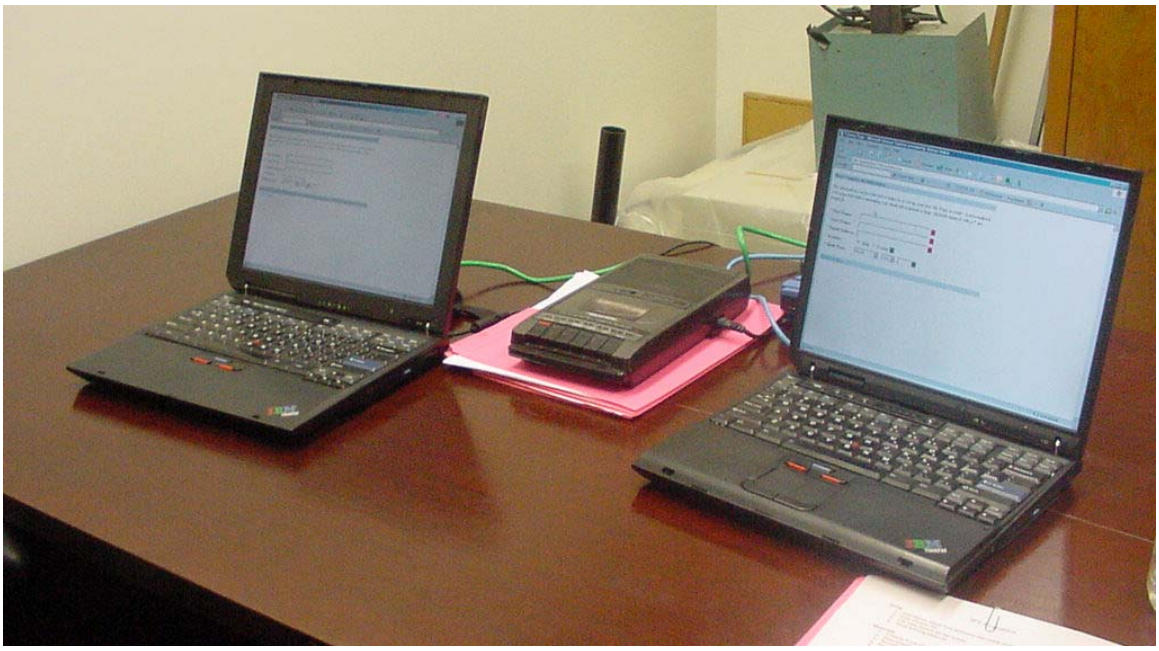


Figure 4.13: AT&T and IPV Evaluation Set-up.

4.3 Results

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

4.3.1 Usability of IPV

This section presents a summary of the participant's responses to each post evaluation survey question, combined with the observations of the study administrator, to provide an overall summary of each question. The questions in the first half of the survey were designed to help determine if IPV effectively communicates privacy policy information, and more importantly, whether it helped the user to feel more secure about disclosing personal information to the websites.

Question 1. Were you able to identify that there were privacy conflicts in a given page?

Success in identifying privacy conflicts on a specific web page was measured through Question 1 of the survey and through observations of actual activity during the study. All six of the study participants reported in response to question 1 that they were able to identify that there were privacy conflicts in a given page. The actual observed results varied for each web page. The results are shown in table 4.3.

Table 4.3: Visibility of Privacy Conflict

Privacy Agent	*Page 1*	*Page 2*	*Page 3*
AT&T	3 (50%)	6 (100%)	2 (33%)
IPV	6 (100%)	6 (100%)	2 (33%)

Three of the participants did not notice the AT&T Privacy Bird was red on page 1 of the website. The experiment administrator pointed this out to the participants. These participants were then able to proceed to determine that there was a conflict. This result indicates that indicators in the browser frame are not always noticed.

All participants determined that there were privacy conflicts on page 2, but not always accurately. Accuracy of privacy conflict source is discussed in Question 2.

Page 3 was the most difficult for the participants to evaluate. Only two of the participants correctly understood that the AT&T Privacy Bird was presenting privacy conflicts for all pages on the website at all times and only two of the participants understood that no indicator by IPV on a web page meant that the page was in conformance or did not require input. This result indicates that a global indicator should always present to indicate the correct conformance state for the specified page.

Q2. Were you able to identify the source of the privacy conflicts?

Participants' success in identifying the source of privacy conflicts on a specific web page was measured through Question 2 of the survey and through observations of actual activity during

the study. All six of the study participants reported that they were able to identify the source of the privacy conflicts in a given page. The actual observed results varied for each web page. The results for each input field are shown in Table 4.4.

Table 4.4: Accuracy of determining input field conflict

Privacy Agent	*Email Address*	*Name*	*Address*	*Home Phone*	*Business Phone*	*Credit Card*
AT&T	4 (66%)	6 (100%)	6 (100%)	6 (100%)	2 (33%)	6 (100%)
IPV	6 (100%)	6 (100%)	6 (100%)	6 (100%)	6 (100%)	6 (100%)

Utilizing the AT&T Privacy Bird, two of the participants thought that the conflicts for all the fields applied to Page 1, even though only one field on the page presented was in conflict. Only two users were able to accurately determine that the business phone was not in conflict using AT&T. These participants actually had to thoroughly review the detailed privacy policy to make this determination.

Q3. Which system made it easier to identify that there were privacy conflicts?

Question 3 measured which system made it easier to identify the existence of privacy conflicts on a specific web page. All six participants identified IPV as the easier system to identify the occurrence of privacy conflicts. The volunteered reasons for the preference are summarized in the Table 4.5.

The visibility of the icon was the number one reason stated for preferring IPV. Two participants stated that the placing of the AT&T Privacy bird in the frame of the border made the indicator much less likely to be noticed. Locating the conformance icon next to the input field was the second most stated reason for preferring IPV. This reason is explored in Question 4.

Table 4.5: Reasons for preferring IPV for identify privacy conflicts

Preference Reason	*Number of participants*
Visibility of icon	6
Location of icon	5
Specific information	2
Color	1

Q4. Which system made it easier to identify the source of the privacy conflicts?

Identifying which system made it easier to identify the source of privacy conflicts on a specific web page was measured through Question 4 of the survey. All six participants identified IPV as the system that better identified the source of privacy conflicts. The volunteered reasons by the participants for this preference are summarized in Table 4.6.

Table 4.6: Reasons for preferring IPV for identify source of privacy conflicts

IPV Preference Reason	*Number of Participants*
Location of icon next to field	6
Visibility of icon	2

Q5. Which System did you prefer overall?

Success in identifying which system was preferred by participants was measured through Question 5 of the survey and through observations of activity. All six participants preferred IPV. The volunteered reasons for the preference are summarized in Table 4.7.

The reasons are consistent with the results of questions 3 and 4. Location and visibility were the most cited reasons for preferring IPV, followed by presentation of field specific information. An unexpected result was that only one participant stated that fewer steps contributed to his preference, even though 5 additional steps were required to use the AT&T Privacy Bird agent.

Table 4.7: Reasons for preferring IPV overall

IPV Preference Reason	*Number of Participants*
Visibility of smiley faces and yuck icons	4
Location of icon next to field	4
Fewer clicks to obtain information	1
Field specific information	3
User friendly	2

Q6. Did you notice that a privacy policy link was available on each web page?

Success in noticing the privacy policy link on each web page was measured through Question 6 of the survey and through observations of activity. The results for Question 6 are shown in Table 4.8. Four of the participants utilized the full policy during the course of the AT&T trial. It was not utilized at all during the IPV trial. This result indicates that privacy policy links are not very visible.

Table 4.8: Privacy Policy visibility

Policy Visible?	*Yes*	*No*	*1 site only*	*1st page only*
Percentage of participants	3(50%)	1(16%)	1(16%)	1(16%)

Q7. Did either system distract you from completing the form-filling task?

Success in determining if either privacy agent distracted the participant from the form-filling task was measured through Question 7 of the survey. The results for question 7 are shown in Table 4.9. Two of the participants were distracted by the additional icons presented by IPV, but not because of visual clutter. One participant was very concerned about disclosing the credit card information once she understood what would be done with the information. The other participant was really interested at first in what information the IPV icons had to display, and spent several minutes exploring the icons and the attached privacy information before

continuing with the task. This is a desired result (at least initially), in that privacy conformance awareness allowed the participant to make an informed decision about completing the task. Nevertheless, the additional distraction may pose a problem for some web sites; this issue is explored further below.

Table 4.9: Distraction from task

System	*Yes*	*No*
AT&T	0 (0%)	6(100%)
IPV	2(33%)	4(66%)

Q8. Did either system help you feel more confident about completing the transaction on the website?

As discussed above, this question was reinterpreted as "Did either system help you feel more confident about *deciding whether to complete* the transaction on the website." Determining whether either privacy agent made the participant more confident was measured through Question 8 of the survey. All six participants reported that IPV made them feel more confident about the transaction; in addition, two participants also reported that Privacy Bird made them feel more confident. The results for question 8 are shown in Table 4.10. The major reason for the increase in confidence was that the systems raised awareness of what was going to happen with their personal information.

Note that this question does not suggest that with IPV, users are more likely to complete a web transaction – only that they can more confidently determine whether or not they would proceed based on a better understanding of the privacy conflicts. This was seen in the study with one participant in particular, who was very uncomfortable about completing the

form after IPV explained what was happening with the information. He appreciated that IPV made the explanation very clear – it was the underlying conflicts that were the problem.

