AN ARCHITECTURE FOR IDENTITY MANAGEMENT

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Masters of Science

in the Department of Computer Science

University of Saskatchewan

Saskatoon

By

Brian Richardson

Keywords: privacy, identity management

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 or 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 this thesis.

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

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

ABSTRACT

Personalization of on-line content by on-line businesses can improve a user's experience and increase a business's chance of making a sale, but with stricter privacy legislation and Internet users' increasing concerns about privacy, businesses need to ensure they do not violate laws or frighten away potential customers. This thesis describes the design of the proposed Identity Management Architecture (IMA). The IMA system allows users to decide on a per business basis what personal information is provided, gives users greater access to their personal information held by on-line businesses, and does not rely on a trusted third-party for management of personal information.

In order to demonstrate the design and functionality of the IMA system a prototype implementation has been built. This implementation consists of the IMA client application and an example participating business to demonstrate the features of the IMA client. To evaluate the design of the IMA system it was compared to three high profile identity management systems: Microsoft .NET Passport, Liberty Alliance Project, and Microsoft Infocards. Through this evaluation each tool was compared based on the access to personal information provided to users and on what areas of privacy legislation compliance are improved for a business that participates.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Jim Greer for his guidance. His enthusiasm for research work and constant encouragement were a great help in completing this project. I would also like to thank my thesis committee members Dr. John Cooke and Dr. Gord McCalla whose comments and suggestions have been very helpful in the development of this thesis. I would also like to thank the students, staff and faculty of the Computer Science Department, and especially the members of the ARIES lab for their support.

TABLE OF CONTENTS

PERMISSION TO USE i

ABSTRACT ii

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS iv

LIST OF FIGURES vii

LIST OF ACRONYMS viii

Chapter 1 Introduction 1

1.1 Introduction 1

1.2 Motivation 2

1.3 Hypothesis and Research Objectives 3

1.4 Organization of Thesis 5

Chapter 2 Background 6

2.1 Privacy Legislation 6

2.1.1 Legislation Compliance 7

2.1.2 Privacy Policies 9

2.2 Privacy of Internet Users 10

2.3 Privacy Preferences 11

2.4 Single Sign-On Systems 12

2.5 Privacy Technologies 13

2.5.1 Platform for Privacy Preferences 14

2.5.2 TRUSTe 16

2.5.3 Privacy Critics 18

2.5.4 Privacy Incorporated Software Agent 18

2.5.5 Privacy Policy Compliance Systems 19

2.5.6 Enterprise Privacy Architecture 20

2.5.7 Enterprise Privacy Authorization Language 21
2.5.8 Security Assertion Markup Language 21
2.5.9 Federated Identity Management 22
2.5.10 PRIME and FIDIS 24
2.5.11 Liberty Alliance Project 26
2.5.12 Microsoft .NET Passport 27
2.5.13 Microsoft Infocards 30
2.5.14 Summary 31

Chapter 3 Identity Management Architecture 32

3.1 Overview 32
3.2 System Architecture 36
3.3 Passport vs. IMA 38
3.4 Managed Relationships 40

Chapter 4 Implementing the IMA System 43

4.1 IMA System Overview 43
4.2 IMA Toolbar 45
4.3 IMA Manager 49
4.3.1 Login 50
4.3.2 Create New Account 51
4.3.3 Identities 52
4.3.4 History 53
4.3.5 Viewing Profile Information 54
4.3.6 Identity Update Forwarding 55
4.4 IMA Web Service 55
4.4.1 Web Service Interface 56
4.4.2 Web Service Implementation 60
4.5 IMA Participating Business 61
4.5.1 Identifying IMA Users and Tracking of Profile Information 61
4.5.2 User's IMA Interactions with a Participating Business 63
4.5.3 Example Participating Business 65
4.6 XML Data 72
4.6.1 User Account 72
4.6.2 Report 74
4.7 Identity Management 76
4.7.1 Identities 76
4.7.2 Personas 77
4.7.3 Identities vs. Personas 78
4.8 Implementation 78
4.8.1 Platform Dependence 79

4.8.2 Setup and Installation 80
4.8.3 Scalability 84
4.8.4 Known Issues 85
4.9 Limitations 86

Chapter 5 Evaluation 89

5.1 Access to Personal Information 89
5.1.1 Information Storage, Management, and Access Comparison 90
5.2 Privacy Legislation Compliance 98
5.2.1 PIPEDA Compliance Comparison 99
5.2.2 Privacy Legislation Compliance in Other Countries 104
5.3 Summary 107

Chapter 6 Conclusions and Future Work 108

6.1 Contributions 109
6.2 Future Work 111
6.3 Conclusions 115

Appendix A Example IE Toolbar Setup 121

Appendix B IE Toolbar Libraries Setup 123

Appendix C User Account Schema 125

Appendix D Example User Account XML 127

Appendix E Report Schema 129

Appendix F Example Report XML 131

Appendix G Identity Schema 133

Appendix H Example Identity XML 134

LIST OF FIGURES

Figure page

Figure 3.1: IMA Identity-Business Associations 34
Figure 3.2: Identity Management Architecture 37
Figure 3.3: Personal Information Flow in .NET Passport 38
Figure 3.4: Personal Information Flow in the IMA System 39
Figure 3.5: Microsoft .NET Passport Relationships 41
Figure 3.6: Liberty Alliance Relationships 41
Figure 3.7: IMA System Relationships 42
Figure 4.1: Start the IMA Toolbar 45
Figure 4.2: IMA Toolbar – Login 46
Figure 4.3: IMA Toolbar and Manager – Login 47
Figure 4.4: Change identity 48
Figure 4.5: IMA Toolbar and Manager – Viewing web site history 48
Figure 4.6: Login Window 51
Figure 4.7: Create New Account 52
Figure 4.8: IMA Manager Identities Tab 52
Figure 4.9: IMA Manager History Tab 54
Figure 4.10: IMA Manager Site History and View Report Windows 55
Figure 4.11: IMA Web Service interface implemented for the example IMA participating business 60
Figure 4.12: Identity Tracking and History Information 62
Figure 4.13: Database Structure 63
Figure 4.14: Sequence of possible actions taken while visiting a participating site 65
Figure 4.15: Navigating to participating business, no identity associated 66
Figure 4.16: Index page of business, no association made yet 67
Figure 4.17: Selecting an existing identity from the list in the IMA Toolbar 68
Figure 4.18: Confirmation before creating the Business-Identity association 69
Figure 4.19: Viewing business products with an identity associated 70
Figure 4.20: Viewing information a business has associated with an identity 71
Figure 4.21: User Account Schema 73
Figure 4.22: Personal Information Report Schema 75
Figure 4.23: Internet Information Services “ima” and “shopping” virtual directories 82
Figure 4.24: IMA Visual Studio solution for client, web service, and example business 83
Figure 5.1: Summary of results from information access comparison of .Net Passport, Liberty Alliance, Infocards, and IMA 97
Figure 5.2: Summary of results from privacy legislation compliance comparison of .Net Passport, Liberty Alliance, Infocards, and IMA 104

LIST OF ACRONYMS

COM	Component Object Model
COT	Circle of Trust
DLL	Dynamic Linked Library
DPA	Data Protection Act
EPA	Enterprise Privacy Architecture
EPAL	Enterprise Privacy Authorization Language
FIDIS	Future of Identity in the Information Society
FIM	Federated Identity Management
FTC	Federal Trade Commission
GAC	Global Assembly Cache
IE	Internet Explorer
IIS	Internet Information Services
IMA	Identity Management Architecture
ISAT	Intelligent Software Agent Technologies
OASIS	Organization for the Advancement of Structured Information Standards
P3P	Platform for Privacy Preferences
PET	Privacy Enhanced Technologies

PIPEDA	Personal Information Protection and Electronic Documents Act
PISA	Privacy Incorporated Software Agent
PPCS	Privacy Policy Compliance System
PRIME	Privacy and Identity Management for Europe
SAML	Security Assertion Markup Language
SOAP	Simple Object Access Protocol
SSO	Single Sign-On
URI	Uniform Resource Identifier
URL	Universal Resource Locator
W3C	World Wide Web Consortium
WS-Security	Web Services Security
WSE	Web Services Enhancements
XML	Extensible Markup Language
XSD	XML Schema Definition

CHAPTER 1
INTRODUCTION

Personalization of on-line content by on-line businesses can improve a user's experience and increase a business's chance of making a sale, but with stricter privacy legislation and Internet users' increasing concerns about privacy, businesses need to ensure they do not violate laws or frighten away potential customers. This thesis describes the design of the Identity Management Architecture (IMA) [46]. The IMA system allows users to decide, on a per business basis what personal information is disclosed. It gives users greater control over their personal information held by on-line businesses, and does not rely on a trusted third-party for management of personal information.

1.1 Introduction

A Single Sign-On (SSO) system is one type of identity management system that allows a user to login to multiple systems and gain access to numerous resources all with a single username and password [60]. Systems like .NET Passport [35], Liberty Alliance [29], and Infocards [21] allow for the use of a single username and password at multiple web sites. The information for a user's account can follow the user from site to site without the user having to re-enter information at each site, as long as that site is a participating member of that SSO service. Both of these systems rely on some third party management of a user's personal information.

One common factor in most SSO systems is that they make the assumption that users always wish to present themselves online with the same identity with the same personal information [13]. However, for privacy reasons people may not want all of their activities online to be linked to a single identity. Many people who use the internet will present themselves with at least three different “Identities”: personal, work, and private [13]. Each of these identities may contain some unique personal information with some overlap, and the owner of these identities would normally not want the activities taken while under each identity to be linked together. For example if someone was considering finding a new job, he or she would not want the job hunt activities to be linked to the work identity for fear of the current employer finding out. It is for reasons like this that SSO systems need to realize the need for supporting multiple identities.

1.2 Motivation

In order to complete any commercial transaction on-line, people must provide personal information such as their name, address, phone number, email, credit card number, etc. On-line businesses often record more information than is actually needed to process a transaction. Some businesses monitor and record what types of products are bought or even the customer’s browsing patterns. This is done to form a detailed profile that will allow the business to target a customer with future advertising of products more closely related to individual interests. As a result, businesses may be inadvertently in violation of privacy law and customers may be unaware of the extent to which their personal data is being stored or used. Individuals sometimes try to counter such actions by supplying false or misleading data in an attempt to conceal their identities. Thus, the following three factors formed the basis of this research:

1. Legislation: Canada's Personal Information Protection and Electronic Documents Act (PIPEDA) and how businesses can readily comply with this law.
2. Personal Concerns: The increasing concerns of Internet users about what information on-line businesses record about them.
3. Tool Support: The lack of an available privacy tool that allows for management of multiple identities.

Even though there may be a wide range of privacy tools available, there is currently no tool for users to manage the information they have provided to an on-line business they regularly visit. This thesis presents the design of the Identity Management Architecture (IMA) that addresses several privacy issues. Privacy management is accomplished by providing Internet users with more control over their personal information while interacting with on-line businesses. This was achieved in the IMA system with four main features:

1. Providing for the creation and management of multiple discrete personal identities
2. Allowing users to restrict and manage the access that businesses have to identifying information
3. Providing users with the ability to request from a business what personal information is stored and
4. Providing businesses with a simple means of answering such requests.

1.3 Hypothesis and Research Objectives

Currently there is no personal information management system that does not require either a third-party or a business to directly pass a user's personal information to another business. It is hypothesized that a personal information management system can

be designed, which does not rely on a third-party, and can provide users with flexibility and control over the management of their personal information, while supporting business compliance with PIPEDA. It is also hypothesized that privacy can be supported through the use of multiple identities by allowing a user to partition personal information into multiple pieces, each referred to as an identity. This allows the user to choose on a per-business basis which identity to present. The use of multiple identities is a feature that is not offered by any personal information management system. The two key questions this research answers are:

1. Does the IMA System provide users with more flexibility and control over the management of their personal information than a third-party system does? This was shown by:
    a. Providing a breakdown of what types of personal information systems such as .NET Passport, Liberty Alliance, Infocards, and IMA allow users to access.
    b. Providing a detailed comparison of how personal information is accessed and updated in .NET Passport, Liberty Alliance, Infocards, and IMA.
2. Does the IMA System support business compliance with current privacy legislation?
    a. This was shown by providing a breakdown of PIPEDA showing what requirements of the act a business would or would not be in compliance with by using the IMA System as opposed to other systems.

1.4 Organization of Thesis

This thesis is organized as follows: Chapter 2 provides an overview of current privacy legislation and provides a review of privacy technologies available. It also provides some comparisons of these tools to the IMA System and shows how it will fill the gap that currently exists. Chapter 3 describes the architecture of the IMA System and compares this design to Passport and Liberty Alliance. Chapter 4 discusses the implementation of all four components of the IMA System, and shows the Extensible Markup Language (XML) format for personal information used by the IMA system. Chapter 5 provides an evaluation of the design of the IMA System and discusses its limitations. Chapter 6 discusses possible future directions for this work and draws conclusions.

CHAPTER 2
BACKGROUND

Management of personal information by identity management systems and the privacy concerns around the use of that information were a motivating factor in this work. The following is a discussion of the current state of Internet privacy in relation to the use and management of personal information. A brief overview of several privacy tools and research projects that involve personal information management is also provided.

2.1 Privacy Legislation

The Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) defines what personal information a business can track about a customer and how the business may use that information [43]. Information such as name, age, weight, height, income, purchases, spending habits, education, home address, and phone number are just a few examples of the types of information covered. This act only applies to information that is collected about an identifiable individual [44]. If an organization retains nothing that will allow the information to be linked to the identity of a user (e.g., use of aliases), this act does not apply. PIPEDA also places the added responsibility on businesses to respond to any requests by customers to view what personal information the business has about them, and to update or remove that information based on customers' wishes. The act defines ten responsibilities that an organization must follow in the course of commercial activity [44]:

- Accountability
- Identifying purposes
- Consent
- Limiting collection
- Limiting use, disclosure, and retention
- Accuracy
- Safeguards
- Openness
- Individual access
- Challenging compliance

PIPEDA has been put into action in three phases: the first phase in January 2001 for businesses that are federally regulated (e.g., banks) [44], the second phase in January 2002 for health care information, and the third phase in January 2004 which extends the laws defined in the act to apply to the handling of personal information gathered in any commercial transaction [44]. It is the third phase of this act that worries many businesses. The act puts restrictions on how information is obtained (i.e., consent must be obtained before information is gathered), and on what types of information businesses are allowed to record.

2.1.1 Legislation Compliance

With the introduction of PIPEDA more restrictions have been placed on how businesses obtain consent when asking a user for personal information. Consent can no longer be obtained by just a general statement asking for a user's personal information, but must be more specific [50]. Businesses can no longer use difficult to understand legal statements when asking for consent, but must now [11]:

- Ask permission from the customer before recording any information

- Explain for what purpose the information will be used and who it will be disclosed to
- Take proper measures to ensure the information is stored in a secure manner

Under PIPEDA a business must disclose to users, upon request, what personal information the business has about them. In order to be able to process these requests a business must ensure it has appointed someone within the organization as a privacy officer who is in charge of fielding requests from customers and also ensuring that the business's handling of personal information is in compliance with PIPEDA [50]. Some companies will need to hire additional employees to deal strictly with customer requests for personal information.

There are several risks that companies face by not complying with privacy legislation, such as having their reputation tarnished in the press for misuse of customer personal information, losing customer trust due to security breaches resulting in stolen customer personal information, being sanctioned for violating privacy legislation, or even being sued by angry customers. For example, Air Canada received bad press when it provided information gathered within the Aeroplan program to business partners without obtaining the consent of the Aeroplan members [50]. In the complaint filed with the Privacy Commissioner by an Aeroplan member, it was stated that Air Canada did not obtain members' consent before sharing members' personal information with any external sources. Air Canada was found to be in violation of PIPEDA [42].

2.1.2 Privacy Policies

A privacy policy is a document describing what an online business does with personal information it collects about people who visit the web site. A privacy policy is responsible for stating [19]:

- What personal information is gathered about users
- How that information is used
- Who that information is shared with
- What options users may have over how their information is used

The current privacy policy from Chapters.Indigo.ca, for instance, provides a description of how a user may modify the information that was originally provided when creating a new account when making a subsequent purchase [12]. This is legally sufficient, but is not necessarily convenient. For example, if a customer's address changes, the address must be updated at each on-line business before making another purchase. This is a process that could be simplified by having a system that could manage a user's personal information across multiple web sites.

Under PIPEDA, businesses must delete any personal information they have stored about a customer upon the customer's request. In the privacy policy from Chapters.Indigo.ca instructions are provided for requesting deletion of personal information [12]. These instructions provide an email address to its privacy officer and require a detailed written description of what information the customer wants to see or has privacy concerns about [12]. This section of the privacy policy states that Chapters.Indigo.ca complies with PIPEDA and helps to maintain confidence with users by showing that there is an established course of action to take to have information removed.

The privacy policy from Amazon.ca explains the types of information that are tracked about customers such as information the customer provides (i.e., address, email, phone number, etc.) and the products purchased and browsed [3]. What makes the privacy policy posted by Amazon.ca different from the policy at Chapters.Indigo.ca is that the privacy policy from Amazon.ca explains the on-line tools it provides for customers to modify account information and even edit or remove the history of products browsed. Customers may access these options in their on-line account. This kind of on-line access to account information is useful, but if there were some way to have common information updated as necessary across several frequently-used on-line businesses, users might find this much more convenient.

2.2 Privacy of Internet Users

The United States Federal Trade Commission (FTC) recently conducted a survey of Internet users and found that 57% of consumers who shop on-line believe that if a business has a privacy policy posted on its web site then it would not gather information from, or share personal information with, anyone else [52]. What is concerning about this finding is that privacy policies from commercial sites generally state that a large amount of information is being kept about customers, while customers believe that the presence of a privacy policy means that a business will not gather or share a user's personal information [17]. This misconception may lead to a user disclosing personal information without fully understanding how a business will use it.

Out of this same study by the FTC, 40% of consumers who use the Internet stated that they had no knowledge of how commercial web sites gather information about their interests and make use of this information [52]. If Internet users are unaware that

information is being recorded about them, they are likely not taking necessary steps to protect their privacy while on-line.

Spiekermann et al. describe a survey where a group of volunteers indicate their privacy preferences in a simulated Internet shopping experience [48]. The results of this experiment showed that many participants, even those who were very concerned about on-line privacy, did not follow their stated privacy preferences while actually shopping. Instead, they were likely to reveal more personal information such as home address, clothing sizes, preferences in brand, cost, usage, etc. than they had intended at the outset.

2.3 Privacy Preferences

On-line businesses already understand the benefits of gathering information about visitors to their web sites since it allows them to get a better understanding of visitors' interests, and thus improve the chances of turning visitors into customers. This has resulted in growing concern among Internet users about what information a business is tracking about them. For example, a survey conducted on Internet users showed several concerns that would cause a visitor to a web site to be less likely to divulge personal information [2].

Rated as highly important by 69% of respondents:

- That the information be used in an identifiable way
- Type of information collected
- The purpose for which this information is being collected

Rated as very important by 62% of respondents:

- Whether a site is run by a trusted company or organization
- Whether a site will allow visitors to access information recorded about them

- Whether a site will remove someone from a mailing list upon request

For on-line businesses, getting Internet users to visit their web site and to be willing to disclose personal information is crucial to survival. After all, if a visitor to a business's web site does not trust how that business handles personal information and is therefore not willing to disclose personal information, then a transaction such as purchasing a product or subscribing to a service will be much less likely to occur. Some businesses seem almost to rely on the lack of awareness by Internet users, assuming that since most are not aware of how businesses gather and use personal information, then they don't need to respect the privacy preferences of users. This is a dangerous attitude to have towards users' privacy. It is also illegal in Canada.

2.4 Single Sign-On Systems

A Single Sign-On (SSO) system is one type of identity management system that allows a user to login to a system and gain access to numerous resources all with the use of a single username and password [60]. Systems like .NET Passport and Liberty Alliance allow for the use of a single username and password at multiple web sites, where the information for that one account can follow the user from site to site without the user having to re-enter information at each site visited, as long as that site is a participating member of that SSO service.

One common factor in most SSO systems is the assumption that users always wish to present themselves online as the same identity with the same personal information [13]. The fact that people may not want all of their activities online to be linked to the same identity is a point that many SSO systems (e.g., .NET Passport) seem to give no consideration to. Often when users are using the internet, they will do so for

different reasons and may sometimes wish to present themselves to others with different identities.