Table 4.10: Transaction Confidence

Privacy Agent	*IPV*	*AT&T*
Participant Confidence	6(100%)	2(33%)

4.3.2 Visual Design Issues in Fine-Grained Conformance Display

Visual design issues were explored through the observations of the study administrator, the second half of the post study questionnaire, and a review of user interface alternatives with the study participants. These evaluation tools were designed to explore the issues of completeness in the match between input fields and indicators, distraction from the task of form filling, clustering and grouping, and to follow up any new issues that arose during the course of the study.

Q9. For IPV, what did it mean to not have an icon next to a field?

The meaning of no IPV icon was explored by Question 9 of the survey and through discussion of the alternate user interface. The results of Question 9 are shown in Table 4.11. One strategy to minimize visual clutter was to remove the green icon (Mr. Smiley) from those fields that were not in conflict. This strategy did not work well, with only 50% of the participants coming to the correct understanding. The other participants believed that either the information was not personal, not protected, or wasn't defined in the privacy policy. This result indicates that all fields should have an appropriate conformance indicator to avoid ambiguity of meaning.

Table 4.11: Meaning of no icon next to an input field

Meaning of no icon on field	*No Conflict*	*Not Protected*	*Not Personal*	*Not defined*
Percentage of Participants	3(50%)	1(16%)	1(16%)	1(16%)

Q10. For IPV, what did it mean not to have any icons on a page?

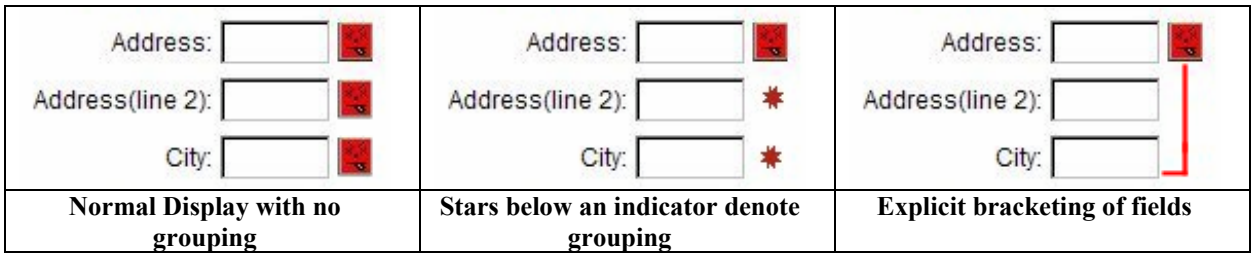
The meaning of no IPV icons at all on a web page was explored in Question 10 of the survey and in discussion of alternate user interfaces. The results for question 10 are shown in Table 4.12. The embodiment of IPV for this study did not have a global indicator. There is no visual indicator present on web pages where no input fields are present on the page. This strategy did not work well for the six participants. Three of the participants felt IPV was not working on that page, and one was unsure of what no icons meant. This result indicates that a global indicator embedded in the web page would avoid ambiguity in meaning.

Table 4.12: No Icons

Meaning of No Indicator Page	*No Problem*	*Not working*	*Not sure*
Percentage of Participants	3(50%)	2(34%)	1(16%)

Q11. For IPV, should fields that have no privacy conflicts be indicated by an icon?

Success in determining if the IPV icon should be present on fields with no privacy conflict was measured by Question 10 of the survey and through discussion of alternate user interface alternatives. All six participants were very emphatic that all fields should have a privacy icon. Visibility and convenience were the major reasons reported. Visual clutter was not an issue. The participants also preferred having the icon repeated for field groups, such as street address, rather than any sort of group or bracket indication (as in Figure 4.14).



Normal Display with no grouping	Stars below an indicator denote grouping	Explicit bracketing of fields

Figure 4.14: Grouping indicators for use when several fields are in conflict.

4.3 Discussion

The first part of the evaluation was designed to determine if visualizing privacy preference conformance at the input field level will improve the understanding of website privacy policies. The second part of the evaluation was designed to explore different visualization issues with IPV. The results indicated that co-locating privacy conformance indication with the input field and providing context sensitive privacy information does improve the understanding of website privacy policies, and that there are display issues when implementing the visual interface that can improve user understanding.

4.3.1 Summary of Usability Evaluation

IPV was the preferred privacy agent of all of the participants, and led to improved recognition and understanding of conflicts in the observational data. Both IPV and the AT&T Privacy Bird had access to the same privacy policy definition in P3P and the same privacy policy preferences in APPEL. The difference to the user was in the way the privacy conformance information was presented and accessed, and this appeared to be the source of users' preferences. The four major reasons cited by the participants (in rank order) were:

1. *Visibility of conformance icon*. The user task is to fill in an input field; during this task, their attention is drawn to the input field and not to the web browser frame. IPV demonstrated that placing the conformance indicator next to the input field increased

visibility. Distraction from task was not a problem in our study, although more work can be done to determine interactions between different visualizations and task attention.

2. *Explicit indication of privacy policy conformance*. This implementation of IPV provided a field-by-field indication of privacy policy conformance or conflict. There was no ambiguity over whether a specific input field privacy policy was in conflict. The AT&T Privacy Bird, by providing a page level indication, indicates a conflict if any field on the page does not conform to the user's privacy preference.

3. *Faster access to obtain conformance information.* The AT&T Privacy Bird requires seven steps to access the website privacy policy statements. IPV requires two steps. Participants liked the reduced effort to find out additional information.

4. *Not obscuring the input field.* Presenting conformance information without obscuring the input field removes the distraction of bringing up and dismissing a full-page dialog. The user can review the information they are entering while at the same time viewing the associated privacy policy statement. Users felt that this reduced the effort required to understand how the privacy policy statement related to the input field.

IPV improved the understanding of website privacy policies. Providing field specific information extracted from the complete website privacy policy saved users the effort of reading entire policy summaries and deciding which input fields the policy statements applied too. As one user put it, IPV removed the fine print and made it obvious what information the privacy policy statement was talking about.

IPV provided users with more confidence in deciding whether to complete transactions. Raising awareness of how the user's personal information was going to be used allowed the user to make informed decisions about disclosing personal information. Most

users were not aware of the policy information that was available to them. In the user's view, IPV presented the information they needed to know when they needed to know it, without any additional effort on their part. The increased understanding did not always lead to the user wanting to complete the transaction because of the content of the actual privacy policy statement. This result is still a positive outcome because it instilled trust in the user that the privacy agent was providing information that they needed to know.

4.3.2 Summary of Visual Design Exploration

This implementation of an IPV agent puts an icon next to every input field to indicate conformance. This choice led to concerns before the evaluation that a web page might appear to be cluttered and that icons present at the bottom of long pages would not be visible. Additional design issues explored during the evaluation were concerned with the misinterpretation of an input field without a conformance icon and the misinterpretation of web pages that had no conformance icons present at all.

Clutter turned out not to be a major design issue. The participants liked the fact that each field was readily identified, even if there was no conformance conflict. Although alternate presentations were presented with grouping, the participants still preferred to see a conformance indicator next to each field. The fact that some conformance icons were only visible when the web page was scrolled was also not found 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.

Misinterpretation of a "missing" conformance icon next to an input field was common. This implies that all input fields should have a privacy policy statement identified, even if the

information is not retained by the website. This is consistent with the participant's responses that all fields should be clearly identified with privacy conformance.

Misinterpretation of no conformance icons on the web page was also common. This implies that a global indicator should be present, even though no information is retained or requested by the website. This indicator should be in the web page, and not in the browser border, in order to ensure its visibility to the user.

4.3.3 Generalization of Results

Although the study was small, there are reasons to suggest that the experiment results will generalize to web transactions in real-world task situations. This evaluation as closely as possible modeled a real world consumer experience. The participants and tasks were chosen to model as closely as possible real experience and activities on the Internet. The participants in this evaluation represent a good sample of typical Internet users. The participants used the Internet frequently, were experienced in filling out forms on web sites, but had never used a privacy agent before. All knew that websites had privacy policies but only a few had ever read them because of their complexity. The task of filling out the registration forms on the WhatsCooking and AllThatJazz represented tasks that the participants were experienced with. None of the users required training in how to proceed through the sites and complete the task.

However, the evaluation did not model long time experience with IPV; the evaluation was a one-time experience. Although the participants were adamant about conformance indicators on every field and a global indicator on every page, longer-term opinions might be different. As the novelty of the icons wore off they might prefer to have fewer icons on the page or different presentations (e.g. smaller icons, highlights) to reduce visual clutter.

Permitting customization of visual style and presence would let IPV serve better both the casual and expert user.

4.3.4 Issues for Practitioners

Utilization of IPV impacts privacy policy designers, web page designers, and IPV privacy agent designers. Each of these groups needs to understand how to correctly utilize IPV to achieve maximum benefit.

Privacy policy designers today must identify the personal data requirements for web application and then develop a P3P privacy policy that reflects the corporate privacy policy. IPV requires the privacy policy designer to refine that privacy policy so that each statement is correctly bound to a personal data group or element. P3P policy development tools already permit this.

Web page designers need to utilize the p3pdataelement attribute to link the work of the privacy policy designer with the input field. WYSIWIG web page design tools, like ECLIPSE (Shavor et al., 2003), could easily be extended to create web page input fields, which automatically read the P3P privacy policy and let the web page designer choose which data element to place in the p3pdataelement. This would minimize the effort to implement IPV in the web page and add the benefit of crosschecking 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 not to render correctly if the input fields are placed 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.

The IPV agent designer needs to provide an agent that takes advantage of the information provided by the privacy policy designer and the web page designer. The conformance indication must clearly be associated with the input field, require minimal steps to obtain additional information, and not obscure the input field while the user is viewing the additional information. This evaluation shows that conformance indication should be present on every input field and a global indicator present on every page.