Most people who use the internet will present themselves with at least three different “Identities”: personal, work, and private [13]. Each of these identities contains some different personal information, and obviously the owner of these identities would not want the activities taken while under each identity to be linked together. For example if someone was considering finding a new job, he or she would not want the job hunt activities to be linked to a work identity for fear of the current employer finding out. It is for reasons like this that SSO systems need to realize the need for supporting multiple identities.

2.5 Privacy Technologies

By law businesses that gather personal information about users must state what information they gather and what they use it for in a legal document called a Privacy Policy. Unfortunately, these documents are often long and difficult to understand. If a visitor to a web site does not read a business’s privacy policy before disclosing personal information, then the visitor may not be making an informed decision. In this section several technologies that have been designed to help protect users’ privacy are discussed and compared.

Internet users want to protect their personal information, yet many people do not make any effort to do so. Over the last few years more than a dozen products to protect users’ privacy (e.g., firewalls, antivirus programs, spyware detection programs, etc.) have been released and have failed to be embraced by internet users [20]. Some of the main reasons for these tools failing to gain acceptance are due to the cost of purchasing, the

inconvenience of setting privacy settings, and the effort required to learn to use these tools [20]. Stephanie Perrin, the president of Digital Discretion Inc., believes that many Internet users are not aware of tools that can help protect their privacy while online and simply do not realize the value of their personal information to people who may wish to steal it:

> "Leaving personal information around ought to be thought of as leaving a bucket of cash, because it's saleable" [20].

2.5.1 Platform for Privacy Preferences

The Platform for Privacy Preferences (P3P) standard was developed by the World Wide Web Consortium (W3C) to allow Internet users to define privacy preferences and have these preferences compared to a business's privacy policy automatically when visiting the web site [55]. Since privacy policies are often long and difficult to understand, P3P offers a simple interface that allows Internet users to define their preferences once, then have these preferences automatically checked at each participating P3P web site they visit.

P3P works by having each participating business define its privacy policy in an XML document, following the P3P standard. The business then posts this XML document on its web site, making it accessible to users with P3P client applications [55]. A user who has a web browser that supports P3P will automatically have the business's P3P XML document retrieved when initially visiting the business's web site. This policy is automatically compared to the user's personal privacy preferences. If there is a discrepancy, the user is alerted to make an informed decision about whether or not to disclose extra personal information to the business [55].

P3P relies on informing visitors to a web site about the contents of a business's privacy policy. The problem with relying on this as the sole form of privacy protection for Internet users is that they may not always change their actions based on the privacy differences shown between their preferences and the business's privacy policy [48]. Although a user may be informed about the ways a business uses personal information, the user may not alter his or her behaviour in terms of browsing activity or when asked questions while shopping, especially when interested in a product or service [48].

With P3P, if there are differences between personal privacy preferences and a business's privacy policy, the user must either accept these differences or not use the business's web site altogether. With P3P there is no ability for someone to agree to certain uses of personal information and not others, it is an all or nothing agreement. The IMA system improves this situation by allowing people to choose, on a per business basis, which identity the business will be able to use, and as a result restrict what personal information the business will be able to access.

Even if a web site states in its P3P privacy proposal that a user's information will only be used for certain purposes, this does not necessarily mean it is true [30]. There is no enforcement of these policies to ensure a business's use of information follows what is stated in its P3P policy. TRUSTe is one organization that monitors businesses privacy practices [51]. TRUSTe acts as a privacy check for consumers by doing some monitoring of businesses' privacy policies to ensure they match the businesses' use of users' personal information. If a business is found to be using information for uses not stated in its privacy policy TRUSTe may remove its trustmark from the business. This is not enough to really straighten out companies that abuse access to users' personal

information and make users feel secure that their P3P privacy preferences match with a business where the user's preferences will actually be respected [30]

The IMA system allows users to define the personal information stored in each identity they create, to decide which identity to allow a business to have access to, and to view what information they have provided. The IMA system is by no means meant to be a replacement for P3P, but rather provides a service which allows users to tailor personal information to a business depending on the level with which they trust that business. For example, if a business's privacy policy stated uses of personal information that the user is not comfortable with, the IMA system will allow the user to select which identity will be used at that web site and how much information will be disclosed.

2.5.2 TRUSTe

TRUSTe is a non-profit organization that provides privacy seals to web sites that follow a set of privacy specifications [14]. When users see a TRUSTe seal on a business's web site it provides assurances to consumers that personal information is handled according to the businesses privacy policy. TRUSTe does not set privacy policies, but rather ensures that a business's privacy policy accurately states what personal information the business gathers and how it is handled.

The idea for TRUSTe was originally started back in 1996, in part by Esther Dyson. Dyson was the driving force behind the establishment of a non-profit organization that would ensure that companies comply with what is stated in their own privacy policies [8]. However Dyson has not been satisfied with the role TRUSTe has ended up taking compared to what was originally envisioned:

> "Rather than revoking seals left and right, TRUSTe officials often seemed to be covering for their clients -– explaining, in one case, that a Real Networks media

> player which reported users' video selections back to Real headquarters in Seattle was "outside of the scope of TRUSTe's current privacy seal."" [8].

Since TRUSTe was formed, it has investigated hundreds of privacy complaints made by customers of very well known businesses such as Microsoft, Yahoo, and Real Networks, but still to this day has never once revoked the TRUSTe certification from a business [31].

In the last few years TRUSTe has come under fire for problems that have occurred. One such problem was Yahoo's decision to change the way it would use personal information, while including information collected previously under a different privacy policy bearing the TRUSTe seal [8]. Similar incidents have also occurred with Real Networks and eBay. This type of misuse of the TRUSTe seal, and the organization's unwillingness or inability to do something about it has brought into question the validity of this TRUSTe seal program. Seth Ross from PC Guardian had this to say about the TRUSTe seal:

> "A trustmark does more harm than good by creating an illusion of privacy where none exists. A meaningless logo may induce people to make information disclosures that they would otherwise avoid" [8].

TRUSTe and the IMA system have a similar goal, which is to help businesses and users build trusted relationships, however they offer very different types of services. TRUSTe ensures that a business discloses how information is handled, while the IMA system discloses the information itself and gives users and businesses more control over managing personal information. Even when a business has a simple, easy to read privacy policy, users may not take the time to read the policy.

2.5.3 Privacy Critics

A privacy critic is an agent that provides the user with warnings of potential privacy risks to the actions the user is attempting to take [1]. It is then up to the user to decide either to take the suggestion, or ignore it. Unlike other privacy technologies that block certain information, a privacy critic takes no action on behalf of the user, but rather just provides the user with warnings without stopping the user from making a bad decision [1]. If a user has decided not to listen to the privacy critic and releases information, there is no way for the user to get that information back. The IMA system does not provide any of the features of a privacy critic, but it does allow a user to view information given to a participating business and modify or remove that information at any time.

2.5.4 Privacy Incorporated Software Agent

Software agents that have a level of intelligence and are used to perform tasks for a user automatically on his or her behalf are called Intelligent Software Agent Technologies (ISAT) [7]. For an ISAT to perform its required tasks it must know personal information about the user to fulfill each request. It is the profile created about the user by the ISAT that creates a potential privacy risk, whether the ISAT is compromised by an outside source, or whether it accidentally releases personal information [7].

With the Privacy Incorporated Software Agent (PISA), Dr. John J. Borking, the vice-president of the Dutch Data Protection Authority, proposed to address areas of potential privacy risks by building a Privacy Enhancing Technology Agent [7]. The main goals of the PISA project were: to demonstrate that Privacy Enhanced Technologies (PET) that rely on agents can be a sound solution for Internet users to be able to protect

their privacy, to work with businesses to develop new privacy services, and to propose standards for agents that protect users' privacy [28].

The PISA agent allows other agents to perform tasks requested by the user yet ensures only the minimal personal information required for a task is provided and nothing more. The PISA agent would be placed either between the user and other agents to prevent those agents from possessing any more information than necessary, or it would be placed between the agents and all outside systems [7].

The PISA agent has an advantage over P3P in that it gives more control over personal information to the user. P3P puts more responsibility on the user to guard his/her own privacy. The downside of the PISA agent is that implementing the agent so that it would not accidentally divulge information to the wrong source and at the same time be compliant with privacy legislation is difficult.

2.5.5 Privacy Policy Compliance Systems

The Privacy Policy Compliance System (PPCS) proposed by G. Yee and L. Korba is based on the idea of allowing web users to check how a business will use his or her personal information [61]. What this system allows a user to do is check if what the business intends to do with the user's information is acceptable to the user before the user discloses his or her personal information.

In the PPCS system both the consumer (i.e., customer) and provider (i.e., business) declare a privacy policy. The privacy policy supplied by the consumer states his or her preferences about uses of his or her personal information that are acceptable for a provider to do. The privacy policy provided by the provider declares how it intends to use the consumer's personal information. In addition to the privacy policy, in the PPCS

system the consumer is also responsible for supplying private data to the system, while the provider is responsible for supplying a list of the data required, an explanation of how the data will be used, and a statement of how the personal information will be managed [61].

The proposed PPCS system allows consumers to see if their privacy preferences match those of the provider before the consumers' personal information is entered. At the same time the system allows businesses to be in compliance with privacy legislation by allowing them to obtain consent by disclosing to a consumer how they plan to use personal information. The IMA system is different from PPCS in that it does not involve privacy policies that explain users' preferences, but instead deals strictly with the personal information itself and allows users to manage their personal information across multiple businesses without the use of a central source (i.e., third party service).

2.5.6 Enterprise Privacy Architecture

The Enterprise Privacy Architecture (EPA) offered by IBM provides a system for businesses to evaluate their existing uses of customer information and locate potential privacy risks that may exist [23]. By evaluating a business's business processes in terms of uses of personal information, EPA allows a business to see how a customer's information is being used and shows how to place protection on that information to ensure any inappropriate disclosure does not occur [23]. With EPA it allows businesses to manage their privacy practices by integrating into their business rules the handling of [49]:

- External privacy rules
    - such as ensuring compliance with privacy legislation

- Privacy preferences of users
    - respecting a user's wishes when handling the user's person information

While the services offered by EPA allow a business to evaluate how information is managed in its business processes, the IMA system looks at how information is managed between a customer and multiple businesses.

2.5.7 Enterprise Privacy Authorization Language

The Enterprise Privacy Authorization Language (EPAL) developed by IBM is a formal language designed for the defining of privacy policies that decide how personal information is handled based on what authorization it is given [5]. EPAL provides businesses with a language to automate the handling of privacy policies between applications and businesses [54].

With EPAL businesses are given the ability to electronically and automatically enforce privacy policy restrictions on the use of personal information and to ensure that a user's personal data is used only under the restrictions stated in the privacy policy. The use of EPAL also allows businesses to give their users greater assurances that what is stated in the privacy policy is actually being enforced [54]. EPAL looks at enforcing privacy policy restrictions in an automated way once information has been provided to a business, while the IMA system looks at offering a way for a user to manage the personal information provided to businesses.

2.5.8 Security Assertion Markup Language

The Security Assertion Markup Language (SAML) is being designed as a protocol for exchanging security information through systems that communicate via web services [15]. This protocol is being developed by the Organization for the Advancement

of Structured Information Standards (OASIS) [27]. The idea behind SAML is to improve interoperability of identity management in a secure way across systems, not just within the same company, but also between companies [15].

The way SSO works with SAML is through an authentication server. A user accesses the authentication server storing personal information using his or her username and password. The authentication server provides authentication for the user. The user then makes a request which is retrieved from a web service. The authentication server opens a connection with the requested service and creates a Uniform Resource Identifier (URI) to the resource and sends it to the user. Finally the user receives the URI and connects directly to the desired resource [27]. This style of SSO once again relies on a central source for authentication of an identity. This authentication service is responsible for verifying the identity of the user and providing access to other resources within the same network. The IMA System on the other hand does not rely on a business or central service to provide authentication, nor does it rely on the use of one company to provide authentication for a user at another company.

2.5.9 Federated Identity Management

Many identity management systems that companies have developed and integrated into their systems deal with helping companies with managing user information within the organization, such as managing usernames, passwords, and access rights for employees and customers across multiple systems. Federated Identity Management (FIM) looks at company-to-company identity management and exchanging of information. FIM is a model for making the sharing of identity information seamless across company barriers [40]. What this means is that FIM is a model for a SSO system

that allows a user to have a single username and password that can be used as an identity at multiple companies and all be linked to the same set of personal information [40].

FIM works as a network of companies. Each company trusts the others participating in FIM to authenticate users and set access rights. When a user visits another company, that company trusts the company that provides the identity and deems it to be correct, granting that user the use of the account and access rights as supplied by the trusted company [10]. An example would be a mechanic at a car dealership being able to access his parts supplier's system based on the identity authenticated and provided by the dealership the mechanic works for.

What makes FIM so challenging to implement is that there must be a standard developed and agreed upon by companies for the exchange of identities. In order for FIM to take off and become widely used it will need to tackle issues of interoperability, such as getting businesses to agree on a standard for communication and security, and it will also need to deal with issues of trust [6]. The push for FIM is being led by the United States Government's E-Authentication project and the government is using SAML to help deal with interoperability issues [6]. Right now FIM is currently being used in a few small test programs managed by E-Authentication.

One issue is that a company is relying on another company to provide authentication for a user. This brings in issues of company-to-company trust and each company needing to know that the other's security and user identity management is handled carefully [10]. Also, like Liberty Alliance, FIM relies on the formation of trusted networks between organizations for identity sharing. In order to get companies to participate they must trust the other companies in the FIM network. If a company does

not want to participate with another company, then it will not be a part of the trusted network and users will not be able to use the same account at both companies.

2.5.10 PRIME and FIDIS

There is currently extensive research work going on in the area of identity management by the Privacy and Identity Management for Europe (PRIME) [41] and The Future of Identity in the Information Society (FIDIS) [18] projects. Work by both of these projects has produced prototype identity management systems which take a similar approach to the handling of identities. Both projects allow users to switch identities (roles) based on which identity they wish to present to a business. While both the PRIME [58] and FIDIS [59] projects are large in scope and are exploring in great depth many privacy and identity management issues, the IMA system really only attempts to address a small portion of the identity management problem.

The FIDIS project presents the prototype iManager which is the Identity Manager for Partial Identities [59]. Each partial identity contains a subset of the user's information that is applicable to the information needed for the user's current role, such as an identity that contains a credit card number and mailing address used when the user is shopping online. This approach is similar to how identities are handled in the IMA system. Basically each identity in IMA, like partial identities, is identified and authenticated by a unique key that allows the user to be authenticated by a business for all future visits using the same identity. This allows a business to be able to track all repeat visits and associate information about the user with that identity, allowing the user to build up a relationship, regardless of whether or not the user has even provided his or her identifying personal information such as name, address, email, etc. A similar approach to the iManager's

partial identities was followed in the design of the IMA system's multiple identity user account. The IMA system splits the user account up into identities which each contain a different subset of the user's personal information such as anonymous, personal, work, school, etc. Each identity allows the user to only provide that set of information contained within the given identity, all other identities and information contained in the account are not disclosed to a business.

The PRIME project presents the prototype IDM system [58]. The IDM system is a much more comprehensive solution than the IMA system; however there are still some areas of IDM that the IMA system touches on. The IDM system improves privacy by allowing the user to remain anonymous, even during a transaction, assuming there is a trusted third party that in the case of a problem (e.g., legal matter) the identity can be recovered. However in the IMA system no attempt to preserve anonymity like this was made. In the IMA system if a user decides to complete a transaction with a business, it is up to the user to decide whether or not he or she wishes to disclose an identity containing the required information. No anonymity is preserved in this type of transaction.

The primary goal of the IMA system was to build a SSO system that did not require third party storage or knowledge of a user's information. As an additional feature the IMA system would also allow a user to create and manage more than one identity from within a single user account where all identities could be accessed by a single username and password. These initial requirements were what the design of the IMA system had to achieve. Rather than comparing the design of the IMA system to the ongoing work in identity management taking place in PRIME and FIDIS, the IMA system was looked at more as an improvement upon existing SSO systems such as .NET

Passport and Liberty Alliance. Both of these existing systems have well defined architectures which allow for a more detailed comparison to be made.

2.5.11 Liberty Alliance Project

The Liberty Alliance Project (www.projectliberty.org) started in 2001 by Sun Microsystems to create a SSO authentication service [16]. The goal of the Liberty Alliance project was to create a system based on an open standard, as opposed to Microsoft's .NET Passport which is a proprietary platform. The Liberty Alliance project has gained support from other well known organizations in the last few years and now has more than 30 companies (i.e., Computer Associates, Hewlett Packard, Novell, etc.) involved in the development of the specification [16].

There is a misconception that Liberty Alliance is a similar service to .NET Passport. This is not the case. While .NET Passport is a SSO service implemented by Microsoft and used by online businesses, Liberty Alliance Project is the development of a specification that businesses can implement if they wish to participate [47]. This specification allows businesses to form identity sharing relationships between each other and each implementation of the specification allows for this communication.

Liberty Alliance is based on the idea of allowing users to connect multiple sets of personal information, that exist across several on-line businesses, into one easy to manage identity. This allows for the convenience of a SSO service, as well as easier management of personal information across multiple businesses [9]. The Liberty Alliance architecture allows an Internet user to store his or her personal information with a trusted business. When the user needs to access a service provided by another business, which is

part of the same group of associated businesses, the user's chosen trusted business provides authentication of the user, as well as the user's identity information [9].

What makes this system architecture unique is that, rather than relying on a trusted third party system such as .NET Passport to provide a user's identity to each business the user accesses, it allows the user to have a business he or she trusts store and pass identity information from one business to another, which is part of the same group of associated businesses [9]. A group of associated businesses who have an agreement to share user identities and act as a SSO service is referred to as a Circle of Trust (COT). In a COT one business may act as the identity provider for a user and provide that identity to other businesses in the COT the user accesses [47]. One downside to this design is that identity management across multiple businesses is restricted to the set of businesses that have formed associations with each other. If a business is part of another group of associated businesses, identity information passing between these businesses is not available.

2.5.12 Microsoft .NET Passport

The Passport system (www.passport.net) was founded in 1999 by Microsoft. This system was designed to provide a SSO service that would allow Internet users to have one account for access to all Passport participating web sites [37]. The Passport system handles authentication of users by having the sign-on page on each participating web site authenticate the user by contacting the Passport system [37].

A user's Passport account information is never provided to a business unless the user gives consent. Consent is obtained when a user signs into a business's web site using Passport [37]. The Passport system stores only the information a user provides on sign-

up. This is the only information that is provided to a business when a user accesses the business's web site. All information gathered about the user while at a business's web site is stored only by the business. It is left up to the business to determine what it will do with the information gathered (i.e., share the user's personal information with other businesses). In the Passport participation agreement, there are no restrictions placed on a business about how the business may use personal information provided by Passport or the user [37]. However, each business must state what information is collected and how it is used in its privacy policy.

A user may update his or her email address, phone number, or postal code, through the Passport home web site and have that information propagated to Passport-approved businesses. This does not give the user access, however, to the personal information a business participating in the Passport system may have gathered. Instead, it is up to the user to read the privacy policy of each partner business's web site to determine what information the business might collect and maintain. If the user had access to the information a business has stored (i.e., profile), this would give the user much more control over personal information.

One of the main features of the Passport system is the SSO service. Although this is convenient, it does not allow for any sort of management of multiple identities [13]. Passport does not allow the creation of multiple identities (i.e., more than one set of personal information) to be associated with a single account to allow a user to choose on a per business basis, what personal information a business receives. Although it is true that this could be accomplished by creating multiple Passport accounts, this defeats the purpose of a SSO service since this approach would require a user to manage several

Passport accounts, to remember multiple usernames and passwords, and to remember which account had been used at each business.

In early 2005, Microsoft announced that it was no longer going to pursue Passport as a solution for businesses to allow customers to manage their credit card as well as other personal information as they move from business to business online [33]. Companies such as eBay and Monster.com, two of the biggest non Microsoft companies to be using .NET Passport, who were initial supporters of the system, have dropped it in the last year. Microsoft pitched .NET Passport as a service that would allow thousands of web sites to be accessed by a single username and password. However, this system failed to gain a following from both the businesses and Internet users that Microsoft had envisioned. Users were concerned about Microsoft storing their personal information and with the few large companies involved with Passport backing out (e.g., eBay) this has forced Microsoft to scale back the plans it originally had for Passport [16].

Users have been skeptical about storing their personal information in the Passport system. Anytime a person logs into another Passport participating site, that site is immediately able to access the information in that person's Passport account [26]. Unfortunately for Microsoft its systems often become the target of attacks by hackers which has caused security holes in .NET Passport to be made public. A report by AT&T labs exposed several security flaws with Passport. All it would take to compromise user accounts would be a site that has a fake Passport login. This would allow usernames and passwords to be obtained providing access to all information in Passport about that user [26]. Microsoft also had to discontinue its use of the .NET Passport Wallet, which is a service that stores a user's credit card information, after it was discovered that all it

would take to steal a person's passport account and gain access to the Passport Wallet would be to get a user to open a Hotmail email [32]. Issues like this have raised continued concerns about security and privacy in the Passport system.

2.5.13 Microsoft Infocards

In response to .NET Passport's failure to be embraced by both Internet users and businesses, Microsoft has taken the lessons learned from that system's failure as the basis for the new MS Infocards system which is currently in development [21]. This new identity management technology will be built into the upcoming Microsoft Windows Operating System Longhorn which is scheduled for release some time in 2006. The two main design features of .NET Passport were its reliance on a trusted third party for storage and management of personal information, and its assumption that people only wish to present themselves online as a single identity. With MS Infocards information is stored and managed locally on the user's computer, while it also allows for multiple identities such as anonymous, work, government, etc [21]. The MS Infocards user interface to be built into Longhorn will allow users to manage identities, create multiple identities, define for what uses each identity may be used, and also restrict what parts of each identity will be shared. Currently no detailed plans for Infocards have been released from Microsoft.

With Infocards it appears that Microsoft is attempting to take a different direction from .NET Passport, by removing the need for third party management of personal information, by allowing for the use of multiple identities, and by defining the usage and disclosure of information for each identity. IMA follows this type of system design. However IMA looks more at the ability to assign individual identities to participating

businesses online and also adds the ability for users to track what information they have provided so they can be aware to whom they have given information and what information the business has associated with them. Since little information has been released by Microsoft about Infocards at the time of the writing of this thesis, a detailed comparison of IMA and Infocards cannot be completed.

2.5.14 Summary

Although there are a number of privacy technologies, either currently in use on the Internet or in the concept stage, .NET Passport and Liberty Alliance are the two that most closely resemble the proposed IMA System. All three of these systems are SSO services that manage a user's personal information across multiple online businesses. The proposed IMA System has a design that takes a different approach to how information is managed than .NET Passport or Liberty Alliance. The design of the IMA system is described in Chapter 3.

CHAPTER 3

IDENTITY MANAGEMENT ARCHITECTURE

In order to understand what makes the IMA system different from existing SSO systems, it is important to understand the overall architecture of the system. The design of both the client side and business side of the system are discussed in the following sections.

3.1 Overview

The IMA system is designed around two main components: the IMA Manager, which is the client application, and the IMA Web Service, which is the web service deployed by participating on-line businesses. A web service is an application that is made available to other applications over HTTP that provides cross platform interoperability by allowing requests and responses to be made using XML messages. Each business that wishes to participate in the IMA approach must follow the standard defined for the IMA Web Service, must implement this service, and must deploy this service on the business's web site. Through this service, all interactions with the IMA Manager client application are handled. Each user who wishes to benefit from the IMA system installs the IMA Manager application on his or her computer. This application ties into the user's web browser (e.g., Internet Explorer). Each person using the IMA Manager can create one or more identities, and then when visiting the web sites of participating businesses, can choose which identity to associate with each business. For future visits, this identity will be used unless changed by the user.

The IMA system also allows users to have access to any information the business has recorded about them. Someone using the IMA Manager can accomplish this by making a request for information from a business. The IMA Web Service that a participating business has deployed handles this request automatically. Upon receiving the request, the IMA Web Service then queries the business's database for information related to the person. This information is returned to the requesting IMA Manager for the user to review.

One of the key features of the IMA system, which is not offered by other personal information management tools, is the ability to create and manage multiple identities from within a single user account. The use of multiple identities does more than just restrict what information a business will see, but also allows people to interact with a business for more than one purpose. For example, if someone shops at an on-line computer parts store, sometimes for work purposes and other times for personal purposes, he or she may want to create a "Work" identity and a "Personal" identity. Creating separate identities allows someone to more easily manage these two separate relationships with a business. This may be beneficial to a business, especially one that personalizes web site content based on the interests of the user. If someone is using his or her "Personal" identity, the business may use the browsed products and recently purchased products to make suggestions to the user about other products that may be of interest. If this same person visits the business's web site at another time using the "Work" identity, the business will be better able to tailor content towards the interests associated with this identity.

Through the IMA Manager a user can create one or more identities. Upon visiting a business's web site for the first time, whichever identity the user has set as the default will be used. The user may change the identity associated with a business at this time. Each time a user visits a business's web site, whichever identity the user has set as the currently active identity will automatically be used by the participating business to identify the returning customer. Figure 3.1 shows an example of a set of identities a user may choose to create and the associations of identities to businesses made by the user. As shown in the figure, a user may associate one or more identities with a business.

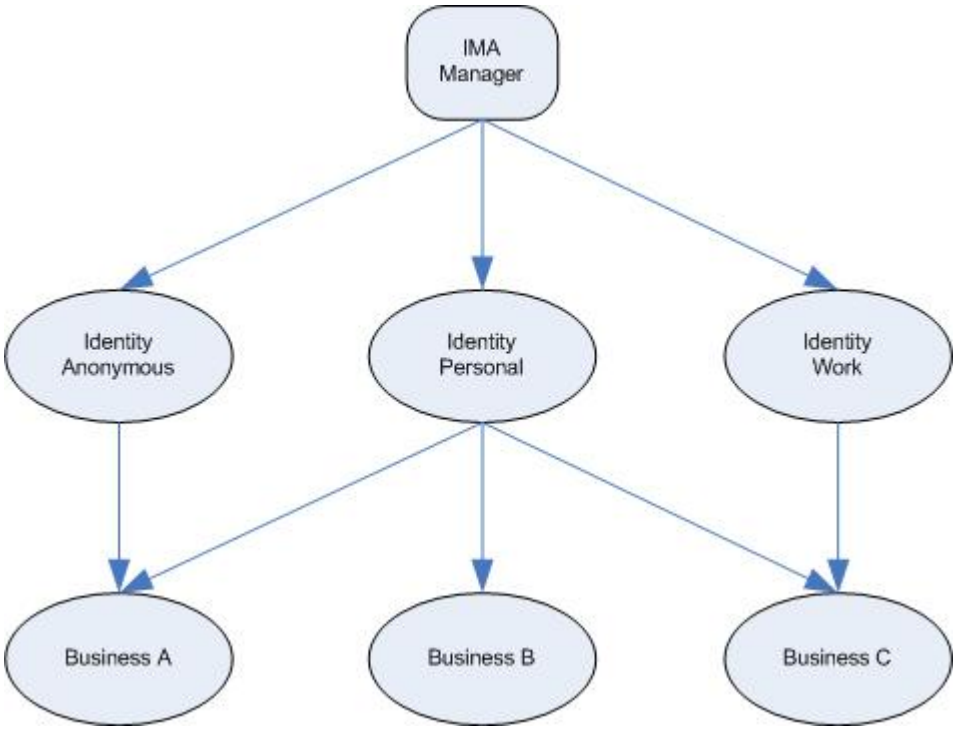

Figure 3.1: IMA Identity-Business Associations

Each identity that an IMA Manager stores is separate and distinct from other identities. A business is not given any information that would allow it to determine whether several identities belong to the same person. Each identity the user creates may be completely independent of the user's account information. The information contained

in an identity is the only information ever transmitted to a business by IMA. The user account, which contains the user's username, password, identities, and associations, is never provided to a business. Each identity is assigned a randomly created unique key, which is used by the business to identify that specific identity. Since no other account information is ever provided, the user can decide freely which identity to use at any given time. If a user was to provide more than one identity to a business, it would not be that difficult for a business to infer that two identities belong to the same person, for example if both identities were provided from the same IP address or the same credit card was used. While this makes it possible for a business to determine if two identities belong to the same person, this is still not legal under PIPEDA since consent has not been obtained from the user and the assumption made may be incorrect.

One of the problems for users, especially when using different sets of information at multiple businesses, is keeping track of what information has been used at each business. This can lead to a person having to remember multiple aliases, which include email addresses, usernames, passwords, etc. What makes the IMA system useful in this respect is that it not only manages the information for each identity, but also the associations of each of these identities to a set of businesses. So for each subsequent visit to a business's web site the user does not have to keep track of what information has been provided, but can rely on the IMA Manager to do this. Through the use of multiple identities users are given the ability to pick and choose what information a business will have access to about them when they browse or make a transaction. For example, if a user has contacted a business for work related purposes, the user only needs to provide an identity which contains his or her work information, but does not have to provide any

information in any other identity, which may include information such as a home phone number or personal email address.

3.2 System Architecture

Personal information management systems such as .NET Passport or Liberty Alliance rely on either a third-party or another business to store and transfer someone's personal information. One goal of IMA is to avoid any use of a third party system and to not require businesses to communicate with each other for the purpose of providing a customer's information. The IMA system has two main components:

1. IMA Manager (Client): An application that attaches to the user's web browser and handles the management of all user identities and web browsing history.
2. IMA Web Service (Business): A web service that each participating business provides to allow users of the IMA Manager to send and receive identity information.

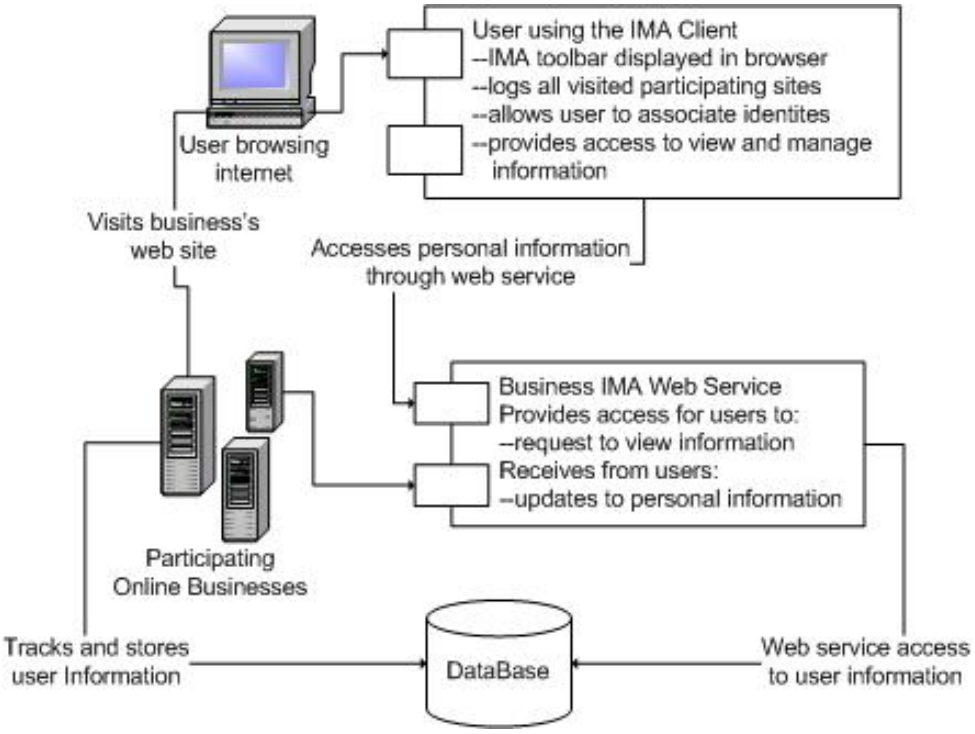

Figure 3.2: Identity Management Architecture

The IMA Manager allows a user to contact a business's IMA Web Service to make a request to see what information the business currently has stored in the user's profile. A user may correct or remove information. If the information to be changed is contained in an identity, a user may modify the identity information stored in the IMA Manager and the application will automatically forward updates to all businesses associated with this identity. The user may associate another identity with a business at any time; this will be used for future visits to that business.

As shown in Figure 3.2, the IMA Manager runs on the user's local web browser. The Manager receives from the web browser the Universal Resource Locator (URL) of each site the user visits while on-line. It then checks to see if the business is participating in IMA by attempting to contact the IMA web service that all participating sites are required to make available. If this business is not participating, then this is shown in the

IMA Manager's display. However, if the web service is available, this URL is stored and the service is contacted. The first communication with a business is the transmission of the user's preferred on-line identity. From this point on, each time the user returns to this web site the user will be identified by this identity, allowing the business to associate information with this identity, such as the products browsed, to determine the user's interests.

3.3 Passport vs. IMA

What makes the IMA System's architecture different from Passport is the lack of a third party participant. The result is several differences in the way personal information is handled in the absence of a third party. In Passport, a user's personal information is stored in two locations: in the Passport system and at each participating on-line business. Figure 3.3 shows the typical flow of personal information in Passport.

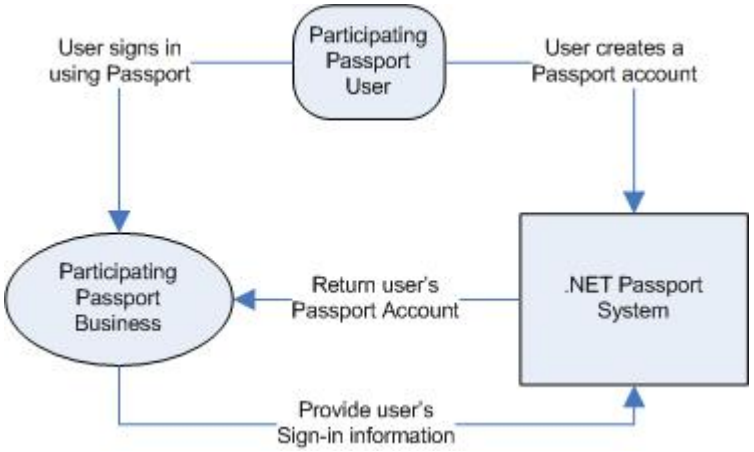

Figure 3.3: Personal Information Flow in .NET Passport

In the IMA system a user's personal information is stored in two places: on the user's local computer and at each on-line business the user has provided with an identity. This is different from Passport's approach since the IMA system does not require a user to provide personal information to a third party in order to participate. The user does not

need to worry about the security of his or her information on the third party system. Figure 3.4 shows the typical flow of personal information in the IMA system.

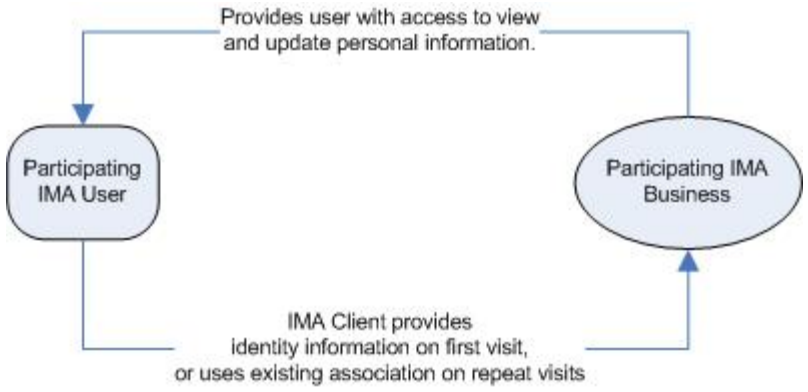

Figure 3.4: Personal Information Flow in the IMA System

Only that information contained in a single identity is sent to a business. This is unlike Passport, which uses only one set of personal information for a user's account and provides it to each partner business. If a user wishes to have more than one identity, multiple Passport accounts must be created which defeats the purpose of a SSO service. The IMA system is designed to provide management of multiple identities on behalf of the user.

One of the key features of the IMA system is that it provides users with the option to directly request what information a business has about them. The only information the user has access to view and update with Passport is the personal information that was entered when the account was created. This does not provide users with any access to additional information a business has about them and it does not provide any benefit to businesses in terms of compliance with information disclosure requirements placed on businesses by PIPEDA. In the IMA application the user has the ability to look at the list of businesses the user has provided one or more identities to, then select a business and an identity that has been used at that business and submit a request for information to that

business. The response from the business is displayed in the IMA application and allows the user to view what information has been provided as well as what profile information has been associated with the identity such as types of products viewed and purchased.

3.4 Managed Relationships

With most personal information management systems (SSO services) it is necessary for either the user or service provider (or both) to establish additional relationships (i.e., with a third party system or another business) in order to be able to participate. The main goal of the IMA system is to provide the same types of services provided by traditional personal information management systems, but without requiring additional relationships to be established or maintained.

In order for users to participate in the .NET Passport system they must create an account with .NET Passport and provide all personal information they wish to be used in that account. This is the first new relationship that must be established outside of the traditional customer-business relationship. The second new relationship required is between .NET Passport and each business that wishes to participate (see Figure 3.5). These two new relationships require both users and businesses to be willing to participate in a personal information management system that involves a trusted third party.

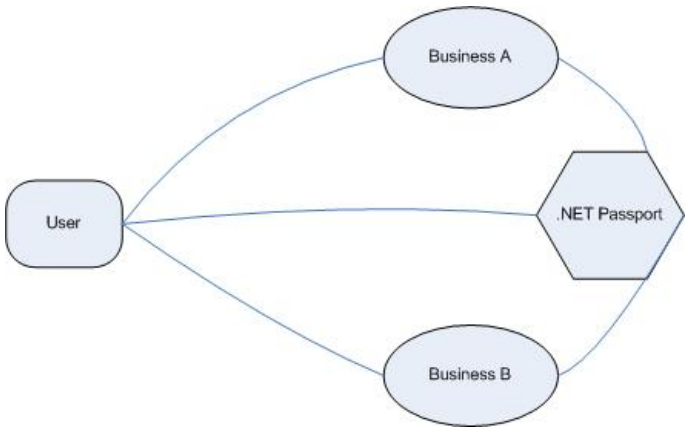

Figure 3.5: Microsoft .NET Passport Relationships

The Liberty Alliance system allows users to select a business they trust to store their personal information. In order for this identity to be used at another business two conditions must be true: the business the user is visiting must also be participating in the Liberty Alliance system and both businesses must have established an identity sharing relationship with each other (see Figure 3.6). If both are true, then the user may use the same account at another business and have personal information transferred from the trusted business.

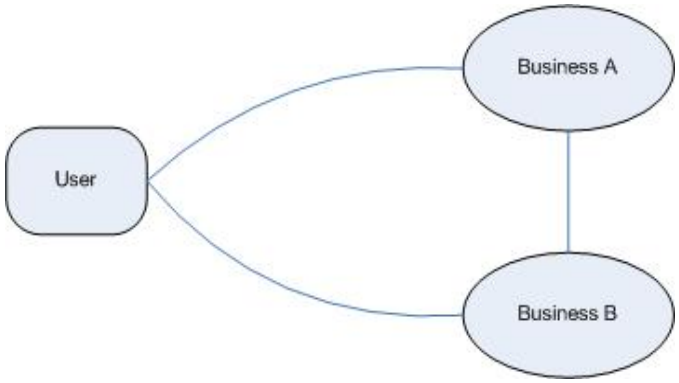

Figure 3.6: Liberty Alliance Relationships

The IMA system relies on the existing customer-business relationship. If both the user and business are participating in the IMA system, personal information the user has in an account can be transferred to each business by the user (see Figure 3.7). The IMA system does not require a user or business to have to form additional relationships with either a third party system or another business.

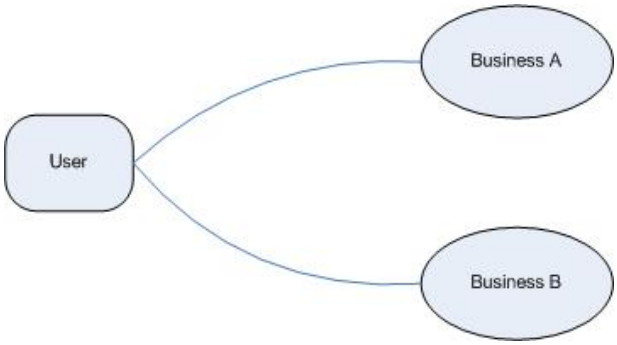

Figure 3.7: IMA System Relationships

By allowing for the use of multiple identities and the management of user account and identity information client side, the IMA system takes an unconventional approach to the design of an SSO system. To provide a better understanding of how the IMA system is used a prototype implementation has been built. This implementation of the IMA system is explained in detail in Chapter 4.

CHAPTER 4

IMPLEMENTING THE IMA SYSTEM

In order to demonstrate the design of this system, a prototype implementation has been built. The implementation includes a functional prototype of the client application as well as the server-side implementation for a sample participating business. The design of each of the four components that make up the IMA System, the user interfaces, and the XML data used are all discussed in detail in the following sections.