The privacy agent should provide customization for both style and content. Colors for conflict and conformance should be selectable. If icons are present in the interface, the user should be allowed to alter or replace them. Customization of conformance indication presence and location would aid expert users in reducing the amount of information presented as desired. This can be explored as future work, as described in the next chapter.

Chapter 5 Conclusions

The problem addressed in this thesis was: **It is difficult for a user to determine the cause and origin of a conformance conflict between the user's privacy preferences and a website's privacy policy**. The main motivation for improving understanding of privacy conformance is to improve user confidence and trust in the disclosure of personal information, permitting more e-commerce transactions to be successfully completed. The solution explored in this thesis is that **user understanding can be improved by visualizing privacy preference conformance at the input field level**. This solution has two main parts: first, refining the mapping of privacy policy to input fields, and second, providing a contextual display of conformance on the web page.

5.1 Summary of Research and Contributions

The implementation of IPV provides three enhancements to P3P conformance display:

1. *User privacy preference conformance is specific to a given input field.* The p3pdataelement attribute allows the vendor to identity the specific policy statements that are associated with an input field. This permits the user to interpret conformance information specific to each input field.
2. *User privacy preference conformance is displayed in context to a given input field.* Conformance indication is presented next to the associated input field, allowing the user to view and interpret the conformance information specific to the adjacent input field.
3. *Fast access to additional conformance information.* Detailed conformance information is available by rolling over the visual indicator with the mouse cursor. This permits

casual access to additional information without changing the context of the user's work area to another dialog window.

The evaluation showed that by providing more visibility, faster access, and context sensitive policy information, the user was able to understand better how their personal information was going to be used. This improved the user's confidence in deciding whether to disclose personal information.

The major contribution of this research is the design and demonstration of a scheme that allows fine-grained comparison and display of privacy conformance information for web transactions. The scheme extends the machine-readable privacy policy standard (P3P) to permit the integration of policy statements at a fine-grained level with the vendor's input form. This allows web form designers to link specific HTML input fields with specific privacy policy statements. This link improves understanding of privacy conflicts by permitting an adaptive presentation of conformance information. This is a better visualization of privacy policy and related conformance information than the current state of the art.

There are also several secondary contributions from this thesis:

- A better understanding of the visual issues involved in presenting conformance visualizations. This thesis documents user requirements of when to show conformance indication, where to show conformance indication, and how to present additional conformance information.
- A better way for the privacy policy designer to collaborate with the web page designer. Implementation of the P3P attribute allows the designers to ensure the privacy policy is complete and reflects all personal data gathered. Missing or inaccurate privacy policy statements will be detected during web page implementation.

- A better understanding of how P3P agents can be characterized as adaptive hypermedia agents. The IPV agent is the first adaptive hypermedia P3P user agent to adapt privacy meta-data into the HTML content presented to the user.
- A reference implementation of the IPV prototype, which shows how a user agent can take advantage of the improved P3P implementation specification, and shows that the fine-grained approach can be carried out without compromising browser performance and without adding undue complexity to design of the user agent.

5.2 Future Work

To continue to develop IPV and explore privacy agent implementations, two main avenues for future work can be considered: changes to the implementation and testing of IPV, and changes in the way that IPV deals with the user's privacy preferences.

5.2.1 System changes and improvements

The first tasks that should be considered in future work are several enhancements to the implementation of the IPV agent so that all Internet users may use it. These are listed here in the assumption that the use of a fine-grained tag such as p3pdataelement will eventually become common in vendor web sites, providing the basic material for an IPV user agent to operate on.

The current implementation of IPV as an HTTP proxy is somewhat unwieldy, and a fully usable solution requires that the IPV agent be easily installable and provide data security. The most widely accepted method of doing this is to implement the agent as a browser helper object. This would permit the IPV agent to be downloaded from the web and installed at the

user's request. Being embedded in the browser avoids data security issues, since the browser handles all the HTTP communication.

The second enhancement is to enable customization of the conformance visualization by the user. The plan is to extend APPEL to include the style and detail of IPV visualizations. This would permit the user to specify:

- The presence of global indicator
- The presence of indicator on a conforming field
- Color to be used in conformance indication
- Choice of conformance icons
- Choice of input field conformance highlighting

This extension would allow the user to tailor the visualization to be as informative as the user's experience required with minimal visual distraction. Other extensions would allow the site designer to specify those visual customizations that still enabled the web page to be rendered properly.

Third, the findings from the evaluation should be tested in a longer-term study where participants have the opportunity to use the fine-grained visualization for their own web transactions. This requires that real Internet vendors implement the p3pdataelement tag, but could be carried out in a small scale trial.

5.2.3 Extending IPV to become an adaptive agent

A larger-scale change to IPV involves the way that users build up their privacy preferences. The current user-adapted approach requires the user to decide their preferences for a large number of personal data requests and is not sensitive to task the user is performing. Extending the user model to be an adaptive one would further decrease the effort to define

privacy preferences. A collaborative user model based on peer or expert experience would enhance user confidence in their privacy preference choices.

One approach is to extend the current definition of a P3P user agent to permit it to join an agent community. The user agent provides information on what privacy policies the user supported by that agent accepting or declining the policies and reporting that to the community. The community maintains profiles of user and web site policies that were accepted or declined. This community then provides assistance to a user when accessing a vendor website. It permits the user to see whether the privacy policy was generally accepted by other users and/or whether users that the consumer trusts have accepted the policy. The user can then decide whether to add this statement of privacy preferences to their own profile. In this matter, the user is provided with guidance on what an acceptable privacy policy is, in a way that is specific to particular vendors, and with the ability to "bootstrap" their own collection of privacy policy preferences (again specific to vendors).

In addition, the community classifies vendor policies into what domain they belong to, such as mortgage and finance companies, realtors, or retailers (booksellers, etc.). For example, a particular bookseller's privacy policy is evaluated against the community of booksellers for conformance. The user is then able to determine whether this particular bookseller's privacy policy is consistent with other booksellers. For instance, it could be determined that a bookseller should not be asking for disclosure of your social insurance number, as this is not generally a required piece of information when completing a sales transaction. Expert policies for each domain are created either by knowledgeable experts or as a result of popular opinion. The expert policies aid the user in deciding whether to worry about disclosure of a particular piece of data for a given domain.

Within this community framework, a user model is maintained based on the user's privacy preferences. This includes setting preferences for each piece of private information, such as email address, mailing address, Social Insurance Number, etc. Four warning levels could be specified:

- Warn always when this data is requested.
- Respect the expert opinion for this domain.
- Respect the opinion of my peers if the confidence level exceeds a specified percentage.
- Don't care—never warn about this field.

For example, an agent for a moderately concerned user might decide to trust the expert's opinion where a more cautious user would always want to be warned. The result is guidance to the user in understanding whether the data requested is appropriate for the domain and how concerned the user should be about its collection. These and other community-based techniques show great potential for overcoming another major hurdle in reducing user effort in dealing with privacy issues in web transactions.

References

1. Aberdeen Group (2002), *Federated Identity Systems – An Executive White Paper*, Technical Report, Aberdeen Group, Boston, MA, June 2002

2. ABA (2000), *American Banking Association Privacy Principles* <http://www.aba.com/about+aba/aba_privprinpublic.htm> Accessed May 15, 2003.

3. Ackerman, M.S., and Cranor L. (1999) *Privacy Critics: UI Components to Safeguard Users' Privacy*. Proceedings of CHI 99. pp. 258-259.

4. Ackerman, M.S., Cranor L.F., and Reagle, J. (1999*) Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences.* ACM conference on Electronic Commerce. pp. 1 - 8

5. Adkinson, W. F., Eisanach, J. A., and Lenard, T.M. (2002) *Privacy Online: A Report on the Information Practices and Policies of Commercial Websites*. March 2002 <http://www.pff.org/publications/privacyonlinefinalael.pdf>. Accessed May 15, 2003

6. Agrawal, R., Kiernan, J., Ramakrishnan, S., Xu, Y. (2003) *An Xpath-based Preference Language for P3P.* WWW2002, May 2014, 2003, Budapest Hungary

7. Altman, I. (1975). *The Environment and Social Behavior*. Monteray California. Brooks/Cole Publishing.

8. Ashley, P., Hada, S., Karjoth, G., and Schunter, M. (2002) *E-P3P Privacy Policies and Privacy Authorization. WPES'02*, November 21, 2002, Washington, DC, USA

9. AT&T (2002) AT&T Privacy Bird <http://www.privacybird.com>. Accessed on May 15, 2003

10. Behrens, L. (2001) *Privacy and Security: The Hidden Growth Strategy.* In Gartner G2, May 31 2001

11. Bohrer, K., Levy S., Liu X, Schonberg E. (2003) *Individualized Privacy Policy Based Access Control.* IBM Research Report RC22756

12. Boyle, M., and Greenberg, S. (2003) *Privacy in Video Media Space*s. Working copy: Draft in preparation.

13. Brusilovsky, P. (1994) *Adaptive hypermedia: an attempt to analyse and generalize*. http://wwwis/win.tue.nl/ah94/brusilovsky.html Accessed May 15, 2004-06-21

14. Brusilovsky, P. (1999) *ADAPTS: Adaptive Hypermedia for a Web-based Performance Support System* http://wwis.win.tue.nl/asum99/brusilovsky/brusilovsky.html Accessed May 15, 2004

15. Byers, S., Cranor, L., Kormann, D. (2003) *Automated Analysis of P3P-Enabled Web Sites* ICEC 2003. Pittsburg, PA

16. Cannataro, M., and Puliese, A. (2001*) XAHM: an XML-based Adaptive Hypermedia Model and its implementation.* University of Calabria Italy. <http://wwwis.win.tue.nl/ah2001/papaers/cannataro.pdf> Accessed May 15, 2004

17. Card, S. K., Mackinlay, J. D., Shneiderman, B. (Eds.) (1999) *Readings in Information Visualization: Using Vision to Think.* San Francisco, California. Morgan Kaufman Publishers.