4.1 IMA System Overview

Each component of the system has been built using the .NET framework [34]. Since the IMA system is based on web services, any platform or programming language could have been chosen. The .NET framework was chosen for this implementation simply for the libraries it provides that make implementing a toolbar for IE relatively straight forward. There were four components that needed to be implemented for this thesis:

- IMA Toolbar
- IMA Manager
- IMA Web Service
- Example participating business web site

The IMA Toolbar is a .NET application that integrates into the user's web browser and provides the user with information on the participation of a web site being visited and the identity currently being used. This toolbar allows the IMA Manager to be

automatically started when the user opens a web browser. This application is a class library that is registered as a toolbar with Internet Explorer (IE) and controls the IMA Manager. The IMA Toolbar created for this project was based on an example .NET Toolbar obtained from The Code Project [62].

The IMA Manager is a .NET application that runs in the background on the user's system. It is displayed as a taskbar icon, unless the user wishes to view the browsing history or modify identities. This application is a standard windows application that displays a window when the taskbar icon is clicked.

The IMA Web Service is a .NET web service that allows the IMA Manager application to communicate with a participating business. Identity information is transferred to and from the IMA Web Service as an XML document. All other information recorded by the business that has been associated with the current identity can be retrieved in an XML document that the IMA Manager can display to the user and the user can make changes if necessary. How a business actually implements this system and how it ties into the business's database is the decision of the business. All that is required is that the URL of the service and the methods offered match the ones required by the IMA Manager.

To demonstrate the IMA system a small prototype E-Commerce web site was built using ASP .NET. This site acted as a participating business providing the IMA Web Service that allowed the IMA Manager to be used to get a feel for how the IMA system would work. This allowed for a better understanding of the IMA system design and also played a role in several design changes. The design of the IMA Manager and the IMA Web Service are described in detail in the following sections.

4.2 IMA Toolbar

The IMA Toolbar provides a compact and easy to use display that appears in the user's web browser. This toolbar offers several basic features of the IMA Manager:

1. Allows a user to login to the IMA Manager
2. Notifies a user if a web site is participating in the IMA system
3. Allows the user to switch identities at any time
4. Launches the IMA Manager application whenever the user wishes to view his or her participating web site history, modify existing identity-to-business associations, or change settings

The IMA Toolbar is registered as a standard toolbar in IE. Once installed, it can be opened by selecting it from the "Toolbars" menu in IE (see Figure 4.1).

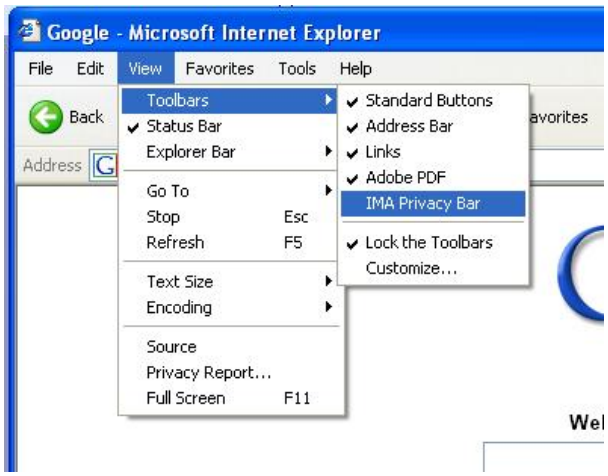

Figure 4.1: Start the IMA Toolbar

The IMA Toolbar once opened appears in the upper right-hand corner of the browser window. The green or red circle, shown on the left side of the toolbar (see Figure 4.2), shows the user if the web site currently being visited is a participating member of the IMA system. The icons used are shown below:

(Green) Participating in the IMA System

(Red) Not participating in the IMA System

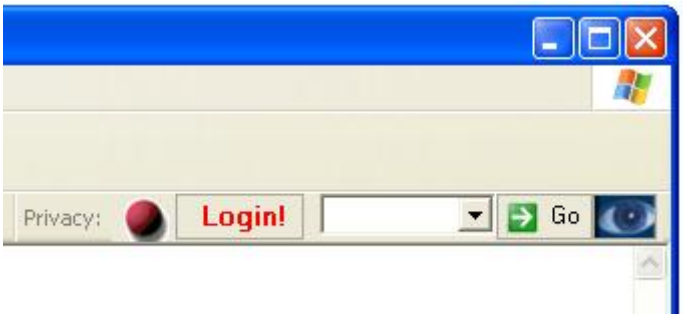

Figure 4.2: IMA Toolbar – Login

If the user clicks the button labeled “Login!” the IMA Manager will be opened displaying the login window (see Figure 4.3). Once a user has logged into the IMA system, it is only necessary to keep the toolbar open to be aware of basic information such as whether each web site visited is participating in the IMA system, what identity is currently being used, and be able to launch the IMA Manager application at any time if needed by clicking the “Eye” button (see Figure 4.3).

Figure 4.3: IMA Toolbar and Manager – Login

If at any time while browsing the Internet users wish to change the current identity they are using, they may do so by using the drop down list (see Figure 4.4). This list makes available all identities the user has created which may be associated with any site the user is visiting. Once the identity is selected the “Go” button is used to send the selected identities to the participating site currently being visited.

Figure 4.4: Change identity

If a user wishes to modify the information contained in an identity, or change the identity associated with a business, these features may be accessed from the IMA Manager. The History tab in the IMA Manager shows the list of participating web sites the user has associated identities with (see Figure 4.5).

Figure 4.5: IMA Toolbar and Manager – Viewing web site history

The IMA Toolbar acts as a browser interface for the IMA Manager; however, this toolbar provides very little functionality. The majority of the work in the client-side part of the IMA system is handled by the IMA Manager application. The features of the IMA Manager are discussed in further detail in the following section.

It is possible that the use of green and red icons, the icons used to show if a business is participating in the IMA system, may pose a problem for users who are red-green colour blind. The green and red icons chosen for the IMA Toolbar could easily be replaced by anything else. For example, these icons could be changed to a check mark (e.g.,✔) for a participating site and a letter x (e.g.,✗) for a site that is not participating if this change is required.

4.3 IMA Manager

The IMA Manager is the client application used by any user who wishes to participate in the IMA system. This application has three main responsibilities:

1. Receiving from the user's web browser the URL of each web site visited
2. Storage of all identity information and the web browsing history associated with each identity
3. Sending and receiving of personal information through the IMA Web Service provided by each participating business

The goal here was to keep the application simple to understand and use. Since the IMA Manager allows a user to create multiple identities and choose what identity a business will see, the user must also set a default identity. If a user visits a business's web site that is not listed in the user's history, the IMA Manager will automatically associate

the new business with the default identity. The user may change the identity associated with a business at any time.

All web browsing history is stored only on the local user's computer and is never disclosed to any business. This gives each user complete control when browsing and allows for the removal or deletion of entries. The only identity information disclosed to an on-line business visited for the first time is the default one, so information stored in any other identity is sent to a business only if the user explicitly permits the business to know him or her by another identity. The anonymous identity is the preferred default identity since it prevents any accidental disclosure of information while the user is online and also allows a user to hide his or her identity information until he or she decides to associate an identity with a business.

4.3.1 Login

The IMA Manager application will always be running in the background on a user's machine unless the user logs out of the IMA Manager, or reboots the machine. When this happens, the next time the user wishes to turn the IMA Manager back on, there will be a prompt to enter the username and password (see Figure 4.6).

Figure 4.6: Login Window

Also provided on this screen is an option to create a new account. This option is available on the login window in case it is either the first time the application has been run after installing it, or there is another person who also uses the same computer and wishes to create his or her own account.

4.3.2 Create New Account

For a new user of the IMA System, all that is required to create a new account is to come up with a username and password (see Figure 4.7). The IMA Manager generates the account for a first time user. Each new account has an anonymous identity created for it when the account is generated. However, once the empty account has been created, the user may use the IMA Manager to create identities. The account is stored in an encrypted file on the user's computer. The username and password used when creating the account is not stored anywhere else, it is left up to the user to manage this piece of information.

Figure 4.7: Create New Account

4.3.3 Identities

The identities tab allows a user to create new identities and manage the information in existing ones (see Figure 4.8). When a user creates new identities he or she is prompted to enter personal information. It is not necessary for a user to enter all of the information requested, but instead to enter only the information a user would want a business associated with this identity to know.

Figure 4.8: IMA Manager Identities Tab

When a user is creating identities, it may be useful to create several standard identities such as the identities: anonymous, personal, work, school, etc. The user must select one identity to be the default identity. This is the identity that will be at the top of the user's list of identities to assign to a participating business on a first visit if the user does not yet wish to provide an identity with more information. For example, a default identity that could be given to each business a user visits for the first time may be an identity called "Anonymous", an identity that does not contain any personal information. This "Anonymous" identity allows the business to begin to gather information on the user's interests over repeat visits, but does not contain any of the user's personal information contained in other identities. As a user begins to trust a business more, the identity associated with this business can be changed using the identity option on the history tab to associate this site with another identity such as one called "Personal".

An identity may store more than just the information shown in the prototype implementation of the IMA client, but may also be used to store information such as language preference, credit card number, etc. To include additional information all that would be required is that the schema that defines an identity be extended for additional information and appropriate security measures be taken based on the type of information being handled (e.g., credit card number).

4.3.4 History

The history tab displays a list of recently visited web sites (see Figure 4.9). The information about each web site provided on this tab is: URL, identity, and view information link.

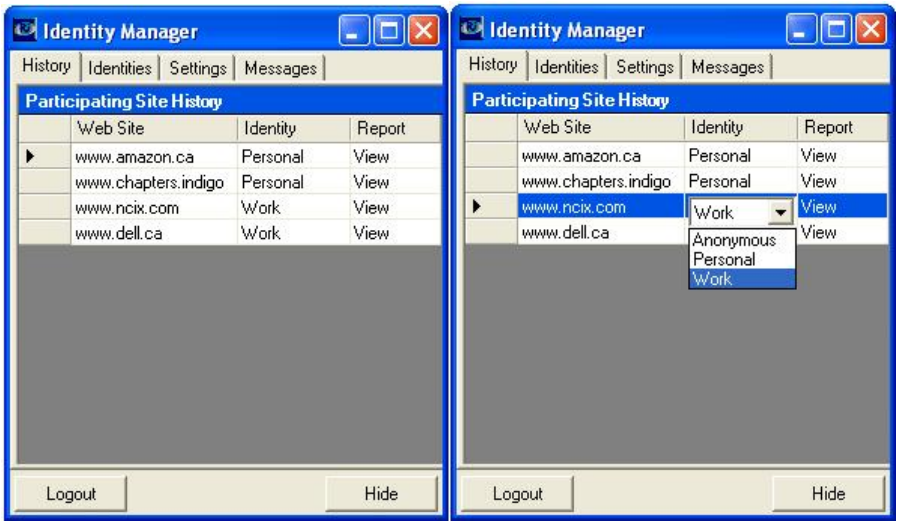

Figure 4.9: IMA Manager History Tab

The main purpose of the history tab is to give the user an up-to-date list of recently visited web sites and allow the user to modify information, such as the identity associated with a web site, while browsing the Internet. With each web site summary on the history tab there is a “View” link. This link is used to retrieve all available information the business currently has about the user and display it in a window. The user may either close the window after viewing this information, or choose to select items to be removed and save these changes. Saving any changes will cause an update to be forwarded to the business.

4.3.5 Viewing Profile Information

The view profile information window is opened by the “View” link on the history tab that is associated with each business visited (see Figure 4.10). By clicking the “View” link all information a business has about the user is retrieved from the business and displayed in the view profile information report window. This information can include identity, browsing history, personal preferences, and any other information the business has tracked about the user. This list of information will depend on the business. If the

user decides to remove some information that the business has shown, this can be done on the view profile information window by marking each piece of information to be removed by using the delete option provided. When this is done the request for removal of information is forwarded to the business.

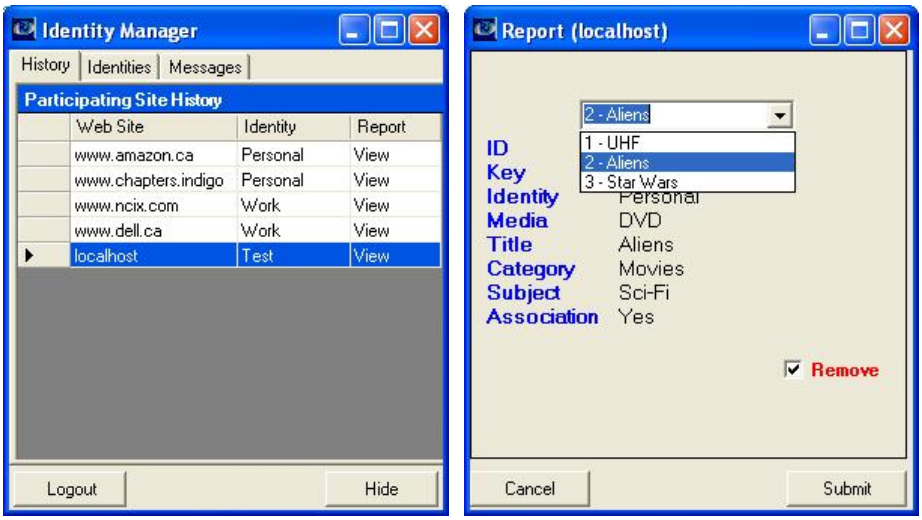

Figure 4.10: IMA Manager Site History and View Report Windows

4.3.6 Identity Update Forwarding

If the information in an identity becomes out of date, it may be updated by using the IMA Manager's identities tab. When the information in an identity is updated and saved, it is not just saved to the user's account, but the IMA Manager also takes the updated identity and automatically forwards it to all participating businesses the user has associated with that identity. This feature removes the hassle of having to update the same information at several businesses. Also, this allows businesses to ensure they have on file accurate contact information for each customer.

4.4 IMA Web Service

Each business that participates in the IMA system must implement the IMA Web Service. All communication between a business and each user's IMA Manager is handled

through this web service. The business decides how to implement this web service and connect it to the user information stored in its database. Each business must provide a uniform software interface required for connection to the IMA Manager.

The web service provided by each business will have a standard URL. For example, if the business has the domain name “www.mycompany.com”, then the address for the business’s IMA Web Service would be:

“www.mycompany.com/ima/imaservice.asmx”.

The reason for this standard IMA Web Service name (e.g., “…/ima/imaservice.asmx”) at each participating site is to allow the IMA Manager to easily connect to participating sites when the user is moving from one web site to another.

4.4.1 Web Service Interface

For the IMA Web Service a basic set of methods has been identified for each business to provide (see Figure 4.11). Each of the following methods are required for a user’s IMA Manager to communicate with the participating business. These methods are:

- `public bool Authenticate( string key )`
    - Checks if the identity (identified by the unique key given) has been used at the current business or not. Returns true if the identity exists.
- `public void AddIdentity( Ima.Manage.Identity identity )`
    - The first time a user visits a participating web site and decides to create an association between one identity and this web site, this method is used to add the new identity.

- `public void AddIdentityParams( string key, string id, string firstname, string lastname, string email, string company, string address, string city, string stateprov, string postalcode, string phone )`
    - Performs the same function as the method AddIdentity(), however it allows the identity to be added without having to create an Ima.Manage.Identity object. The only fields required are the key and id (identity name).
- `public Ima.Manage.Identity GetIdentity( string key )`
    - First checks to see if an identity for the key given exists. If it does it gathers all the information for that identity and returns it.
- `public void UpdateIdentity( Ima.Manage.Identity identity )`
    - When a user updates and saves information for an identity, the IMA Manager calls this method at each business this identity has been associated with to ensure each business has the updated information.
- `public void UpdateIdentityParams( string key, string id, string firstname, string lastname, string email, string company, string address, string city, string stateprov, string postalcode, string phone )`
    - Performs the same function as the method UpdateIdentity(), however it allows individual items to be updated without having to pass an Ima.Manage.Identity object. Of course, with any update, the identity's key is always required.
- `public void AddProfile( string identity, string key, Ima.Manage.Profile profile )`

- When a user who has an identity associated with a participating site is using the site, the business may add new profiles to that identity at any time. For example, if the user seems to be showing an interest in a specific product by browsing and reading about it on the web site, a new profile could be added to this identity for future reference.

- `public void AddProfileParams( string key, string media, string title, string category, string subject, string identity, string association )`
  - Provides the same functionality as method AddProfileParams(), however it allows for fields in a single identity to be updated without requiring the creation of an Ima.Manage.Profile object. The identity key is the only required field.

- `public Ima.Manage.Profiles GetProfile( string key )`
  - Retrieves all profile and history information a business has associated with the identity identified by the key.

- `public void UpdateProfile( Ima.Manage.Profiles profiles )`
  - If a user chooses to remove profile information using the IMA Manager, this method is used to pass the profile changes back to the business so these requests for removal of information can be handled.

- `public void UpdateProfileParams( string historyID, string key, string media, string title, string category, string subject, string identity, string association, bool remove )`
  - Performs as the same function as the method UpdateProfile(), however it allows individual items in a single profile to be updated without requiring

the use of the Ima.Manage.Profiles collection of Ima.Manage.Profile objects to be created. Both the historyID and key are required fields.

- `public void AddHistoryItem( string ipAddress, Ima.Manage.Profile profile )`
    - Each time a user views a product on the web site a new history item is added to the user's profile.

- `public void AddVisitor( string ipAddress, string identity, string key )`
    - When a user visits a participating web site and an association has already been made for this site, the IMA Manager sends the current IP Address of the user, along with the name and ID of the associated identity. This information is then used to track history information for the user's profile during the current online session.

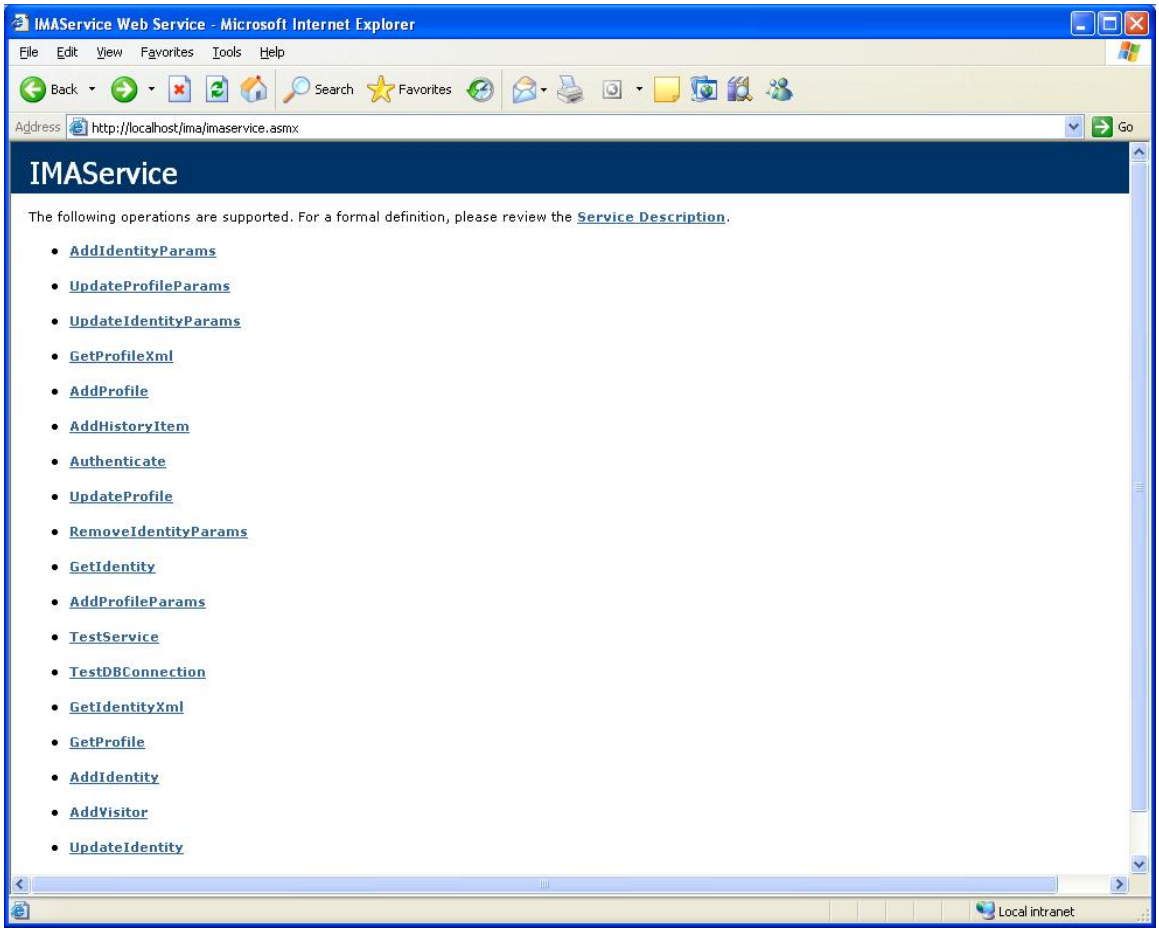

Figure 4.11: IMA Web Service interface implemented for the example IMA participating business

4.4.2 Web Service Implementation

How a business implements the IMA web service is dependent on the data structures or database architecture of its own system. Ideally, with a well-designed system, the implementation of each web service method would be accomplished by a single query to a database, although in a complex system this may require more effort to implement. One of the main reasons the list of web service methods is kept short and simple is to keep the work required for a business to participate in the IMA system to a minimum.

Information transferred using web services between users of the IMA system and businesses must be handled with a reasonable degree of security. XML web services are

based on the Simple Object Access Protocol (SOAP) specification [53]. There is an extension to SOAP defined in the Web Services Security (WS-Security) language specification designed by both IBM and Microsoft [22]. The WS-Security specification describes extensions to the SOAP standard that provide support for multiple security tokens and also add support for multiple forms of encryption [22]. Support for the WS-Security extension is offered to .NET developers in the Web Services Enhancements (WSE) package [36]. The WSE package makes it straight forward to implement in .NET a secure SOAP message that is signed with credentials (i.e., username and password) and has the body of the SOAP message encrypted [4].

4.5 IMA Participating Business

For a business to be considered participating in the IMA System it must implement the IMA Web Service. However, for a business to make use of information provided by IMA users, the IMA System must allow the business to receive associated identities and to track the actions of users in order to build a profile.

4.5.1 Identifying IMA Users and Tracking of Profile Information

There are several ways information about the current user can be passed to the business as the user browses the business's web site, such as browser session state, cookies, request parameters, etc. However, any of these approaches would make the Identity's key insecure. The solution to this is a simple approach. Each time the user visits the first page within a business's web site, the current IP Address as well as the associated Identity's name and key (if an identity is currently associated) is provided to the business. The business stores the IP Address to Identity key mapping in a database table. As the user views products and takes actions on the site, the business gathers

information it would like to add to the Identity's profile. This is done by simply passing the IP Address and the information to be added to the profile over the IMA Web Service. The business then does a look up of the Identity key based on the IP Address and adds the profile information to the Identity in the business's database. A diagram of this interaction is shown in Figure 4.12.

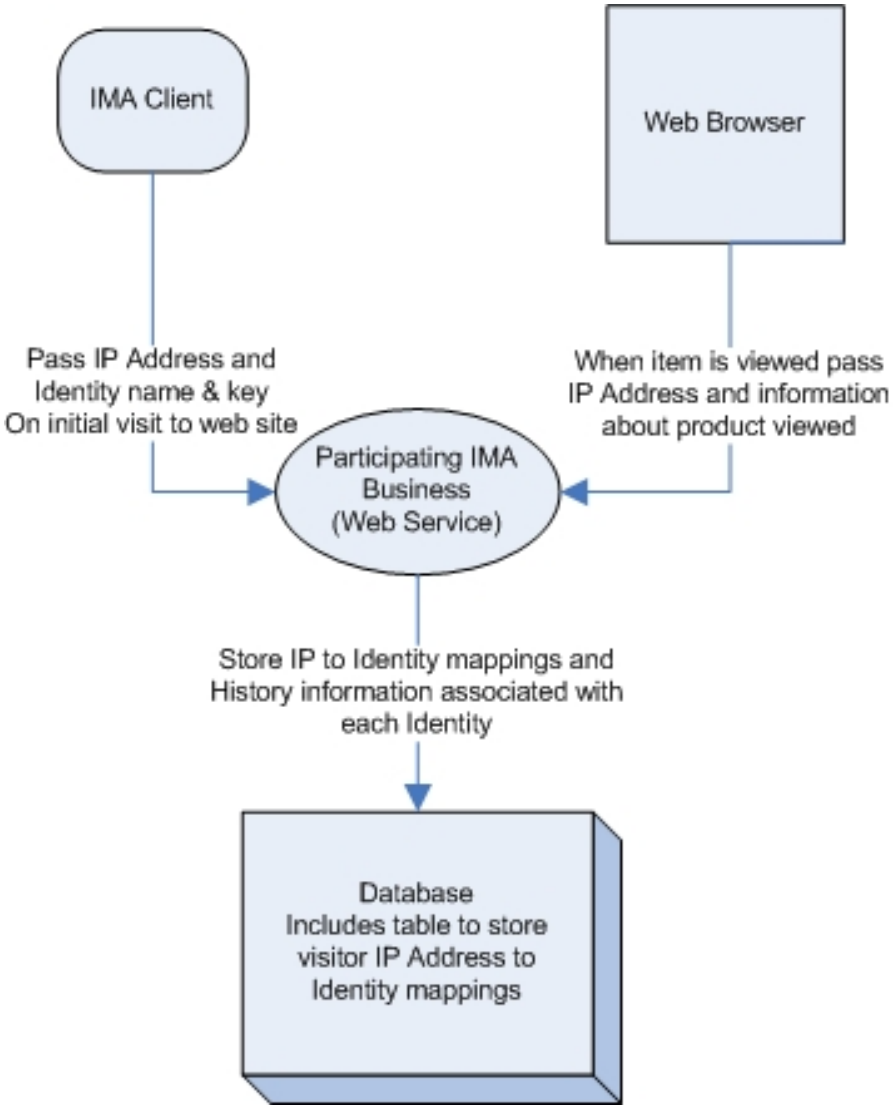

Figure 4.12: Identity Tracking and History Information

Information from IMA users only needs to be used in a maximum of three tables: Visitor, Identity, and History (see Figure 4.13). The Visitor table stores the mappings of

the current IP Address the IMA Client is using to the name and key of the Identity the user currently has associated with the business. When the user associates an Identity with the business, a new entry is added to the Identity table with the Identity's information. At the same time the Visitor table is updated with the IP Address and the current Identity name and Key being used. Now whenever an IMA user is accessing a participating business's web site, entries are added to the History table by using the IP Address provided to look up the Identity being used from the Visitor table and to add the History item and associated Identity to the History table.

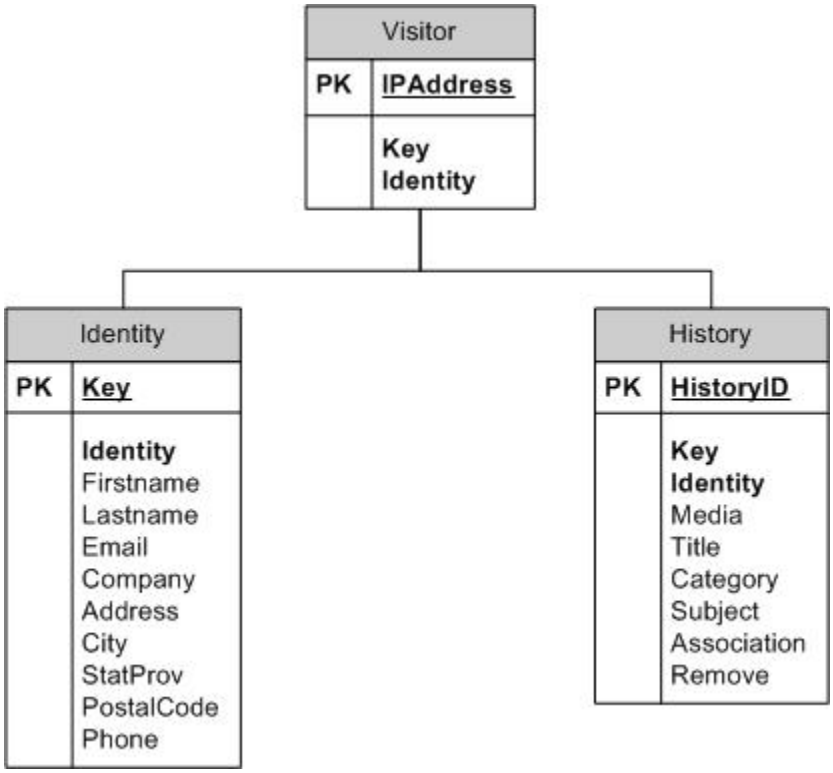

Figure 4.13: Database Structure

4.5.2 User's IMA Interactions with a Participating Business

Upon visiting the business, the IMA client application first determines if the business is participating, if so then the IMA client determines if the user has an association. If this is true then the IMA Manager provides the client IP Address, and the

associated identity's name and key. This allows the business to track the user's activities. For any item that is selected, the business's web page takes the client IP Address along with the information the business wishes to add to the user's profile and passes it to the business by the IMA web service and adds that information to the user's profile. This is done by looking up the current identity using the IP Address provided. In Figure 4.14 is a decision tree showing the possible courses of action that may be taken when a user of the IMA system visits a participating web site.

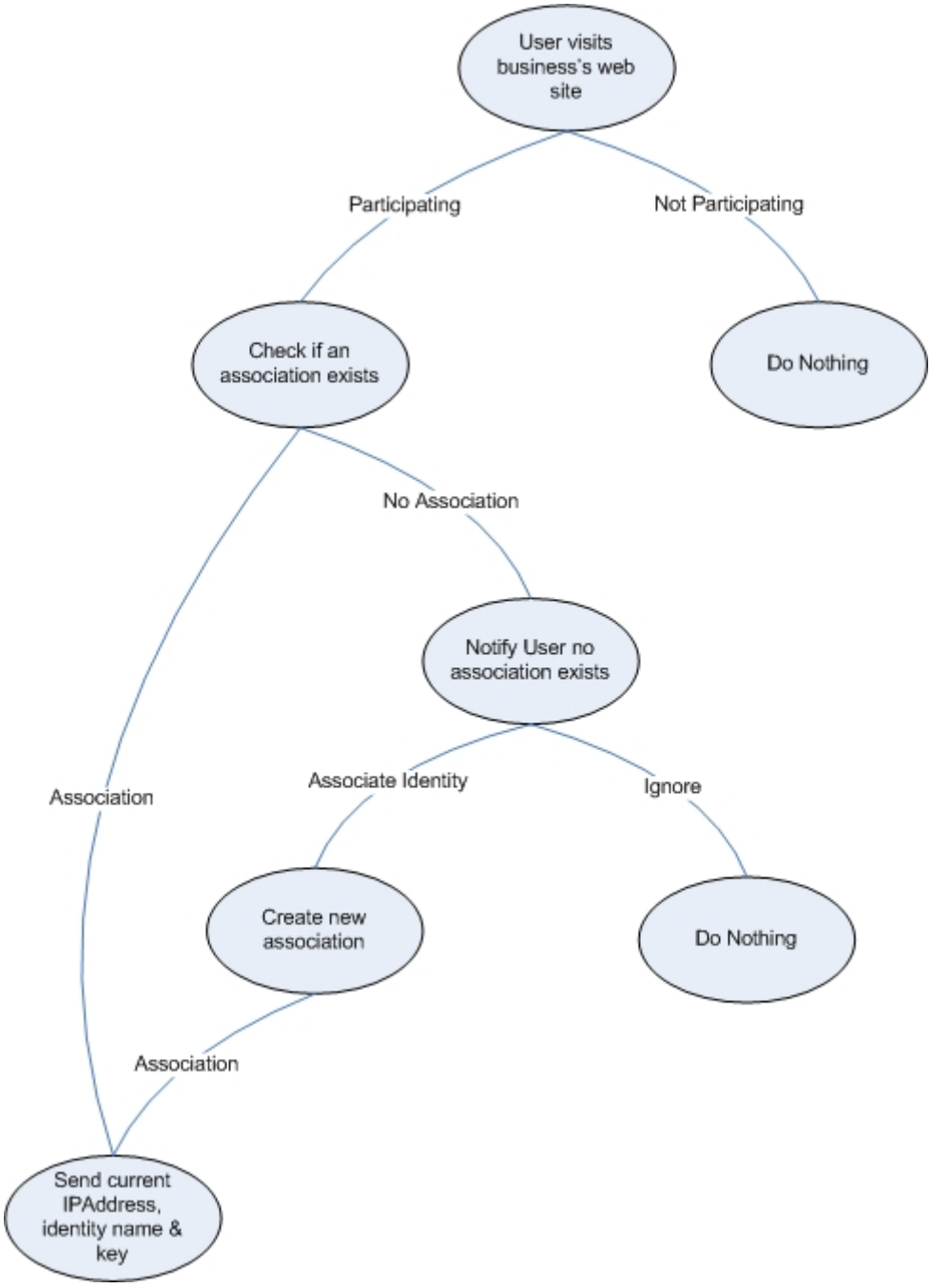

Figure 4.14: Sequence of possible actions taken while visiting a participating site

4.5.3 Example Participating Business

To demonstrate the functionality of both the client side and business side of the IMA system, a demo E-Commerce web site that sells movies, books, and software was created to act as a participating IMA web site. When a user of the IMA system visits a

participating site, the site must be given some basic information to be able to track the visitor while browsing the web site. This is done for the purpose of allowing a business to identify a returning IMA user and build a profile of information associated with the given identity.

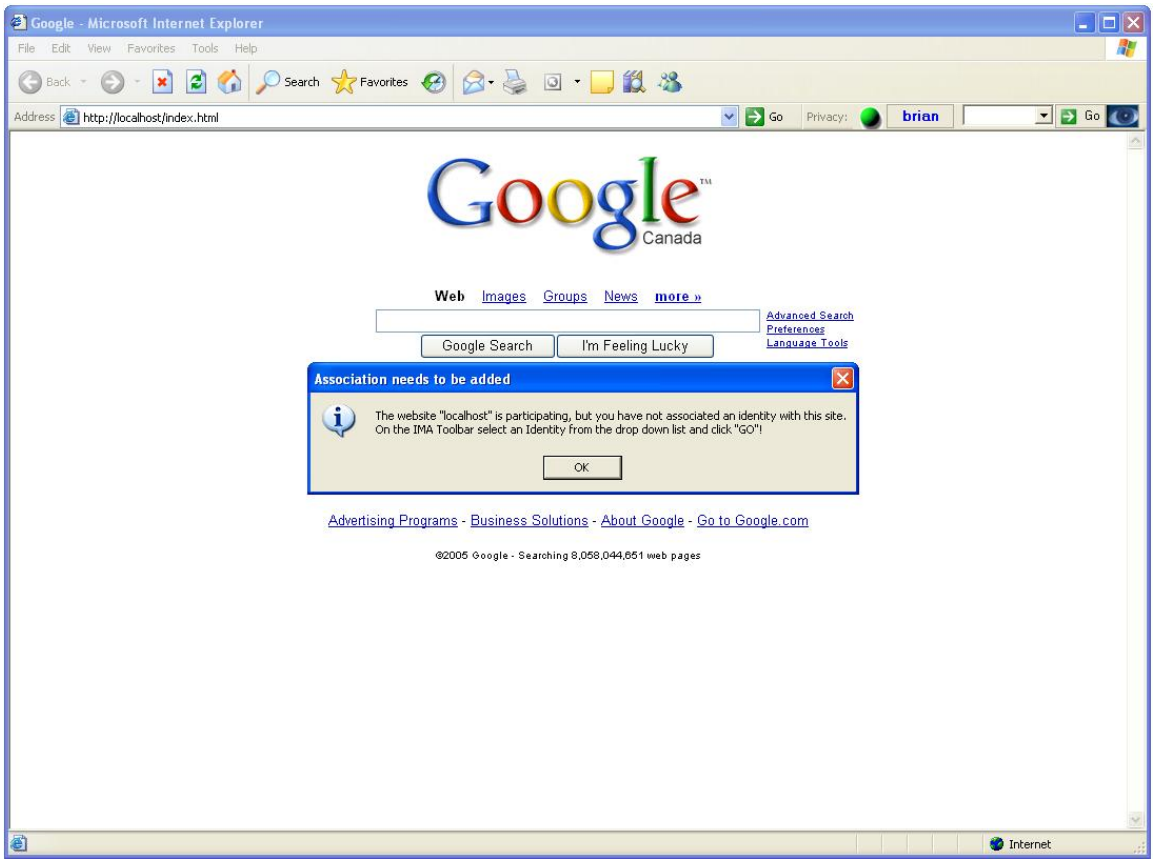

Figure 4.15: Navigating to participating business, no identity associated

When a user enters a URL of a business, the IMA Manager checks to see if the site is participating before the browser is forwarded to that site. If the site is participating and the user currently has an identity associated with that site, then that identity is automatically used. However, if no identity is associated with that site, the user will be notified by the popup message (see Figure 4.15). If the user chooses to ignore this message he or she will still be able to view the business website. No identity information will be sent to the business and the user will remain anonymous.

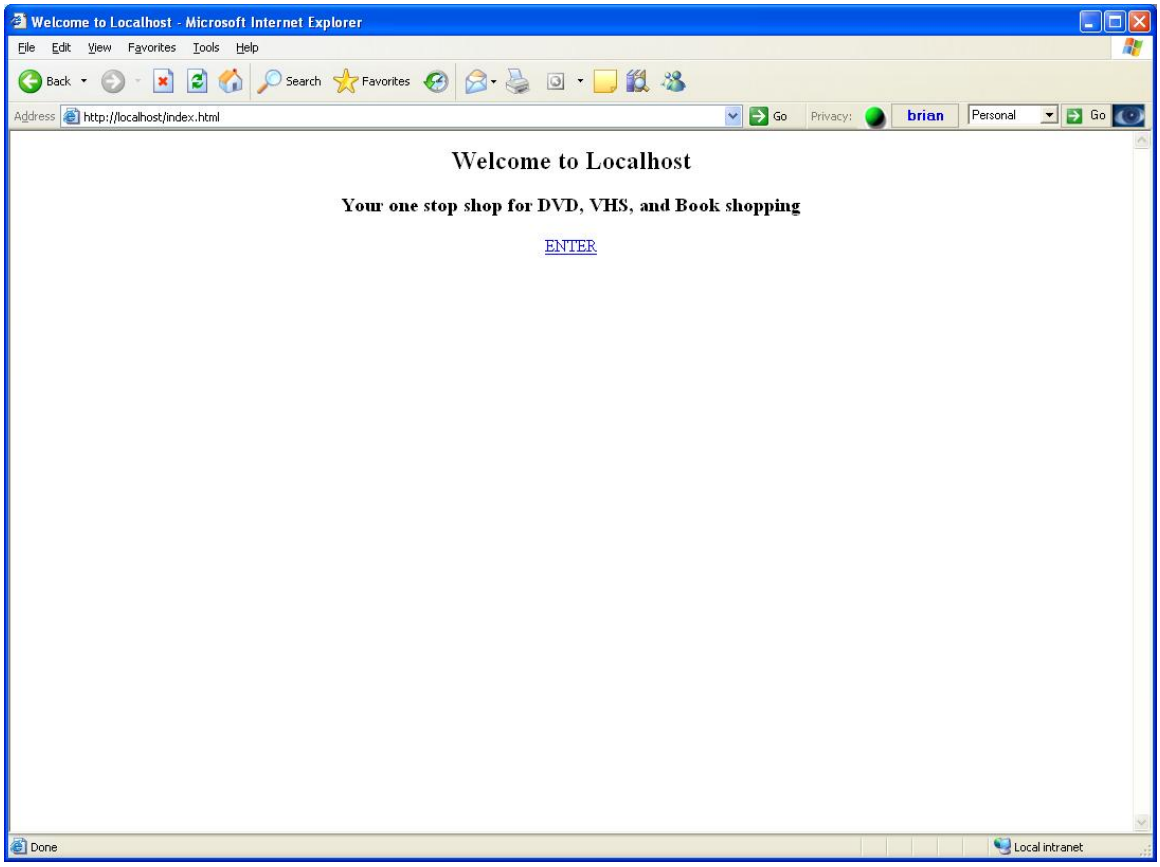

Figure 4.16: Index page of business, no association made yet

Once the IMA user has reached the first page of the business web site with which an identity was previously associated (see Figure 4.16), that identity has already been forwarded along with the IP address to the business. If there was no previously associated identity, at this point no identity information has been given to the business and the user's actions are not being added to a profile associated with any identity.