18. Clarke, R. (1999) *Internet privacy concerns confirm the case for Intervention*. Communications of the ACM. February 1999. Vol. 42, No. 2, pp. 60 – 67

19. Clark, R. (1997) *Introduction to Dataviellance and Information Privacy and Definition of Terms* <http://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html>. Accessed May 15, 2003

20. Congress (1996) *Health Insurance Portability and Accountability Act of 1996* USA. Public Law 104-191. <http://www.aspe.hhs.gov/adminsimp/pl104191.htm>. Accessed May 15, 2003.

21. Constantine, L. (2002) *Devilish Details: Best Practices in Web Design* forUse 2002 Proceedings <http://www.foruse.com/articles/details.pdf > Accessed May 15, 2004

22. Coyle, K. (2001) *P3P: Pretty Poor Privacy? A Social Analysis of the Platform for Privacy Preferences (P3P).* <http://www.kcoyle.net/p3p.html>. Accessed May 15, 2003

23. Cranor L. and Reidenberg J. (2002) *Can user agents accurately represent privacy notices? Discussion draft*, August 2002. <http://intel.si.umich.edu/tprc/papers/2002/65/tprc2002-useragents.PDF>. Accessed May 15, 2003

24. Cranor, L.F. (2002) *Web Privacy with P3P.* O'Reilly and Associates, Cambridge Massachusetts

25. Cranor, L., Arjula, M., and Praveen G. (2002) *Use of P3P User Agent by Early Ado*pters, in Proceedings of the ACM Workshop on Privacy in the Electronic Society, November 21, 2002. pp. 1-10

26. Culnan, M., and Milne, G. (2001) *The Culnan-Milne Survey on Consumers & Online Privacy Notices: Summary of Responses.* In Interagency public workshop (ed.) *Get Noticed: Effective Financial Privacy Notices,* Washington, D.C.

27. Cyber Dialogue (2001) *UCO Software to Address Retailers 6.2 Billion Privacy Problem,* Press Release http://www.cyberdialogue.com/news/releases/2001/11-07-uco-retail.pdf Accessed May, 15, 2004

28. Earp, J.B., and Baumer, D. (2003) *Innovative Web Use To Learn About Consumer Behavior and Online Privacy*, Communications of the ACM, April 2003. Vol. 46, No. 4, pp. 81-83

29. Esposito, D. (1999) *Browser Helper Objects: The Browser the Way You Want It*. MSDN Library. <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnwebgen/html/bho.asp> Accessed May 15, 2004

30. European Union (2000) *On the protection of individuals with regard to the processing of personal data*. European Parliament and the Council of the European Union Act No. 77/2000 <http://personuvernd.is/tolvunefnd.nsf/pages/1E685B166D04084D00256922004744AE>. Accessed May 15, 2003

31. Findlater, L., and McGrenere, J. (2004) *A comparison of static, adaptive, and adaptable menus* Proceedings of the 2004 conference on Human factors in computing systems pp. 89-96

32. Fink, J., A. Kobsa and Schreck, J. (1997): *Personalized Hypermedia Information through Adaptive and Adaptable System Features: User Modeling, Privacy and Security Issues*. In: A. Mullery, M. Besson, M. Campolargo, R. Gobbi and R. Reed, eds.: Intelligence in Services and Networks: Technology for Cooperative Competition. Berlin Heidelberg: Springer, 459-467. http://www.ics.uci.edu/~kobsa/papers/1997-IS&N'97-kobsa.pdf Accessed May 15, 2004

33. Fox, S. and Rainie, L. (2000) *Trust and Privacy Online: Why Ammericans Want to Rewrite the Rules.* Pew Internet & American Life Project, Washington DC. <http://www.pewinternet.org/reprots/toc.asp?report=19> Accessed May 15, 2004

34. Furnell, S.M., and Karweni, T. (1999) *Security Implications of Electronic Commerce: A Survey of Consumers and Businesses.* Internet Research. Vol. 9, No. 5, pp. 372-382

35. GBDE (2001). *Consumer Confidence – Personal Data Privacy Protection*. Global Business Dialog on Electronic Commerce. <http://www.gbde.org/acrobat/recommendations01.pdf>. Accessed May 15, 2003

36. Hoffman, D.L., Novak, T.P., and Peralta, M. (1999) *Building consumer trust online.* Communications of the ACM. April 1999. Vol. 42, No. 4, pp 80-85

37. Internet Society (1999) HyperText Transfer Protocol – HTTP/1.1 <http://www.w3.org/Protocols/rfc2616.html> Accessed May 15, 2004

38. Ipsos-Reid and Columbus Group (2001) *Privacy Policies Critical to Online Consumer Trust.* Canadian Inter@active Reid Report

39. Jensen, C., Potts, C. (2004) *Privacy Polices as Descision-Making Tools: An Evaluation of Privacy Notices.* CHI 2004, April 24-29, 2004, Vienna, Austria.

40. Karjoth, G., Schunter, M., and Waidner, M. (2002*) The Platform for Enterprise Privacy Practices – Privacy-enabled Management of Customer Dat*a. Proceedings of the Privacy Enhancing Technologies Conferenc*e*, San Francisco, CA, April 14-15, 2002

41. Kaufman, J.H., Edlund, S., Ford, D.A., Powers, C. (2002*) The Social Contract Core*. Proceedings of the eleventh international conference on World Wide Web 2002. pp. 210-220

42. Langheinrich, M, Nakamura, A., Abe, N., Kamba, T., Koseki, T. (1999) *Unintrusive Cusomization Techniques for Web Advertising* NEC Corporation, C&C Media Research Laboratories <http://www8.org/w8-papers/2b-customizing/unintrusive/unintrusive.html>

43. Mackay, W. (1991) *Triggers and Barriers to Customizing Software*, Proc. ACM CHI 1991, 153-160.

44. Mackay, W. (1990) *Patterns of Sharing Customizable Software*, Proc. ACM CSCW 1990, 209-221.

45. MacLean, A., Carter, K., Lovstrand, L., and Moran, T. (1990) *User-Tailorable Systems: Pressing the Issues with Buttons*, Proc ACM CHI 1990, 175-182.

46. Maes, P. (1994) *Agents that Reduce Work and Information Overload.* Communications of the ACM. July 1994. Vol. 37, No. 7. pp. 30-40

47. Manbar, U., Patel, A., Robinson, J. (2000) *Experience with Personalization at Yahoo!* Communications of the CAN, Special issue on Personalization, (August 2000), pp. 35-40

48. Microsoft Office 2003 (2003) <http://www.microsoft.com/office/editions/prodinfo/default.mspx> Accessed May 15, 2004

49. Morch, A. (1997) *Three Levels of End-User Tailoring: Customization, Integration, and Extension.* In Computers and Design in Context. M. Kyng & L. Mathiassen (eds.). MIT Press, Cambridge, 51-76.

50. OECD (1980), *Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*. Organization for Economic Co-operation and Development. <http://www.cdt.org/privacy/guide/basic/oecdguidelines.html>. Accessed May 15, 2003

51. Page, S., Johnsgard, R., Albert, U., Allen, C. (1996) *User Customization of a Word Processor*, Proc ACM CHI 1996, 340-346.

52. Reagle, J., and Cranor, L.F., (1999) *The Platform for Privacy Preferences*. *Communications of the ACM*. February 1999. Vol. 42, No.2, pp. 48-55.

53. Roy Morgan Research (2001) *Privacy and Community*. Prepared for the Office of the Federal Privacy Commssioner, Sydney, July 31, 2001 <http://privacy.gov.au/publications/rcommunity.html> Accessed May 15, 2003

54. Rubin, J., (1994) *The Handbook of Usability Testing*. John Wiley and Sons, Inc, NY 1994

55. Shavor, S., D'anjou, J., Fairbrother, S., Kehn, D., Kellerman, J. McCarthy, P. (2003) *The Java Developer's Guide to Eclipse.* Addison-Wesley, Boston Massachussetts

56. Shneiderman, B. (1997) *Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces*. Proc. ACM International Workshop on Intelligent User Interfaces'97. pp. 33-39

57. Shneiderman, B. (2000), *Designing Trust Into Online Experiences*, Communications of the ACM. December 2000. Vol. 43, No. 12, pp. 57-59

58. Siau, K., and Shen, Z. (2003) *Building Customer Trust in Mobile Commerce*, Communications of the ACM. April 2003. Vol. 46, No. 4, pp 91-94

59. Spiekermann, S., Grosslags, J., Berendt, B. (2001) *E-privacy in 2nd Generation E-commerce: Privacy Preferences versus actual Behaviour*. Proceeding of the 3rd ACM conference on Electronic Commerce. October 2001. pp 38-47

60. Teltzrow, M. and A. Kobsa (2004) *Impacts of User Privacy Preferences on Personalized Systems: a Comparative Study*. In: C.-M. Karat, J. Blom and J. Karat, eds: Designing Personalized User Experiences for eCommerce. Dordrecht, Netherlands: Kluwer Academic Publishers, 315-332. <http://www.ics.uci.edu/~kobsa/papers/2004-PersUXinECom-kobsa.pdf>

61. Trigg, R., and Bodker, S. (1994) *From Implementation to Design: Tailoring and the Emergence of Systemtization in CSCW*, Proc. ACM CSCW 1994, 45-54.

62. Trigg, R., Moran, T., and Halasz, F. (1987) *Adaptability and Tailorability in NoteCards*, Proc. IFIP INTERACT 1987, 723-728.