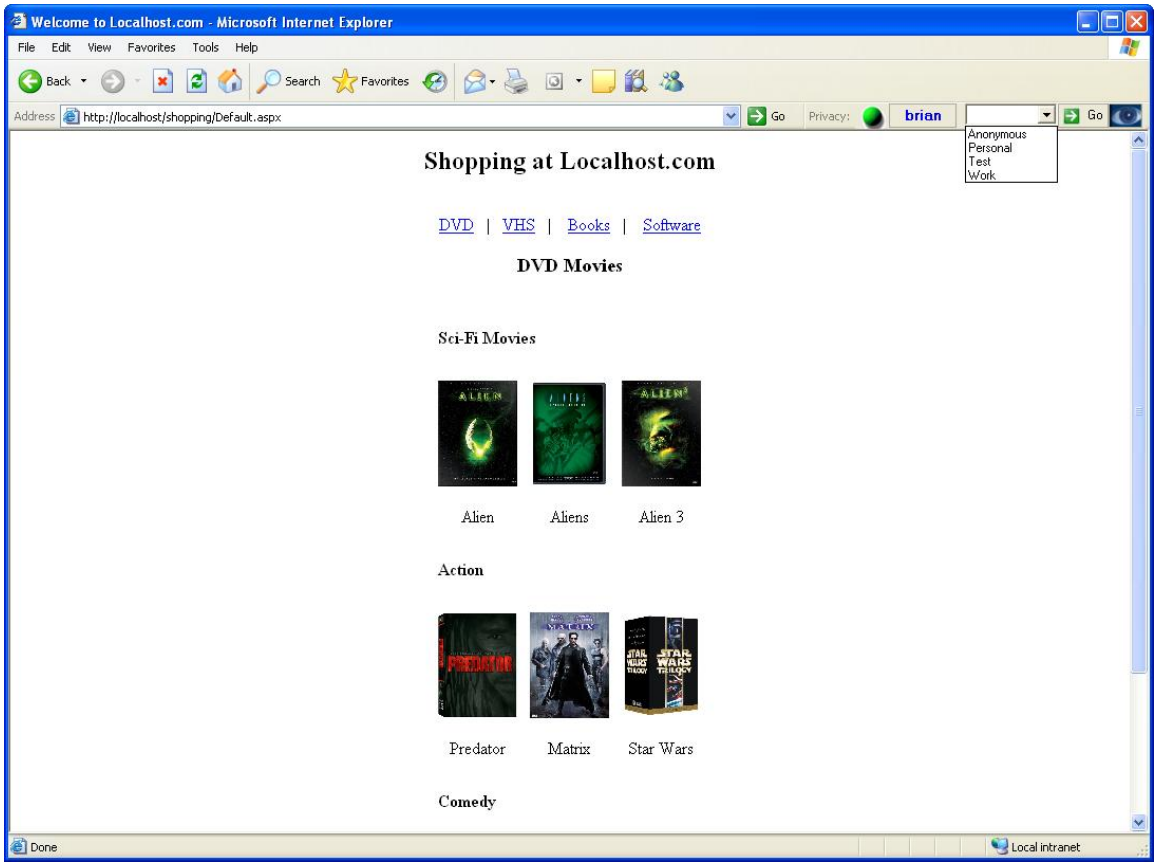

Figure 4.17: Selecting an existing identity from the list in the IMA Toolbar

If the current user has no identity associated with the business, an identity may be added at anytime while viewing the participating business's web site. The user selects one of the existing identities from the drop down list in the IMA Toolbar and clicks the "Go" button (see Figure 4.17). The IMA Manager then sends the current identity to the business along with the user's IP address, and adds a business to identity association to the user's user account.

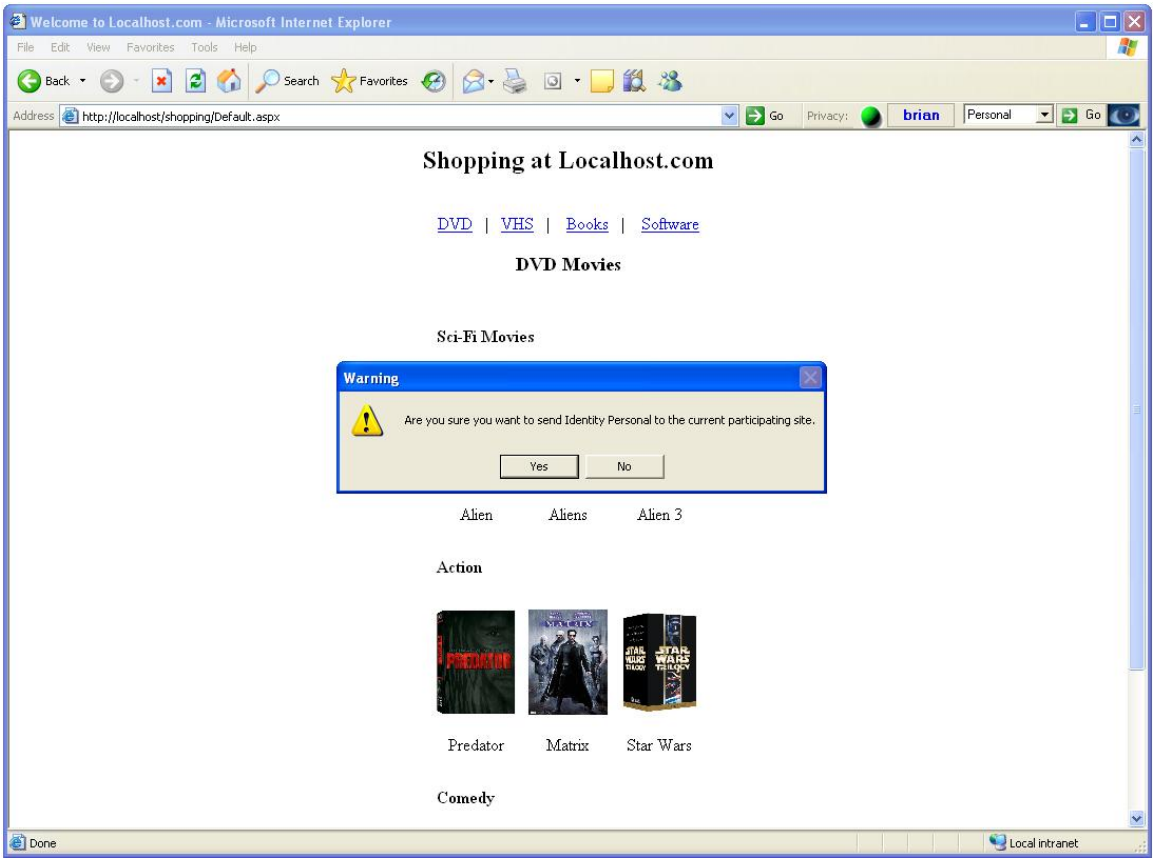

Figure 4.18: Confirmation before creating the Business-Identity association

Before an identity is associated with a business and the identity information is provided, the user is asked to confirm this action (see Figure 4.18). If the user says “Yes”, the identity is provided and the association is now made. This association will be used automatically on all subsequent visits to the business’s web site until the user changes the associated identity.

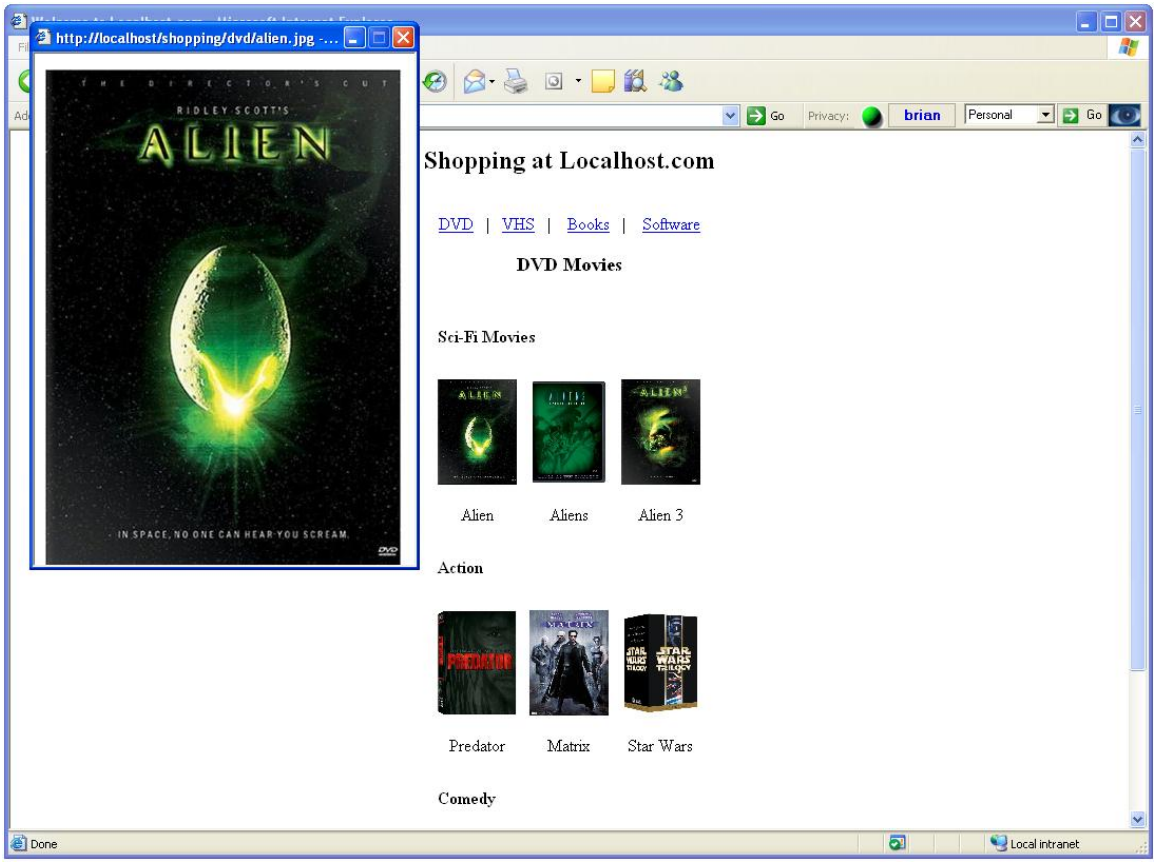

Figure 4.19: Viewing business products with an identity associated

Now that the association has been established, all actions the user takes on the business's site may now be tracked by the business and used to build a profile of the user's interests and preferences to be linked with the current identity. Figure 4.19 shows an example of the user clicking on a movie to view. The business will use this action to track information such as the user is interested in Sci-Fi movies on DVD and the movie "Alien". This profile information may be used by the business for a number of reasons, for example, to make recommendations about other products the user may be interested in based on the information gathered in the identity's profile.

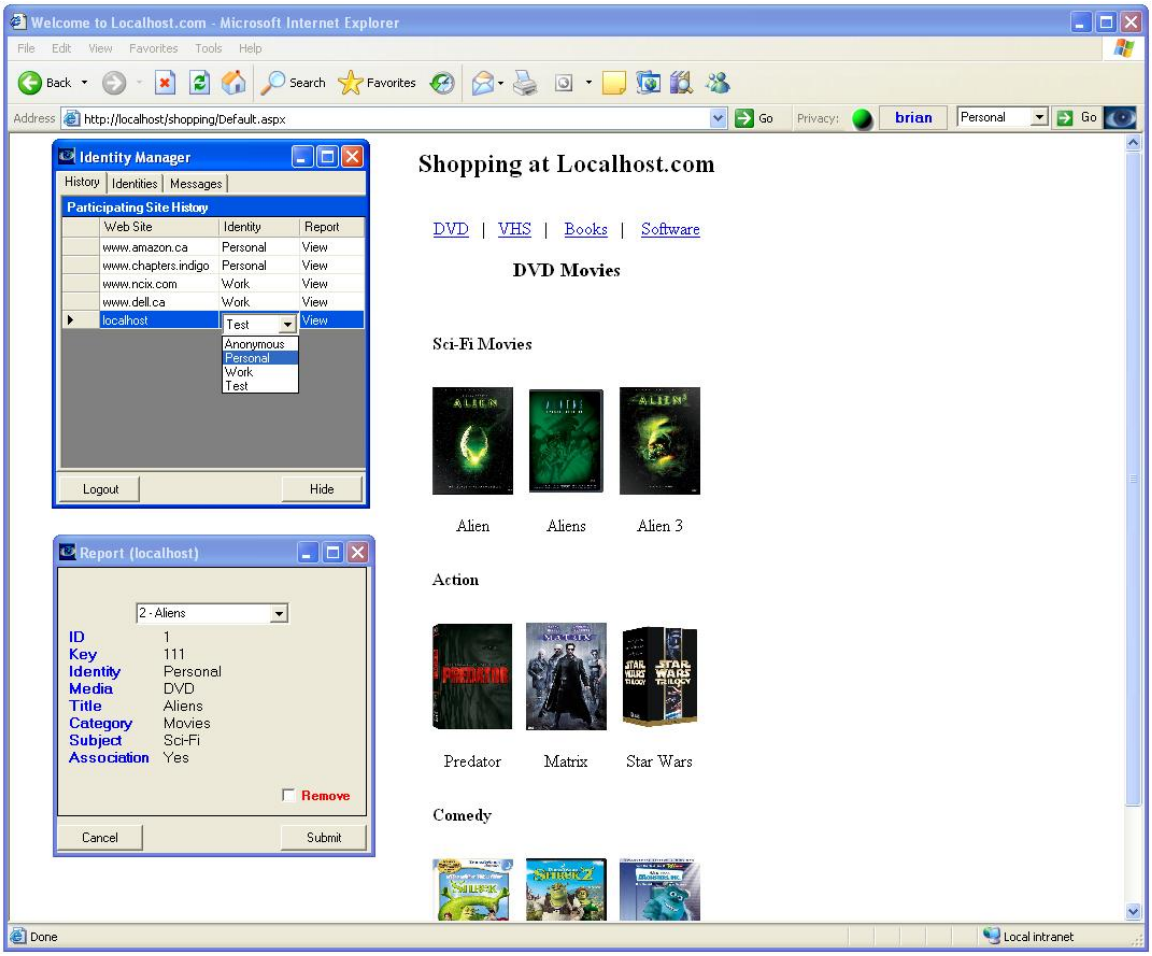

Figure 4.20: Viewing information a business has associated with an identity

Now that an identity has been associated with the business and the user has taken some actions that allow the business to add information to the user's identity to help build a profile of the user's interests, the user may use the IMA Manager application to view this information. By clicking the "Eye" logo on the right side of the IMA Toolbar the IMA Manager application is opened. In the list of participating sites that the user has visited the user selects one of the identities that has been associated with a participating site and then clicks the "View" link under the "Report" column. This sends a request to the business's IMA web service and displays the returned information in a separate window as shown in Figure 4.20.

4.6 XML Data

4.6.1 User Account

When a user creates a new account for the IMA system, all information for this user is stored on the client side in a single XML document [57]. The structure of the user account consists of three main parts: the user's basic account information, the user's identities, and the user's associations of businesses to identities. The structure for the user account's XML document is defined in an XML Schema Definition (XSD) [56].

Each account contains one or more identities that are created by the user. As the user visits web sites, one or more identities will be associated with each site. One of these identities must be selected as the default identity. This default identity is used for any participating web sites the user visits that have not yet been associated with an identity. Each identity that is created contains a unique key generated by the IMA Manager. As the user visits participating online businesses, this key is given to each business to allow it to identify the user's identity on subsequent visits. Each time a user selects a new identity for a business, a new association is made between the user's identity and the business. Each association is added to the user's account under the "associations section". An example user account schema is shown in Appendix C and an example user account XML file is shown in Appendix D. Figure 4.21 provides an overview of the design of a user account schema.

When a user visits an on-line business, information is automatically forwarded to the business in an XML document called an Identity. This document is a subset of the user account, since it contains only the information for a single identity and allows the user's other identities and associations to remain unknown to the business. An example identity schema is shown in Appendix G and an example identity XML file is shown in

Appendix H. The “Identity” element in the schema shown in Figure 4.21 provides an overview of the design of the identity schema.

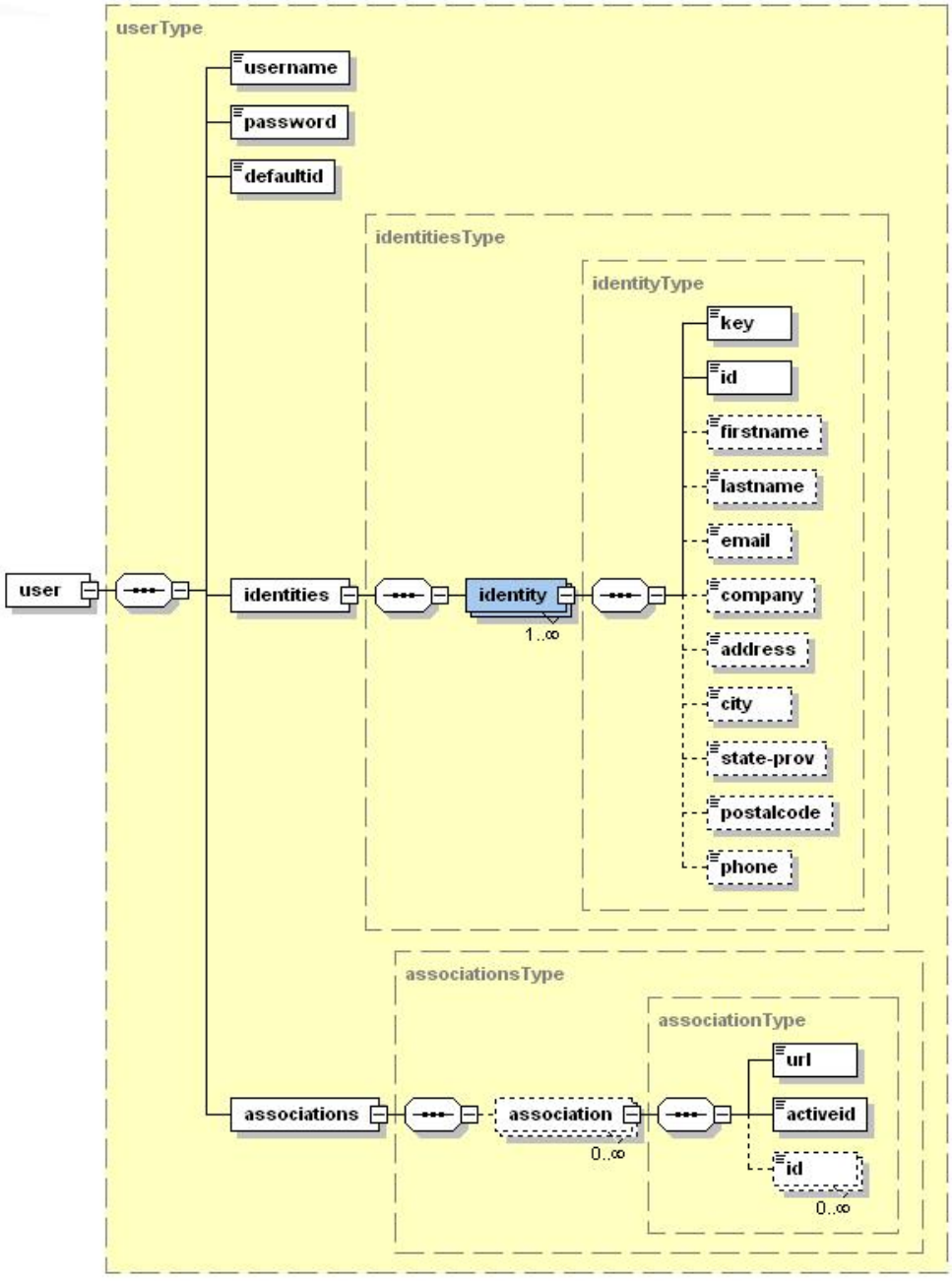

Figure 4.21: User Account Schema

4.6.2 Report

When a user requests to see what personal information a business has about him or her, the business responds to this request with an XML document. The structure of this document is divided into two sections: business information and identity information. The business information returned is just basic contact information to be displayed with the report. For each identity the report returns the unique key used for the identity, the name of the identity, and the profile information for that identity. The profile section of each identity contains information on what types of products the user has purchased or browsed. This is the information the business has used to determine the user's personal interests. An example report schema is shown in Appendix E and an example report XML file is shown in Appendix F. Figure 4.22 provides an overview of the design of a personal information report schema.

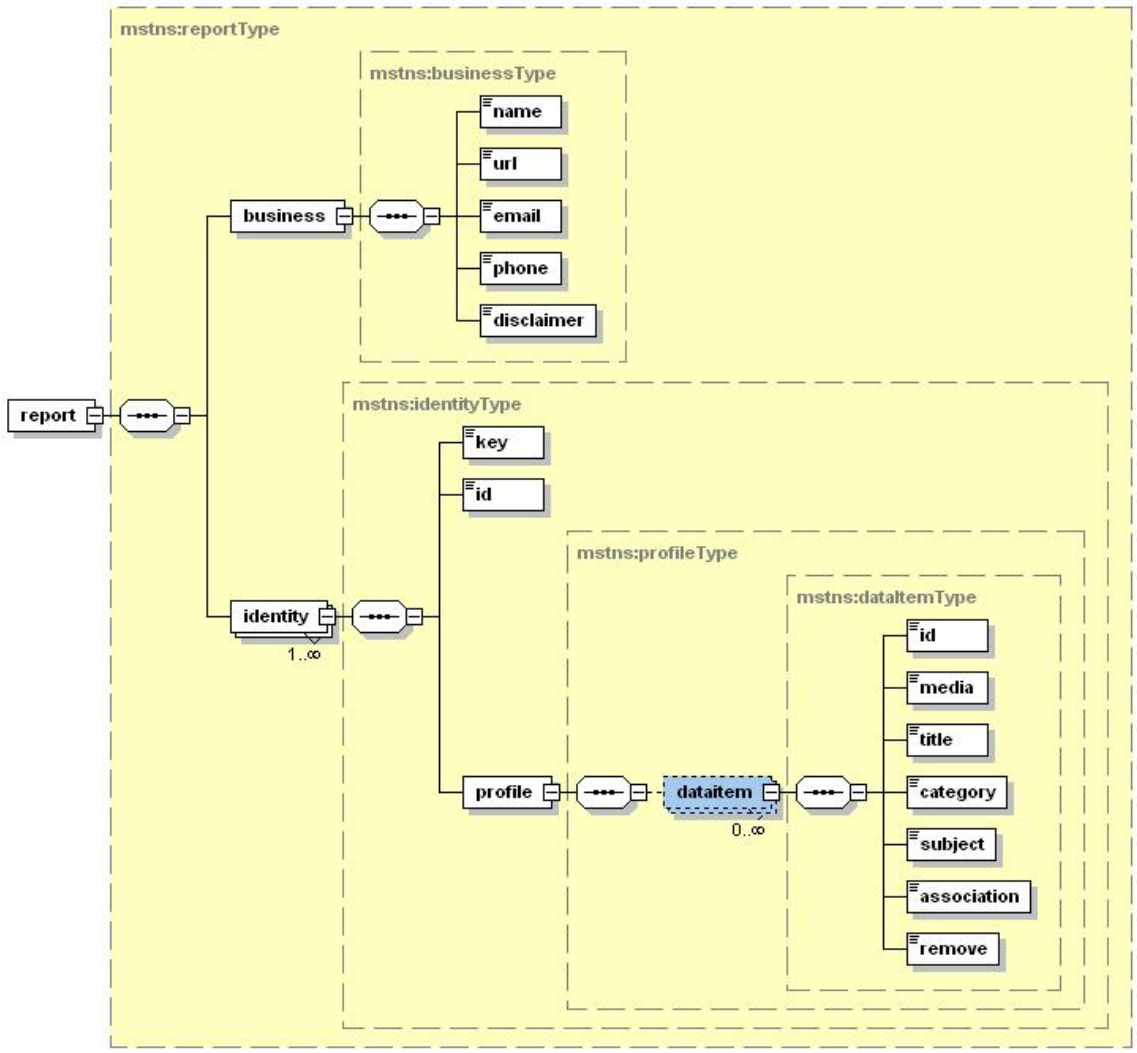

Figure 4.22: Personal Information Report Schema

The “DataItem” element shown in Figure 4.22 represents a single item from the user’s profile that the business has created about the user’s potential interests based on his or her activities. For example, a DataItem could be a category of books, such as science fiction, that the user has browsed on several visits to the business’s web site. Each of these profile items is returned in a simplified form to allow the user to look them over and remove any items that no longer represent the user’s interests.

If a user has associated more than one identity with a business, and the user makes a request for a report, information for each identity is received in a response to a request for information sent to the participating business's web service. One request is made for each identity. Once all requested information is received, it is compiled by the IMA Manager and displayed to the user. For each profile information item displayed to the user, a remove option is provided. Selecting this option for an item will mark it for deletion. Once the user is finished, the report is returned to the business indicating all items to be removed.

4.7 Identity Management

4.7.1 Identities

The IMA system allows users to set different identities, for example for personal, work, and private related activities [13]. Having the ability to easily separate these identities allows web users to ensure that the information recorded about them at certain web sites is based on the identity they have assigned for that site. Users are able to select a default identity to be used for all web sites they visit for the first time, but will be able to switch to another identity when desired. The IMA Manager also allows users to view a list of recently visited web sites and attach some alternate identity to any site at any time. This change results in the information for the alternate identity being immediately forwarded to the business. This alternate identity is then used for subsequent visits to this site.

When shopping at a number of different on-line businesses, it is convenient for customers to have each business store some of their personal "account" information such as name and mailing address and connect to that account with a username/password.

However, when personal “account” information changes, the user is normally required to update this information manually for each business. The IMA Manager allows a user to update an identity and have those changes automatically forwarded to all businesses with which this identity has been associated.

If a business is participating in the IMA system, there may be no need for a first time visitor to the web site to have to fill out a long form to access a feature the business offers, as is often required for download of trial software. Instead, as the user accesses the site, he or she may choose to associate an identity with the business. Only after the IMA Manager receives confirmation from the user is the identity forwarded to the business by the IMA Manager.

4.7.2 Personas

When people shop on-line it is not always for just one purpose but sometimes for many reasons such as purchasing an item for themselves or purchasing a gift for someone else. The IMA Manager allows users to create multiple identities that contain much the same information but refer to different on-line browsing activity. These are called Personas. For example, users can have one persona when shopping for products for themselves and another persona when shopping for someone else. This alternative persona can be created with the same standard account information as the personal identity (i.e., name, mailing address, email, etc.), but is associated with a different product browsing and buying history.

Since multiple personas at a single business all contain the same identifying information such as name, address, email, etc, the business is left to decide how it is going to use these distinctions. This gives a business the opportunity to choose how to

custom tailor the content the user sees. The business may choose to keep the information gathered on the different personas completely separate to allow the user to receive more content based on the interests established by active use of the persona. However, the business could also make a combined use of the personas, such as a shopping site that uses the history of the currently used persona to make recommendations of products the user may be interested in, but could also use the history of another persona to decide what types of products to display in the banner ads on the site. This is just one example. The danger here is that a business may completely ignore the separation of histories between personas and result in driving away the user with content that does not target the user's current interests.

4.7.3 Identities vs. Personas

A persona is essentially an identity. It is just that multiple personas can each use the same personal information from a single identity. If a user wishes to create two sets of personal information, the user would create two separate identities for this information. However, if the user would like to use the same identity information, but perhaps wanted a business to associate two separate sets of personal interests with the same set of personal information, the user would create two personas that share the same set of personal information. As far as the IMA Manager and the business is concerned a persona is an identity and is treated as such.

4.8 Implementation

The implementation phase of this thesis was where a great deal of time was spent. A number of design issues were raised during the implementation phase due to

limitations with the tools being used. The following sections provide an overview of issues with the IMA system implementation.

4.8.1 Platform Dependence

For the implementation of the IMA system for this project, both the client and business side of the system were implemented using the Microsoft .NET Framework and therefore are dependent on Microsoft Windows Operating Systems [34]. The prototype of the IMA system could have been written in any language, however for convenience the .NET framework was chosen due to its ease of integration of custom made toolbars into IE.

The IMA client consists of two separate applications that have been combined to form the client: the IMA Toolbar and IMA Manager. The IMA Toolbar is an IE Toolbar that when registered with IE appears in the list of toolbars available in IE. The IMA Toolbar only works with the IE web browser. The IMA Manager is a Windows Forms application that is accessed through the IMA Toolbar. For these reasons the IMA client must be used on a computer with a Microsoft Windows operating system installed that has the .NET Framework version 1.1 or better installed, and the only browser that may be used with the client is IE. Although .NET was chosen for this implementation, the IMA system is not dependent on any one programming language or platform. Since the IMA system uses web services for the client and business to communicate, client applications may be implemented using any programming language (e.g., Visual Basic, C++, Java, etc.) and any platform (e.g., Windows, Linux, etc.).

The example implementation of the IMA Business Web Service built for this project used .NET Web Services as well as a Microsoft Access Database. However, this

was done as a demonstration of the business side of the system. The IMA Web Service specification outlines the web methods a business must make available in order to participate in the IMA System, but there are no restrictions as to what platform, programming language, or database the business may use. Web Services were used intentionally for this purpose. It is left up to the business to implement this web service interface in the manner that best works with the implementation of its existing system.

4.8.2 Setup and Installation

The prototype implementation of the IMA system is contained in a single visual studio solution. This solution file contains all projects for both the IMA client application as well as an example participating IMA business that has implemented the IMA web service. For testing purposes, and due to limited resources, a single machine was used to demonstrate the IMA Manager (Client), the IMA Web Service (Business), and the example business web site. No Microsoft Windows installation file has been created for the IMA application, so this section describes step-by-step how to setup the prototype IMA implementation on a single computer.

The first step after obtaining the software is to set two paths in the code before the project is built. After opening the IMA Visual Studio solution, go to the “IMA Toolbar” project, then to the “Manage” folder and open the “Account.cs” file. At the top of the file is a constant called “BASE_PATH”. Set this path to be to the absolute path on the local machine to the “Schemas” folder which is located in the root folder of the IMA project. This path is used by the IMA Toolbar and Manager to access new account template XML files and is also where the XML files for newly created user accounts are stored. Next the IMA Web Service needs to have the absolute path to the example business database set.

Open the IMA Visual Studio solution, go to the “IMA Service” project, and open the “IMAService.asmx” file. At the top of the file is a constant called “SOURCE”. Set this path to be the absolute path on the local machine to the “personalmanager.mdb” database which is located inside the “imaservice” folder which is in the root folder of the IMA project.

The second step is to add two virtual directories required for the example IMA business. Open Internet Information Services (IIS) (as shown in Figure 4.23). The first virtual directory that needs to be added is for the business’s IMA Web Service that will allow users using the IMA Manager to communicate with the participating business. Create a new virtual directory called “ima” and point it to the location on the local machine of the “imaservice” folder located in the root folder of the IMA Project. The second virtual directory that needs to be added is for the example business’s web site. Create a new virtual directory called “shopping” and point it to the location on the local machine of the “Business” folder located in the root folder of the IMA Project. This will allow the user to browse to the example business web site using a URL such as: “http://localhost/shopping”. It may be useful to add an “index.html” page to the root folder in IIS that takes users to the index.html page located in “http://localhost/shopping”.

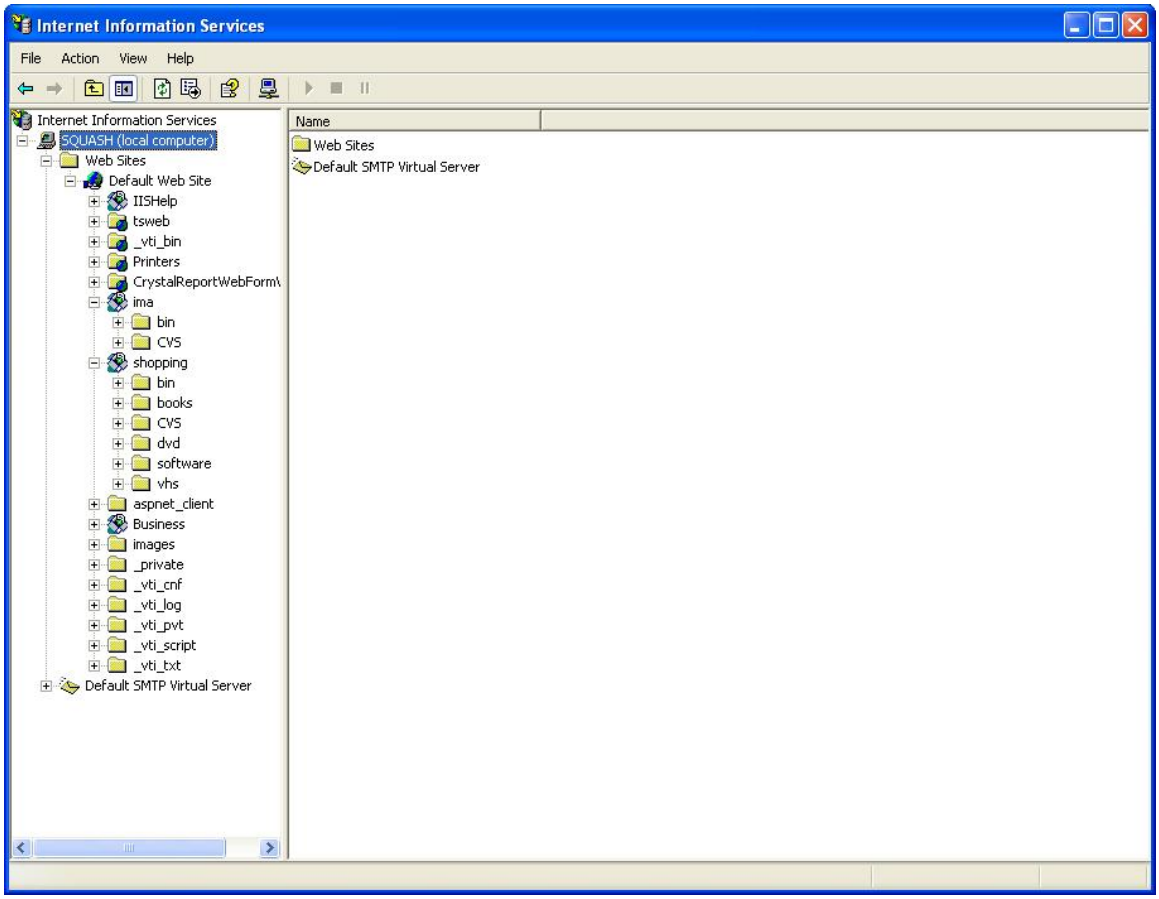

Figure 4.23: Internet Information Services “ima” and “shopping” virtual directories

The third step is to build the IMA solution. Open the “IMA.sln” solution file in Visual Studio. This solution file contains all projects used in both the client and business (as shown in Figure 4.24). In Visual Studio’s “Solution Explorer” right click on “Solution IMA” and select “Rebuild”. This will build the IMA Web Service, the example participating business, IMA Toolbar, all supporting libraries required for the IMA client application to work with IE, and finally it adds all required libraries to the Global Assembly Cache (GAC) along with the IMA Toolbar Dynamic Linked Library (DLL). Now the sample participating business is ready to be accessed by the URL “http://localhost/shopping/” and the IMA Toolbar is now accessible from the menu in IE under “View → Toolbars → IMA Privacy Bar” (as shown in Figure 4.1).

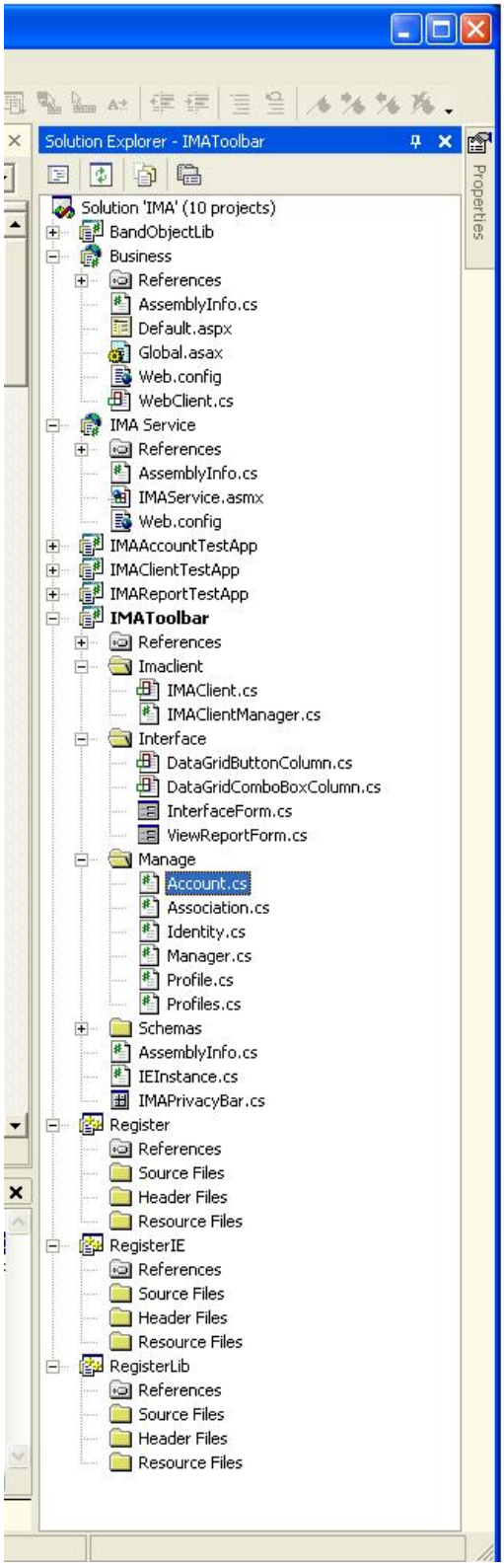

Figure 4.24: IMA Visual Studio solution for client, web service, and example business

All projects used in the IMA solution are shown in Figure 4.24. The "BandObjectLib" project provides the library of base functionality for creating Toolbar Bands. This library is basically a .NET wrapped for the Component Object Model (COM) components used to create add-ons for IE. The "Business" project provides the example participating business web site. The "IMA Web Service" project provides the web service implementation for the example business to allow IMA client applications to connect to a participating business. The "IMA Toolbar" project is the complete client application which includes the IE toolbar for the IMA client, as well as the windows forms IMA Manager application. The "Register" project places the IMA Toolbar project's DLL in the GAC and also registers it as a COM component. This allows the IMA Toolbar to be available to IE. The "RegisterIE" project adds the shdocvw.dll (Shell Document Object and Control Library) to the GAC. The "RegisterLib" project adds the BandObjectLib.dll to the GAC.

4.8.3 Scalability

In a SSO that relies on a trusted third party, such as .NET Passport, scalability can be an issue especially as the system grows in popularity and the resources required by the third party to operate increase. The backbone of the .NET Passport system is a centralized service owned by Microsoft that handles the management of users' personal information and provides the third party service that acts as the link between customers and businesses. With this type of system, the more businesses and customers that are involved the greater the resources the third party must provide.

What makes scalability with the IMA system so simple is that it does not rely on a centralized third party service for management of personal information or for

authentication. This is possible because of the two main design decisions made in the development of the IMA system:

1. All identity sharing and authentication is done directly between the user and the business
2. Client side management of the user account, identity information, and identity to business associations

4.8.4 Known Issues

There is currently a bug in the .NET framework that prevents the "BeforeNavigate" event from firing whenever an instance of IE attempts to navigate to a new web page [39]. Without this event firing the IMA Client cannot perform any of its basic functions. If this event is not being fired by the user's IE browser, the IMA client is rendered useless since it cannot perform any of its basic functionality such as:

- obtaining the URL of the current website being visited
- determining if the current website is a participating member of IMA
- checking if an identity is currently associated with the business being visited
- automatically providing an associated identity to a business

If the system the IMA Client is being installed on currently has .NET Framework Service Pack 2, the Post-Service Pack 2 .NET Framework Core CLR Hotfix Package may be installed to fix this problem [38].

Another issue related to the "BeforeNavigate" event occurs when a second IE browser window is opened, which creates a second instance of the IMA Manager application. Each instance of the IMA Manager must register with an open IE browser window to receive the "BeforeNavigate" event. Then the second instance takes control of

the “BeforeNavigate” event from both browsers causing the first instance of the IMA Manager to lose its event notifications and therefore it no longer receives updates when the browser is navigated to another website address. So the user closes the second browser window that was opened and this also closes the second instance of the IMA Manager. Now the user is left with the original IE browser window open and no “BeforeNavigate” events being received by the first IMA Manager that was opened.

A solution would be to suppress the IMA Manager from opening automatically when a single instance already exists, however this will not work in IE. If the IMA Toolbar is selected to be displayed in the first instance of IE, then it will automatically be opened in any additional browser windows that are opened. IE does not allow for one instance of IE to have different initial settings from another instance of IE.