63. UMR (2001) *Privacy Concerns Loom Large.* Conducted for the Privacy Commisioner of New Zealand. <http://www.privacy.org/nz/privword/42pr.html> Accessed May 15, 2003

64. Westin, A. (1967). *Privacy and Freedom*. New York, NY: Bodley Head Publishers

65. W3C (1998) Document Object Model(DOM) Level 1 Specification, W3C Recommendation http://www.w3.org/TR/REC-DOM-Level-1/ Accessed May 15, 2003

66. W3C (2002a) *A P3P Preference Exchange Language 1.0 (APPEL 1.0) Working Draft*, <http://www.w3.org/TR/P3P-preferences>. Accessed May 15, 2003.

67. W3C (2002b) *The Platform for Privacy Preferences 1.0 (P3P 1.0) Recommendation*, <http://www.w3.org/TR/P3P>. Accessed May 15, 2003

68. Xerces2 Java Parser (2002) The Apache XML Project, <http://xml.apache.org/xercws2-j/index.html> Accessed May 15, 2003

69. Xalan-Java (2002) The Apache XML Project, <http://xml.apache.org/xalan-j/index.hmtl> Accessed May 15, 2003

70. XML Path Language (XPATH) Version 1.0 Specification, (1999) W3C Recommendation. <http://www.w3.org/TR/xpath> Accessed May 15, 2003

71. XSL Transformation (XSLT) Version 1.0 Specification, (1999) W3C Recommendation. <http://www.w3.org/TR/xslt> Accessed May 15, 2003

Appendix A: Evaluation Materials

1. **Informed consent form**
2. **Demographics and privacy survey**
3. **Instructions to participants and fictitious personal information**
4. **Post-study questionnaire**
5. **Alternate interface presentations of conformance**

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: Improving Understanding of Website Privacy Policies
Investigators: Steve Levy, Department of Computer Science
Dr. Carl Gutwin, Department of Computer Science (966-8646)

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This study is concerned with investigating ways computer systems can be used to improve the understanding of website privacy policies. The computer system was developed at the University of Saskatchewan in the Department of Computer Science.

In this experiment you will be asked to navigate to a simulated website for downloading software. You will be asked questions about what the website will do with any personal information that is requested and how that compares to privacy preferences. Both the privacy preferences and personal information to be used will be supplied to you.

Your responses may lead to improvements in preserving personal privacy on the Internet. No personally identifiable information will be collected about you. Recording of your opinions and actions during the experiment by the observer will be stored only by participant number. There is no personal risk to you in this experiment.

You are free to withdraw from this study at any time without penalty and without losing the 10 dollar honorarium. If you withdraw, your data will be deleted from the study and destroyed.

The results of this study will be used in Steve Levy's Master of Science Thesis and may be included in journal articles or department technical reports.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Dr. Carl Gutwin, Associate Professor Department of Computer Science (306) 966-8646
  gutwin@cs.usask.ca
- Office of Research Services University of Saskatchewan (306) 966-2084

Participant's signature: _____ Date: _____

Investigator's signature: _____ Date: _____

A copy of this consent form has been given to you to keep for your records and reference. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on (date).

Demographic and Privacy Survey

To help us interpret the results of this survey we would like to ask some personal information about you.

Please feel free to skip any question in the survey you prefer not to answer.

1. Please circle your gender:

   Female *Male*

2. Please circle your age range:

   18 -26 *27-35* *36-42* *43-50* *over 50*

We would like to ask you some questions about conducting personal business over the Internet.

1. How would you rate your experience in using the Internet to purchase goods and services?

   Beginner *Novice* *Average* *Sophisticated* *Expert*

2. Please circle how often you make purchases on the Internet:

   Never *Occasionally* *Once or twice a month* *3 or 4 times a month* *Whenev possible*

3. Do you look for a privacy policy or a security statement on a vendor's website?

   Yes *No* *Sometimes*

   If you have read a vendor privacy policy, what were you looking for the policy to t you?

4. Do you know when your browser is in secure mode?

   Yes *No* *I don't know what that is*

   If yes, do you look for the secure mode indicator every time you disclose personal information?

   Yes *No* *Sometimes*

5. Do you take steps to protect your email address when a website asks for it?

   Yes *No*

   If yes, what steps do you take to protect your email address?

6. Do you take steps to protect your credit card information when requested to provid your credit card information to a website?

   Yes No

   If yes, what steps do you take to protect your credit card information?

7. Have you used a user privacy agent?

   Yes *No* *I don't know what that is*

If yes, which one?

Personal Identity and Information

This study looks at two ways information about vendor privacy policies may be presented to a user to help the user to protect their privacy. The systems being tested are user privacy agents that compare your privacy preferences to the privacy policy of the website and indicate where there are conflicts between your privacy preference and the vendor privacy policy.

The interfaces the privacy agents present are experimental. There will be an opportunity for you to comment on the interfaces and make recommendations at the end of the study.

For the purpose of this study, please assume that you are very concerned about disclosing your:

Name
Address
Home Phone
Email Address
Credit Card Information

You **always** want to understand how the above information will be used by a website so you can make an informed decision of whether to disclose the information or not.

I would like you to assume that you would not like to disclose the information above for any purpose. You may decide to disclose the information if the website is offering you something you really want. **For the purpose of this study, after understanding what information and for what purpose you are disclosing information to the website, you will provide the information the website.**

The information to use is provided below. **Please be sure to use the information provided!** I do not want you to disclose any of your own information.

You will be trained on three ways of determining how your personal information will be used by a website. You will then be presented with a task to supply the personal information provided, determining for each task if there is a privacy concern, what personal data is involved, and what the reason for the concern is.

You will be completing the task twice, once for each sample website, utilizing a different privacy agent for each site.

Please turn over for personal information to use

Personal Information to be used for the study

Name:	Jane Smith or John Smith
Gender:	Female or Male
Date of Birth:	May 21, 1980
Address:	101 Paradise Place Saskatoon, SK s7k 4m7
Home phone:	306 242 2389
Business phone:	306 242 5734
Email Address:	jsmith@paradiseplace.ca
Credit Card:	VISA 1111 3333 5555 6666 Expires October, 2004
Password:	t00ntown

Post Study Survey

We would like to ask you some questions about your experience using the two user privacy agents. (Administered by the study conductor)

1. Were you able to identify that there were privacy conflicts on a given page? If, not why?

2. Were you able to identify the source of the conflicts? If not , why?

3. Which system made it easier to identify that there were privacy conflicts? Why?

4. Which system made it easier to identify the source of the conflicts? Why?

5. Which system did you prefer overall? Why?

6. Did you notice that a privacy policy link was available on each web page?

7. Did either system distract you from completing the form filling task?

8. Did either system help you feel more confident about completing the transaction on the website?

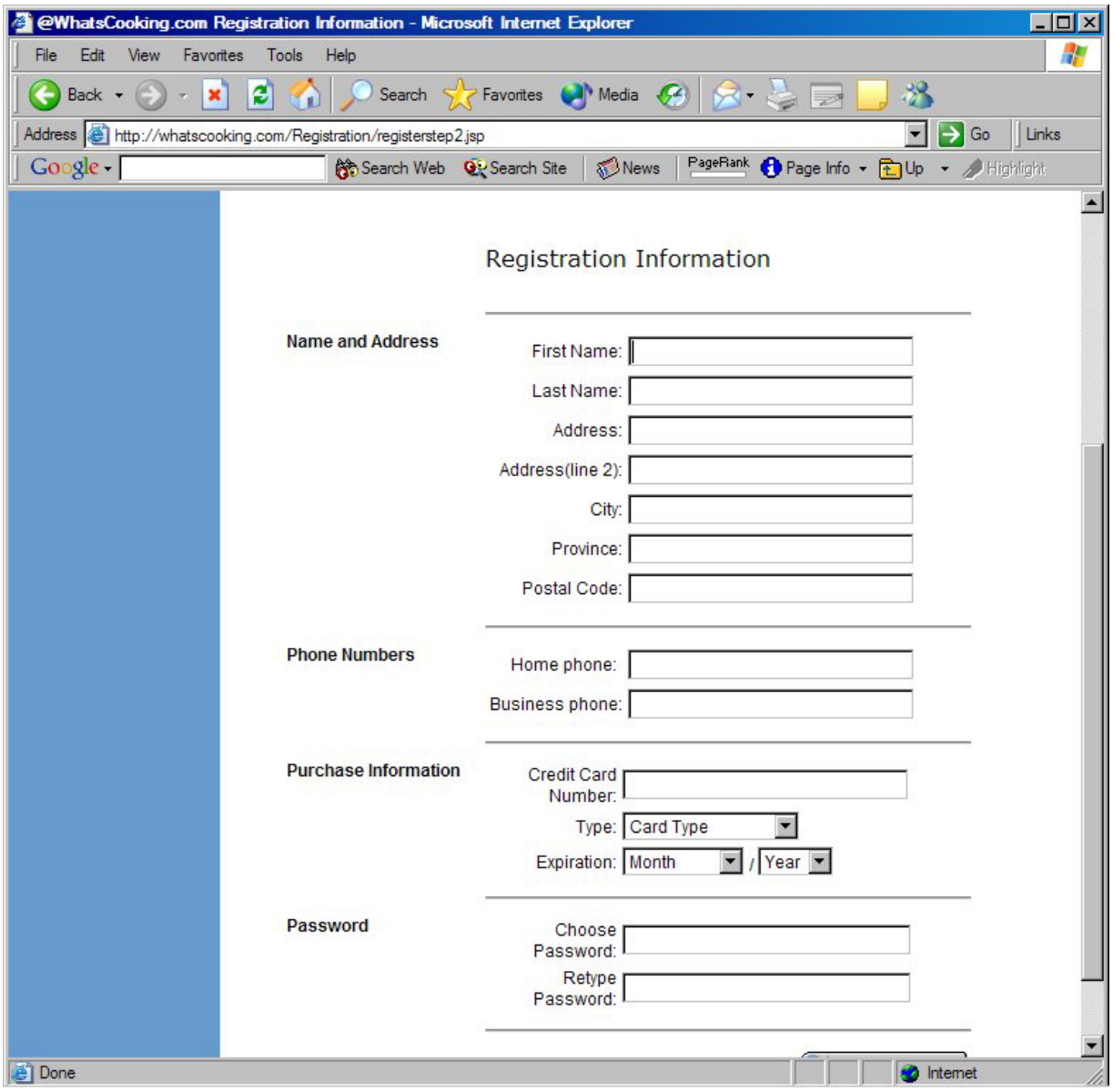
9. For system 2, what did it mean to not have an icon next to a field?

10. For system 2, what did it mean not to have any icons on a page?

11. For system 2, should fields that have no privacy conflicts be indicated by an icon?

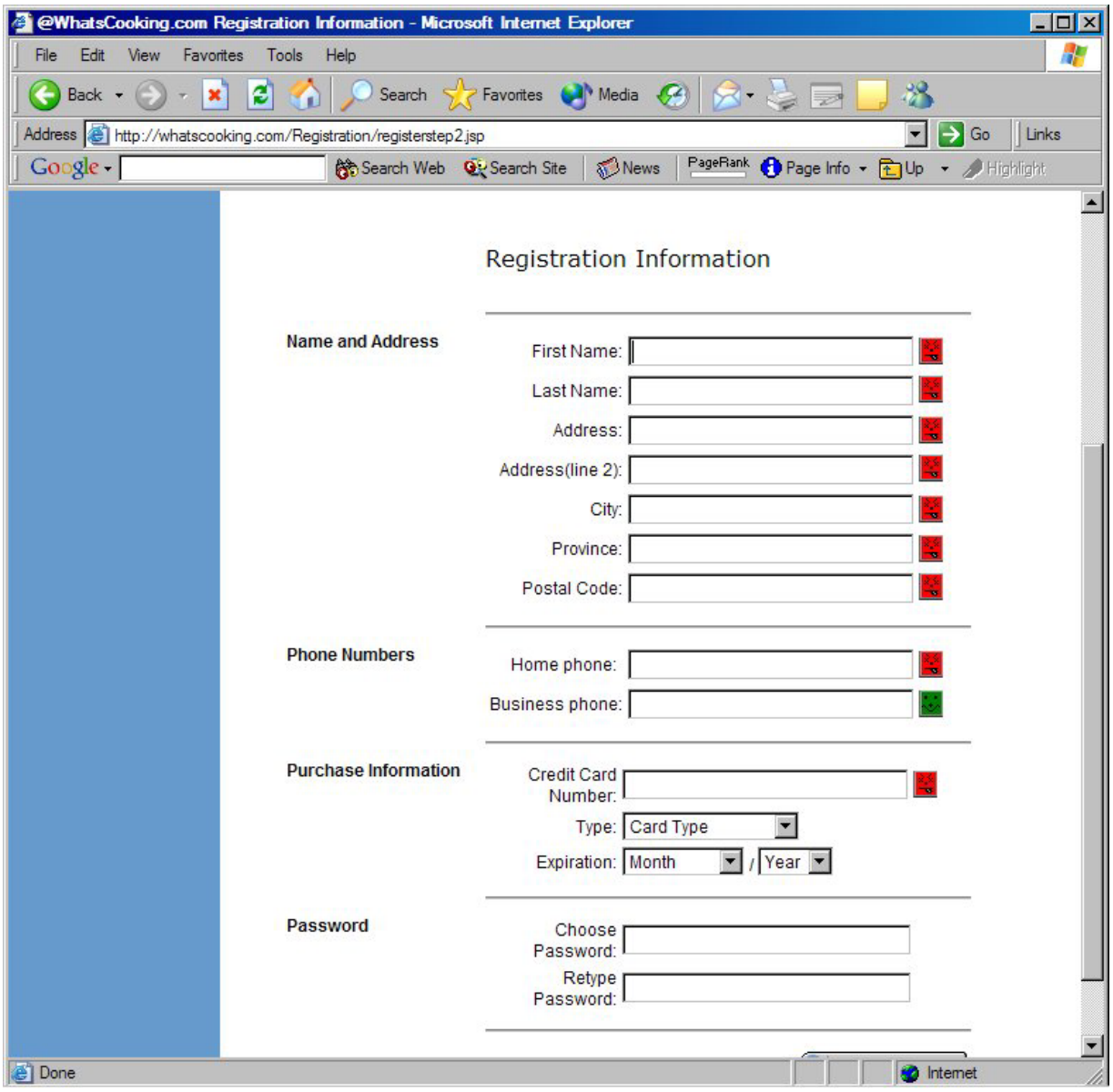
12. Do you have any additional observations about either of the privacy agents?

Alternate Presentation 1



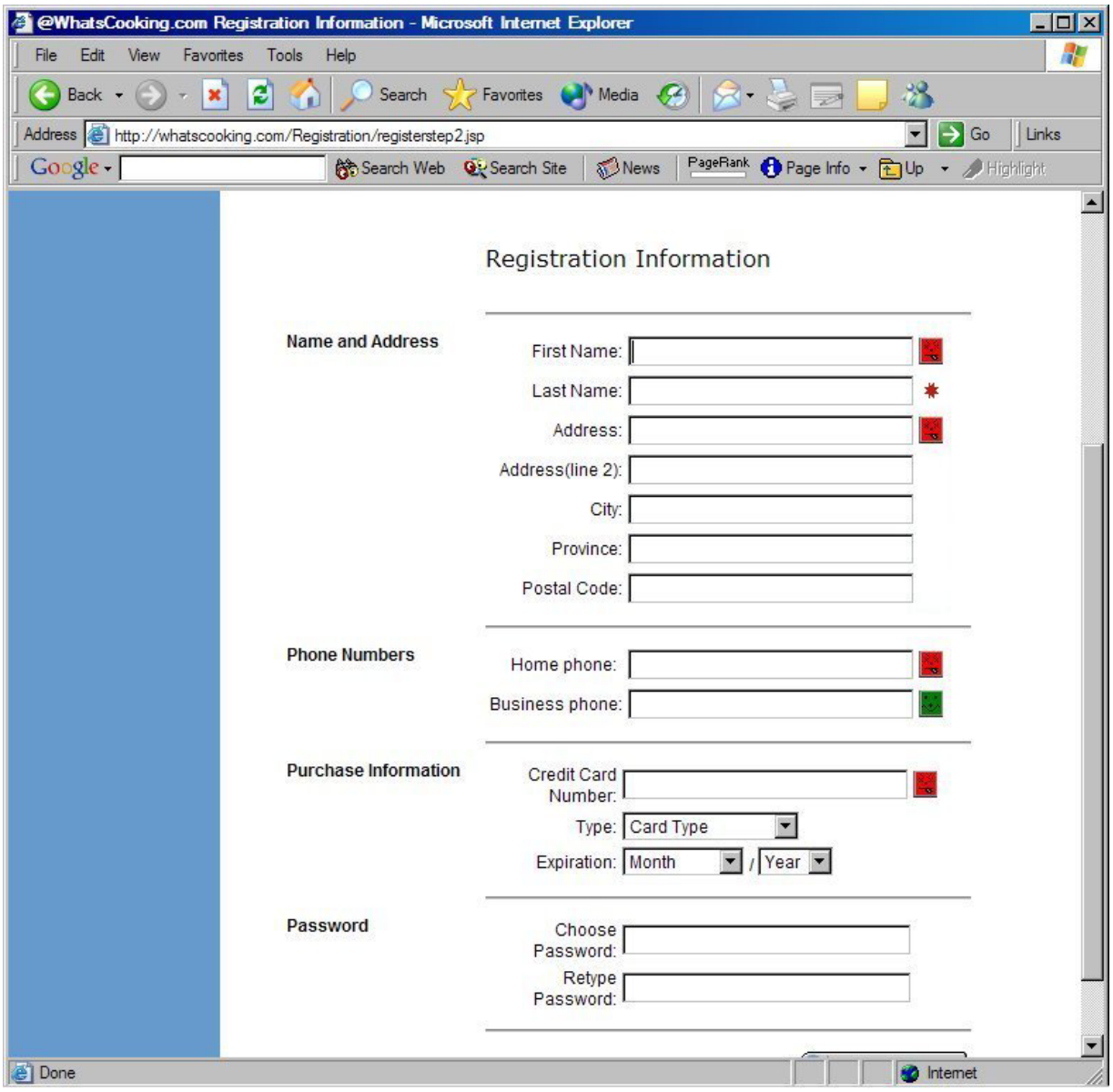
The participant is presented with a blank page and asked for their opinion of have conformance indication should be presented. Marking pens were provided to allow markup of the alternate presentations.

Alternate Presentation 2



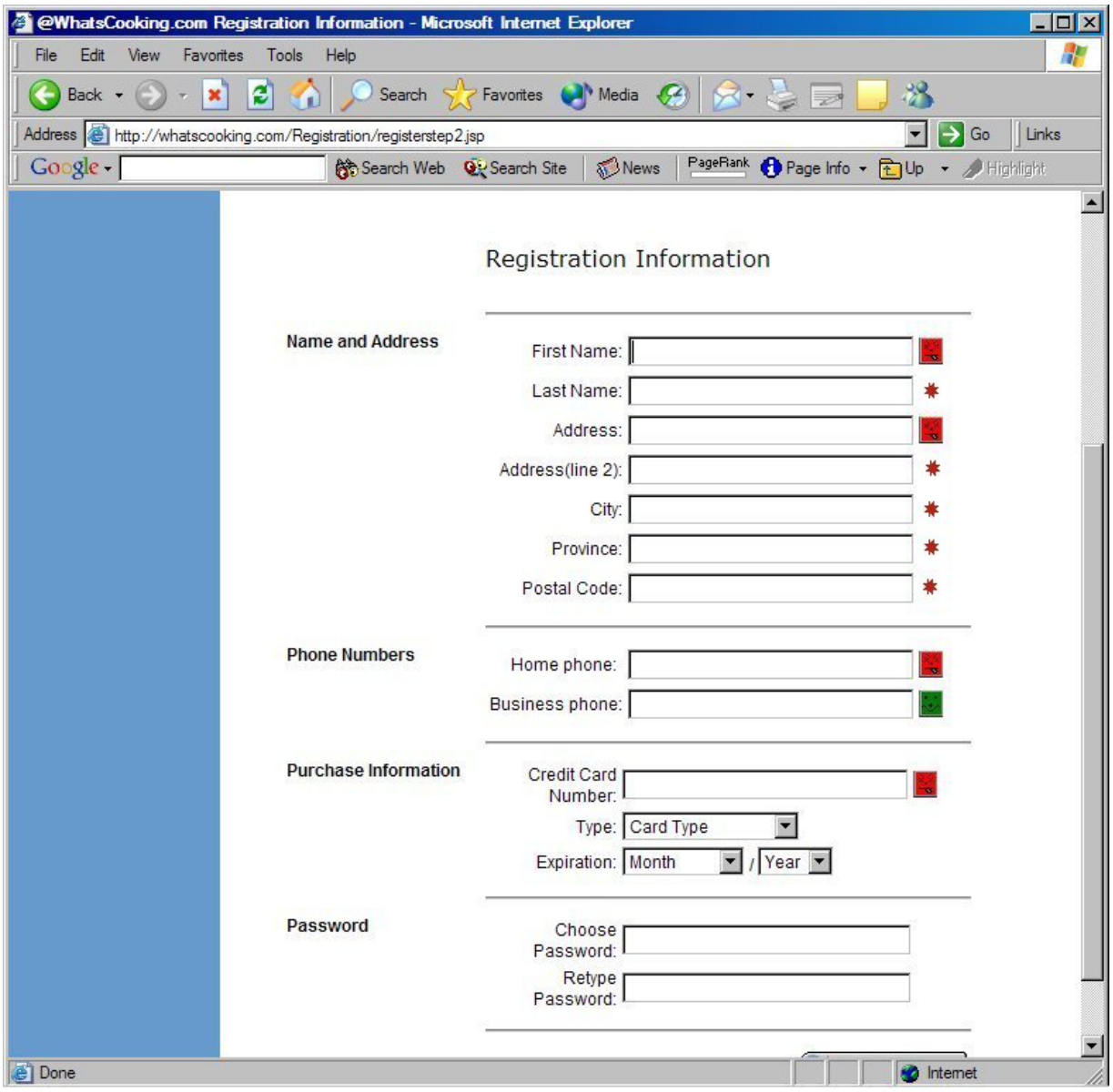
The participant is presented with a page displaying conformance in the same way as the evaluation. Issues of clutter, icons, and color are discussed.

Alternate Presentation 3



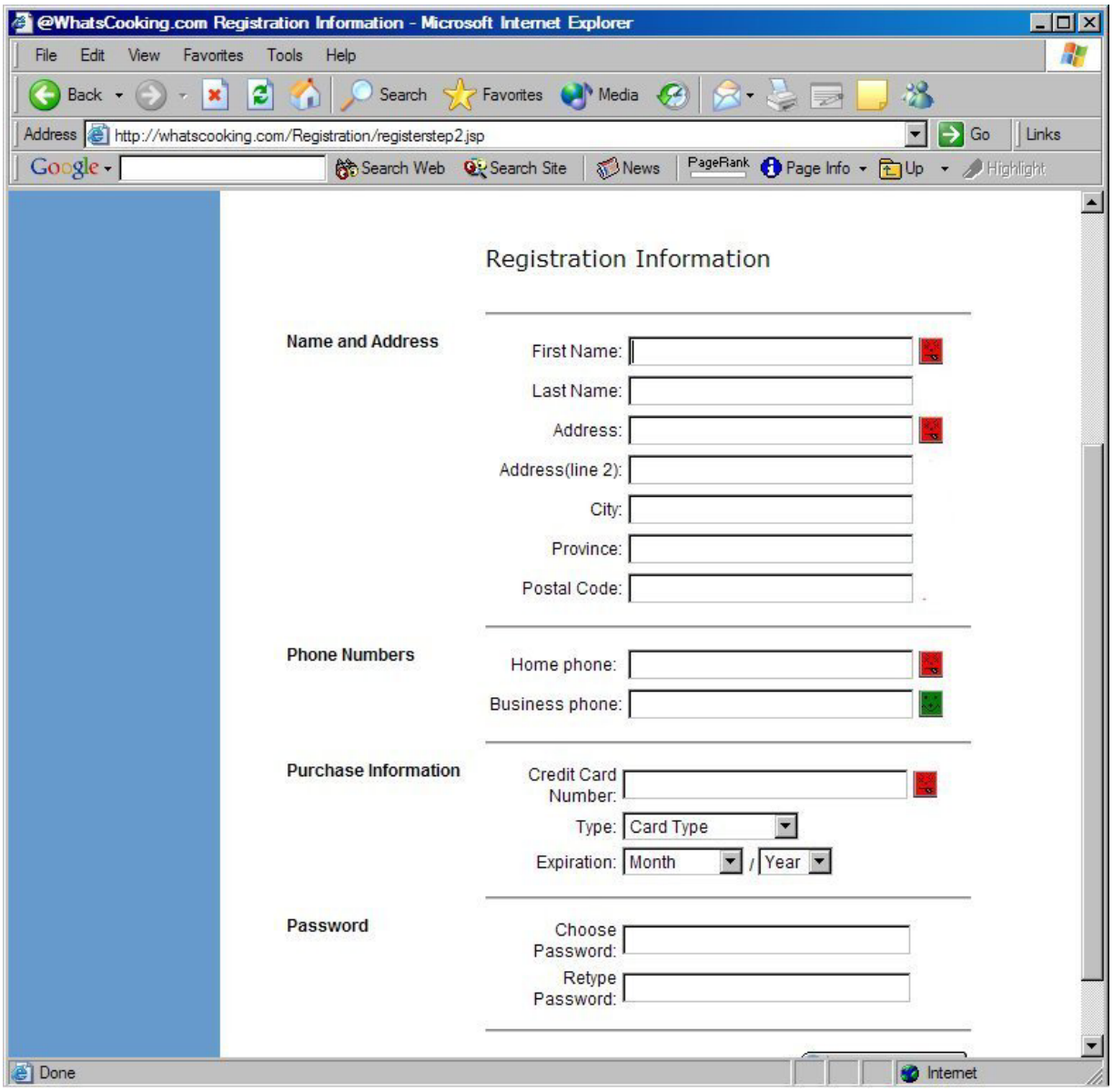
This presentation is used to discuss grouping. The issue of whether grouping should be implicit or explicit was discussed..

Alternate Presentation 4



This page was used to discuss explicit grouping and issues of clutter.

Alternate Presentation 5



This page was used to discuss implicit grouping of fields.

Appendix B: Privacy and Preference Policies

1. **WhatsCooking privacy policy**

2. **Individual privacy preference policy**

WhatsCooking privacy policy

```
<?xml version="1.0"?>
<POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1">
    <!-- Generated by WC P3P Policy Editor version Beta 1.11 built 6/4/02
11:23 AM -->

    <!-- Expiry information for this policy -->
    <EXPIRY max-age="604800"/>

    <!-- Custom data elements defined by this policy. -->
    <DATASCHEMA>

    <DATA-DEF name="WC" short-description="WC Data">
        <CATEGORIES><uniqueid/></CATEGORIES>
        <LONG-DESCRIPTION>Custom data elements defined by WC.</LONG-
DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.computerinfo" short-description="Computer
information">
        <CATEGORIES><computer/></CATEGORIES>
        <LONG-DESCRIPTION>Information about your WC computer system (for
example, a computer serial number for an WC personal computer you may have
registered online).</LONG-DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.registration" short-description="Registration
information">
        <CATEGORIES><uniqueid/></CATEGORIES>
    </DATA-DEF>
    <DATA-DEF name="WC.registration.userid" short-description="WC User ID">
        <CATEGORIES><uniqueid/></CATEGORIES>
        <LONG-DESCRIPTION>User ID created by registering for an WC
application or service.</LONG-DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.registration.password" short-description="WC
Password">
        <CATEGORIES><uniqueid/></CATEGORIES>
        <LONG-DESCRIPTION>Password created by the user when registering for
an WhatsCooking.com application or service.</LONG-DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.purchaseinfo" short-description="Purchase
information">
        <CATEGORIES><preference/><purchase/></CATEGORIES>
        <LONG-DESCRIPTION>Information about products being purchased and
payment method.</LONG-DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.purchaseinfo.payment" short-description="Payment
information">
        <CATEGORIES><purchase/></CATEGORIES>
        <LONG-DESCRIPTION>Information needed to process payment for an
online purchase, including credit card type, number, and expiry
information.
</LONG-DESCRIPTION>
    </DATA-DEF>
    <DATA-DEF name="WC.postings" short-description="Forum posting content">
        <CATEGORIES><content/></CATEGORIES>
```

```
        <LONG-DESCRIPTION>Content posted to public forums or discussion
groups.</LONG-DESCRIPTION>
    </DATA-DEF>
    </DATASCHEMA>

<POLICY
    name="WC-default"
    discuri="http://whatscooking.com/Registration/privacypolicy.jsp"
    opturi="http://whatscooking.com/Registration/privacypolicy.jsp"
    xml:lang="en">
    <!-- Description of the entity making this policy statement. -->
    <ENTITY>
    <DATA-GROUP>
<DATA ref="#business.name">WhatsCooking.com Corporation</DATA>
<DATA ref="#business.contact-info.online.email">prvcy@us.WC.com</DATA>
<DATA ref="#business.contact-
info.online.uri">http://whatscooking.com/</DATA>
<DATA ref="#business.contact-info.postal.organization">Customer Information
Privacy Practices</DATA>
<DATA ref="#business.contact-info.postal.street">WC Corporation
44 S. Broadway</DATA>
<DATA ref="#business.contact-info.postal.city">White Plains</DATA>
<DATA ref="#business.contact-info.postal.stateprov">NY</DATA>
<DATA ref="#business.contact-info.postal.postalcode">10601</DATA>
<DATA ref="#business.contact-info.postal.country">USA</DATA>
    </DATA-GROUP>
    </ENTITY>

    <!-- Disclosure -->
    <ACCESS><none/></ACCESS>


    <!-- Disputes -->
    <DISPUTES-GROUP>
        <DISPUTES resolution-type="service"
service="http://whatscooking.com/Register/privacypolicy.jsp" short-
description="Customer Service">
            <LONG-DESCRIPTION>Questions regarding this statement should
first be directed to WhatsCooking.com. You may contact us by e-mail at
prvcy@us.WC.com, or by postal mail at:
Customer Information Privacy Practices
WhatsCooking.com Corporation
44 S. Broadway
White Plains, NY USA
10601
</LONG-DESCRIPTION>
            <REMEDIES><correct/></REMEDIES>
        </DISPUTES>
        <DISPUTES resolution-type="independent"
service="http://www.truste.org/"
verification="http://www.truste.org/validate/331" short-
description="TRUSTe">
            <LONG-DESCRIPTION>TRUSTe - building a Web you can believe
in.</LONG-DESCRIPTION>
            <IMG src="http://whatscooking.com/trustmark.gif" alt="Reviewed
by TRUSTe - click to verify"/>
            <REMEDIES><correct/></REMEDIES>
```

```
        </DISPUTES>
    </DISPUTES-GROUP>

    <!-- TELEPHONE home -->
    <STATEMENT>
    <!-- Consequence -->
        <EXTENSION optional="yes">
            <GROUP-INFO
xmlns="http://www.software.WC.com/P3P/editor/extension-1.0.html"
name="Telephone Number"/>
        </EXTENSION>
    <CONSEQUENCE>
    Your telephone number will be used for telemarketing.
    </CONSEQUENCE>

    <!-- Use (purpose) -->
    <PURPOSE><telemarketing/></PURPOSE>

    <!-- Recipients -->
    <RECIPIENT><ours/><delivery/></RECIPIENT>

    <!-- Retention -->
    <RETENTION><indefinitely/></RETENTION>

    <!-- Base dataschema elements. -->
    <DATA-GROUP>
    <DATA ref="#user.home-info.telecom.telephone"/>
    <CATEGORIES>
       <telecom/>
    </CATEGORIES>
    </DATA-GROUP>
</STATEMENT>

    <STATEMENT>
    <!-- Consequence -->
        <EXTENSION optional="yes">
            <GROUP-INFO
xmlns="http://www.software.WC.com/P3P/editor/extension-1.0.html"
name="Telephone Number"/>
        </EXTENSION>
    <CONSEQUENCE>
    Your telephone number will be used only for administration of website.
    </CONSEQUENCE>

    <!-- Use (purpose) -->
    <PURPOSE><admin/></PURPOSE>

    <!-- Recipients -->
    <RECIPIENT><ours/></RECIPIENT>

    <!-- Retention -->
    <RETENTION><indefinitely/></RETENTION>

    <!-- Base dataschema elements. -->
    <DATA-GROUP>
    <DATA ref="#user.business-info.telecom.telephone"/>
    <DATA ref="#user.bdate"/>
```

```
    <DATA ref="#user.gender"/>
    <CATEGORIES>
       <telecom/>
    </CATEGORIES>
    </DATA-GROUP>
</STATEMENT>

    <!-- HOME ADDRESS -->
    <STATEMENT>
        <EXTENSION optional="yes">
            <GROUP-INFO
xmlns="http://www.software.WC.com/P3P/editor/extension-1.0.html" name="Home
Address"/>
        </EXTENSION>

    <!-- Consequence -->
    <CONSEQUENCE>Home Address</CONSEQUENCE>

    <!-- Use (purpose) -->
    <PURPOSE>
    <current/>
    </PURPOSE>

    <!-- Recipients -->
    <RECIPIENT><ours/><delivery/></RECIPIENT>

    <!-- Retention -->
    <RETENTION><indefinitely/></RETENTION>

    <!-- Base dataschema elements. -->
    <DATA-GROUP>
    <DATA ref="#user.home-info.postal.name.given"/>
    <DATA ref="#user.home-info.postal.name.family"/>
    <DATA ref="#user.home-info.postal.street"/>
    <DATA ref="#user.home-info.postal.city"/>
    <DATA ref="#user.home-info.postal.stateprov"/>
    <DATA ref="#user.home-info.postal.postalcode"/>
    <CATEGORIES>
       <telecom/>
    </CATEGORIES>
    </DATA-GROUP>
</STATEMENT>

    <!-- EMAIL ADDRESS -->
    <STATEMENT>
        <EXTENSION optional="yes">
            <GROUP-INFO
xmlns="http://www.software.WC.com/P3P/editor/extension-1.0.html"
name="Email Address"/>
        </EXTENSION>

    <!-- Consequence -->
    <CONSEQUENCE>Email Address</CONSEQUENCE>

    <!-- Use (purpose) -->
    <PURPOSE>
    <contact/>
```

```
    </PURPOSE>

    <!-- Recipients -->
    <RECIPIENT><ours/><delivery/></RECIPIENT>

    <!-- Retention -->
    <RETENTION><indefinitely/></RETENTION>

    <!-- Base dataschema elements. -->
    <DATA-GROUP>
    <DATA ref="#user.home-info.online.email"/>
    <CATEGORIES>
       <telecom/>
    </CATEGORIES>
    </DATA-GROUP>
</STATEMENT>



    <!-- Credit Card -->
    <STATEMENT>
        <EXTENSION optional="yes">
            <GROUP-INFO
xmlns="http://www.software.WC.com/P3P/editor/extension-1.0.html"
name="Purchase Information"/>
        </EXTENSION>

    <!-- Consequence -->
    <CONSEQUENCE>
     Your credit card information will be used for purchassing services
from us and related vendors, including delivery services</CONSEQUENCE>


    <PURPOSE>
    <individual-decision/>
    </PURPOSE>

    <!-- Recipients -->
    <RECIPIENT><other-recipient/></RECIPIENT>

    <!-- Retention -->
    <RETENTION><business-practices/></RETENTION>

    <!-- Base dataschema elements. -->
    <DATA-GROUP>
    <DATA ref="#user.jobtitle"/>
    <CATEGORIES>
       <purchase/>
    </CATEGORIES>
    </DATA-GROUP>
</STATEMENT>


<!-- End of policy -->
</POLICY>
</POLICIES>
```

Individual privacy preference policy

```
<?xml version="1.0"?>
  <appel:RULESET xmlns:appel="http://www.w3.org/2001/02/APPELv1"
xmlns:p3p="http://www.w3.org/2000/12/P3Pv1" crtdby="AT&amp;amp;T Privacy Bird"
crtdon="Fri May 02 13:55:17 2003
" >
    <!-- rule 1 -->
    <appel:RULE behavior="limited" description="This site will provide your
telephone number to third parties for the purpose of telemarketing" >
      <p3p:POLICY >
        <p3p:STATEMENT >
          <p3p:PURPOSE >
            <p3p:telemarketing />
          </p3p:PURPOSE>
        </p3p:STATEMENT>
      </p3p:POLICY>
    </appel:RULE>

    <!-- rule 2 -->
    <appel:RULE behavior="limited" description="This site will use your
email address to send you information about additional products and
services" >
      <p3p:POLICY >
        <p3p:STATEMENT >
          <p3p:PURPOSE >
            <p3p:contact />
          </p3p:PURPOSE>
        </p3p:STATEMENT>
      </p3p:POLICY>
    </appel:RULE>

    <!-- rule 3 -->
    <appel:RULE behavior="limited" description="This site will provide your
name and mail address to third parties for informational mailings" >
      <p3p:POLICY >
        <p3p:STATEMENT >
          <p3p:PURPOSE >
            <p3p:current />
          </p3p:PURPOSE>
        </p3p:STATEMENT>
      </p3p:POLICY>
    </appel:RULE>

    <!-- rule 4 -->
    <appel:RULE behavior="limited" description="This site will provide your
credit card to third parties who may not have the same business practices
as this site" >
      <p3p:POLICY >
        <p3p:STATEMENT >
          <p3p:PURPOSE >
            <p3p:individual-decision />
          </p3p:PURPOSE>
          <p3p:RECIPIENT >
            <p3p:other-recipient />
          </p3p:RECIPIENT>
          <p3p:RETENTION >
```

```
          <p3p:business-practices />
        </p3p:RETENTION>
      </p3p:STATEMENT>
    </p3p:POLICY>
  </appel:RULE>

</appel:RULESET>
```