4.9 Limitations

Personal information in the IMA system is always stored in two places; at the participating businesses the user has visited, and on the user’s computer. Any responsible business will have security measures in place to protect stored personal information. However, the information stored on an individual’s computer may be at risk of being compromised. In order to ensure no one other than the owner of the account has access to this information the account is stored in an encrypted, password-protected file. The user sets the username and password for an account when it is created. No one other than the owner of the account knows the password. The current implementation does not provide a password retrieval feature to help a user who has forgotten his or her account password. However, this type of feature would not be difficult to add since it could be built into the IMA Manager and use a simple question the user can setup when the account is created to

retrieve his or her password. The question could be something like “What is the name of your pet dog?”.

Even if an IMA participating business has excellent security measures in place to ensure each user’s personal information is protected, there may still be increased security risks since this architecture promotes a more open exchange of personal information between users and businesses. If someone tried to make a request to a business for personal information while posing as another user, this could lead to the business disclosing a user’s personal information to the wrong person. The way the IMA system attempts to reduce this risk is through the use of unique keys stored in each identity. When an identity is created, a unique key is generated and added to it. This unique key is provided to the business to authenticate an identity each time it is used. The key is never known to the user, but instead remains in the identity of the user account and is provided to the business along with the information in the identity. In order for someone to pose as another user to retrieve identity information from a business, the impostor would have to know the key for that user’s identity.

Another potential threat to the IMA system is that a business may not publicly state that it is participating yet may secretly deploy an IMA Web Service. Through this service the business may attempt to receive personal information from visitors to the business’s web site who are using the IMA Manager. The result is that a business may try to collect personal information from users, yet not provide the required access for users to stored personal information. For the IMA Manager to release an identity to a business this action must be initiated by the user. First the user must be currently visiting the business’s web site. Second the user must select an identity and attempt to associate it

with the business. Third the user is asked by the IMA Manager to confirm the sending of the identity. The IMA Manager will never automatically release information to a business. If the user confirms that the identity should be sent to the business, then only at this time is this action taken by the IMA Manager. The IMA Manager makes every effort to prevent a business from receiving identity information without full knowledge of the user.

CHAPTER 5
EVALUATION

In order to evaluate the success of the design of the IMA system, the two main design goals must be considered. First, does the IMA system provide greater access for a user to his or her personal information than existing systems do? Second, does the IMA system provide businesses that use it with greater compliance with certain areas of current privacy legislation than other systems do?

5.1 Access to Personal Information

This evaluation looks at how personal information is managed in the IMA system compared to .NET Passport, Liberty Alliance, and Infocards. This evaluation will be carried out through a detailed comparison of each system by examining the following criteria:

1. What information does the system handle? For example, does the system provide users with access to history information, such as what products the user has browsed while visiting the business's web site?
2. Where is each piece of information stored? For example, is the user's information such as mailing address, email address, phone number, etc., stored on the client, with a trusted third party, a business, or somewhere else?
3. Which pieces of information does the system allow the user to access? Does the system provide access only to basic account information, or does the system provide the user with access to information each individual business stores?

4. How does the system provide access to a user's personal information? Is access provided through a third party, another business, or directly from the business to the user?

5.1.1 Information Storage, Management, and Access Comparison

The IMA system not only offers many of the same features of other identity management systems, but also offers features other similar systems do not. The following is a summary of how each system handles storage, management and access to personal information, as compared to the IMA system.

1. Ability to edit existing user account information.
    a. .NET Passport – Yes

        Updated at a trusted third party. Microsoft provides access to edit a user account from its main Passport web site.

    b. Liberty Alliance – Yes

        Updated at the business where the account was initially created. In a Liberty Alliance COT, the trusted business users select to manage their account takes on the role of the trusted third party.

    c. Infocards – Yes

        Updated in Info-Card Interface provided in the Longhorn operating system.

    d. IMA – Yes

        Updated in IMA Manager on the user's computer.

2. Tracking of business to identity associations made.
    a. .NET Passport – No

Passport does not provide any ability for users to track which businesses they have visited using their account.

b. Liberty Alliance – No

Liberty Alliance neither provides any ability to track which businesses users have visited using their account nor which businesses in the COT they have not visited before.

c. Infocards – No

The Infocards system allows a user to specify specific uses for an identity and what information will be shared in an identity. At this time no indication has been given that would suggest it will allow associations between specific organizations to be made, tracked, and modified.

d. IMA – Yes

The IMA Manager allows a user to view a list of identity-to-business associations made for each identity to allow the user to track which businesses have been visited using a specific identity. It also allows the user to remove any associations made that the user no longer wishes to keep. When an association is deleted, the next time the user visits that business online the user will need to associate a new identity with that business.

3. Viewing of information associated with an identity at a specific business (e.g., history information).

a. .Net Passport – No

.Net Passport does not provide any access to information associated with a user's account either by Passport itself or by businesses where the account has been used.

b. Liberty Alliance – No

Liberty Alliance does not provide any access to the information a business has associated with the user's account. With Liberty Alliance a user chooses a trusted business to store his or her account information and act as a third party for authentication. Any information another business decides to associate with that user, will most likely not be shared with the trusted third party.

c. Infocards –No

With Infocards the user has the ability to assign uses to an identity and decide what information in that identity will be shared, but there is no ability to retrieve information a user has associated with an account.

d. IMA – Yes

The IMA Manager allows a user to view all items a business has associated with a given identity, such as history information the business has tracked about the user's activities at the business's web site that have been used to form a profile. The user has the ability to request this information and view it.

4. Removing of information associated with an identity at a specific business.

a. .Net Passport – No

Just as .Net Passport provides no ability to view the information a business has associated with an identity, it also provides no ability to remove information associated with an identity.

b. Liberty Alliance – No

Liberty alliance provides no ability to update or remove information a business has associated with a user's account.

c. Infocards – No

Infocards provide no access to update or remove information a business has associated with one of the user's identities.

d. IMA – Yes

IMA Manager allows a user to view profile information a business has associated with an identity and also allows the user to mark any items as removed and send the updated list back to the business. Marking an item "Remove" does not necessarily guarantee the business will eliminate the information. After all, information is valuable and the business may want to keep this information for research. The act of pointing out information to the business that is not accurate, allows the business to have a more accurate profile on the user. This can be used for features such as suggesting products of interest to the user, but with more accurate and up to date information that the user has corrected for the business.

5. The creation of multiple discrete identities.

a. .Net Passport – No

Passport only allows for the creation of a single set of personal information in one user account.

b. Liberty Alliance – No

Liberty Alliance only allows for the creation of a single set of personal information under one user account.

c. Infocards – Yes

Infocards allows for the creation of multiple discrete identities.

d. IMA –Yes

The IMA Manager allows a user to create multiple discrete identities inside a single user account, where each identity may be updated and used from the same username and password.

6. The ability to create associations of where an identity is used and attach these associations to identities.

a. .Net Passport – No

Passport provides no ability for users to explicitly create account to business associations that the user has control over to edit or remove.

b. Liberty Alliance – No

Liberty Alliance provides no ability for the user to create account to business associations that the user has control over to edit or remove.

c. Infocards – Yes

The Info-Card system allows the user to set specific types of uses for an identity (e.g., identity used for accessing government information web sites).

d. IMA –Yes

The IMA Manager allows the user to associate an identity with any participating business online. This identity is then used for all future visits until the user decides to change it.

7. Avoids reliance on third party storage for management of personal information.
    a. .Net Passport – No

    The basis of the .Net Passport system is the use of a trusted third party which is responsible for the management of all user accounts, authentication of users at participating sites, and for providing account information to participating businesses.

    b. Liberty Alliance –No

    A business the user trusts plays the role of a third party and handles the authentication of the user and provides the account to other businesses accessed by the user in the participating business's COT.

    c. Infocards – Yes

    All user account (identity) information is created and managed on the user's computer using the tools provided in Longhorn.

    d. IMA – Yes

    All user account (identity) information is managed locally on the user's computer using the IMA Manager application.

8. Allowing users to keep track of personal information that has been provided to an organization.
    a. .Net Passport – No

Passport does not provide a history that users can access showing them what businesses they have visited and what account information has been provided to each business.

b. Liberty Alliance – No

Liberty Alliance also does not provide this sort of functionality.

c. Infocards – Unknown

Infocards has not provided any indication this will be a feature offered in this system.

d. IMA – Yes

The IMA Manager allows the user to request from a business what history information the business knows about the user's identities.

9. When an identity is updated, pushes these changes automatically out to all businesses associated with that identity.

a. .Net Passport – No

Passport allows the user to update his or her user account information, but these changes are not immediately provided to businesses where the account has been used. This information is updated the next time the user logs into one of these businesses.

b. Liberty Alliance – No

Liberty Alliance provides these updates the next time the user visits a business.

c. Infocards – Unknown

At this time it is not stated in any information available that this will be a feature of Infocards.

d. IMA – Yes

The IMA Manager pushes any updates to an identity out to all participating businesses the identity has been associated with.

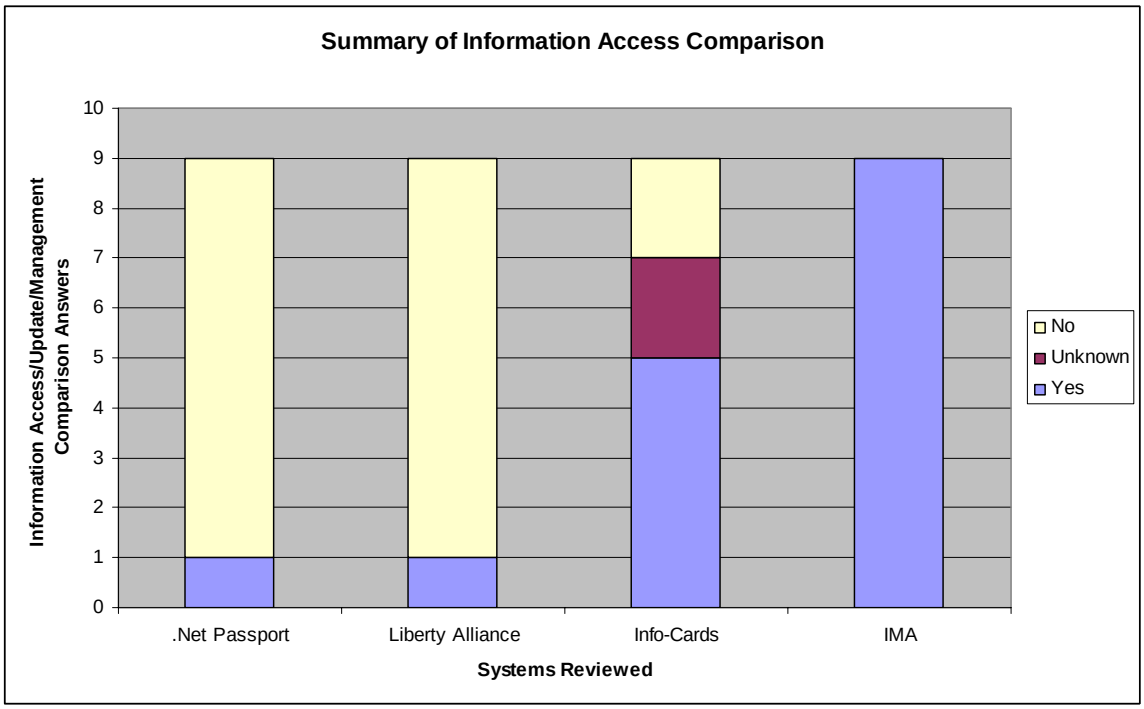

Figure 5.1: Summary of results from information access comparison of .Net Passport, Liberty Alliance, Infocards, and IMA

The graph shown in Figure 5.1 shows the summary of the access to information comparison of the three existing systems/specifications as compared to the IMA system. Since the IMA system offers many of the same features as the other SSO identity management systems discussed, the main focus for this comparison was to look at what information each system gives users the ability to access, update, and manage. One of the main goals in the development of the IMA system was to go beyond what existing single sign on services offer in terms of personal information management. This was done to

allow users to be more aware of who has their personal information and what other information has been associated with them.

As shown in Figure 5.1, both .Net Passport and Liberty Alliance allows a user to access only the basic identity information provided in his or her user account. Microsoft's Infocards system takes a similar approach to IMA in that it allows for client side management of personal information and for the use of multiple identities. However, based on what information is currently available about the features Infocards will offer, it provides only about half of the information access features looked at in the comparison. The IMA System was the only system to provide all information access features evaluated in the comparison.

5.2 Privacy Legislation Compliance

This evaluation examines the current privacy legislation (using PIPEDA as an example) and shows the areas of the legislation where .NET Passport, Liberty Alliance, Infocards, and IMA help to improve compliance. The evaluation was carried out by looking at the following points:

1. Breakdown of the ten privacy principles defined in PIPEDA and explaining which principles the various systems help support.
2. Showing what areas of the legislation where these systems can potentially provide compliance. Included in this discussion are the areas of the legislation where these systems do and do not provide compliance, as well as suggestions about how these systems could be adapted to improve compliance.

5.2.1 PIPEDA Compliance Comparison

The IMA system, as compared to other SSO identity management systems, does more than just provide users with increased access to their personal information, but also allows businesses to more easily comply with certain restrictions placed on businesses by privacy legislation. Of the ten principles defined in PIPEDA, SSO systems can directly help businesses improve PIPEDA compliance with five principles, which are consent, limiting collection, limiting use, accuracy, and access. The following is a discussion of each of these five PIPEDA rules along with a comparison of each tool [45].

1. Consent - must be obtained by an organization before it can collect, use or disclose a user's personal information.
    a. .NET Passport – Yes

        Passport obtains consent when a user uses a Passport account to login to a participating site. It is only after a user has logged in that personal information stored in the user's account is transferred to the participating site.

    b. Liberty Alliance – Yes

        Liberty Alliance also uses the moment a user logs in at another business in the same COT as the moment consent is obtained for account information to be provided to that business.

    c. Infocards – Yes

        Infocards allows a user to manage identities client side, so when a user provides an identity to an organization that is taken as consent.

    d. IMA – Yes

IMA also allows a user to manage identities client side, so when a user provides an identity that is considered consent.

2. Limiting Collection – of information to only that which is required, do not gather excess information without purpose or reason for doing so.
    a. .NET Passport – No

        Passport does not place any restrictions on personal information a business can gather. That is left up to the business to decide.

    b. Liberty Alliance – No

        Liberty Alliance also does not place any restrictions on personal information a business can gather. That is left up to each business in the COT to decide.

    c. Infocards – Yes

        Infocards allows a user to restrict what information in an identity is shareable with those the identity is provided to. This helps to restrict information given.

    d. IMA – Yes

        IMA Manager allows a user to specify which identity in his or her account is used at a given business; this is done to only provide the information necessary, by providing one identity, rather than all account information.

3. Limiting use, disclosure, and retention – only use information for the reason it was collected, only keep personal information until it is no longer needed, only

keep personal information that has been used to decide something about the person for a reasonable amount of time, remove data that is no longer relevant.

a. .NET Passport – No

   Passport does not provide any access to information a business has collected about the user that has been associated with a Passport account.

b. Liberty Alliance – No

   Liberty Alliance also does not provide any access to information a business has collected about the user.

c. Infocards – No

   Infocards also does not provide any access to information a business has collected about the user.

d. IMA – Yes

   IMA Manager allows a user to request to view any information the business has associated with his or her identity. This also allows a user to remove information that is no longer relevant.

4. Be accurate – ensure the accuracy of information when it is being used to make a decision about a user.

   a. .NET Passport – No

      Passport does not allow a user to check the accuracy of information that has been associated with the user's account.

   b. Liberty Alliance – No

Liberty Alliance also does not allow a user to check the accuracy of information that has been associated with the user's account.

c. Infocards – No

Infocards also does not allow a user to check the accuracy of information that has been associated with the user's account.

d. IMA – Yes

The IMA Manager allows a user not only to manage identity information on the client to ensure it is correct, but each time information in an identity is updated, these updates are automatically forwarded to all businesses to which this identity has been associated. This is done to ensure that all businesses knowing an identity have the most accurate and up to date information about that identity. Also when a user requests to view what information has been associated with an identity by a business, there is an option to remove any information that is not accurate.

5. Give individuals access – allow users to view their information upon request and correct any inaccurate information.

a. .NET Passport – No

Passport allows a user to access and correct information contained in his or her Passport user account, but nothing else. A user does not have the ability to access information gathered by any business the user has accessed with his or her Passport account.

b. Liberty Alliance – No

Liberty Alliance allows a user to access and update account information, but as far as accessing information stored by a business, this is not handled through the Liberty Alliance system.

c. Infocards – Partially

Infocards provides a user with more direct access to account/identity information since it is stored and managed locally on the user's computer, however this does not give a user access to information stored by businesses that a user's account/identity has been used at.

d. IMA – Yes

IMA Manager stores a list of all identities used at each business and allows a user to request and view any information a business has associated with the identity and remove any information that is not accurate. Also, any changes made to an identity are automatically forwarded to all businesses where the identity has been used.

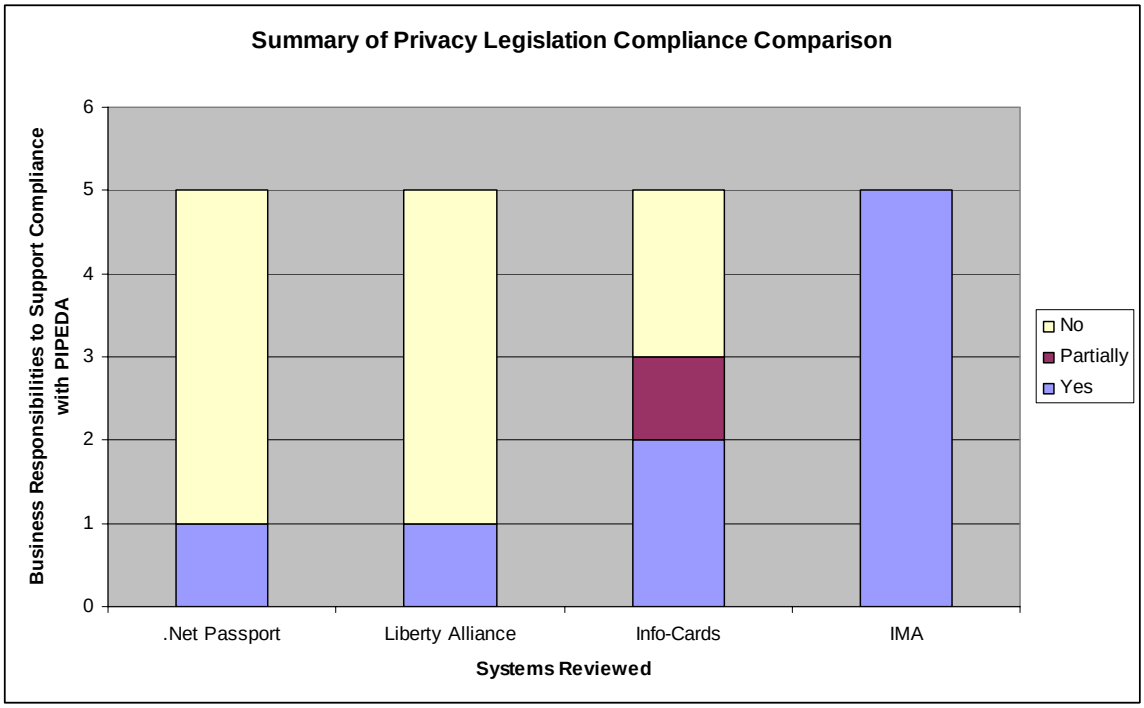

Figure 5.2: Summary of results from privacy legislation compliance comparison of .Net Passport, Liberty Alliance, Infocards, and IMA

5.2.2 Privacy Legislation Compliance in Other Countries

The IMA system can provide benefits to both users and businesses in more than just Canada where PIPEDA applies. Privacy legislation in other countries also provides citizens with similar rights and protections as PIPEDA on how a business may use someone's personal information. One example is Europe's Data Protection Act (DPA) [24].

The evaluation in section 5.2.1 discusses how the IMA system provides benefits for users and helps improve business compliance with some of the main rights of users as defined by PIPEDA. However these rights are not ideas defined only for Canadian privacy legislation, but are common to rights defined in privacy legislation in other countries as well. The DPA written by the Information Commissioner's Office of the United Kingdom defines the obligations a business must meet when handling personal

information. DPA defines eight best practices that businesses must abide by when handling users' personal information and also defines seven rights that people have when accessing businesses about how information is gathered, processed, and used.

The eight best practices for businesses defined by DPA are [25]:

1. Data must be processed fairly and lawfully. For data to be considered fairly processed it must meet one of six possible conditions, the first of which is:
    a. Consent must be obtained from the individual before his or her personal information can be processed.
2. Data can only be processed for limited purposes.
3. Processing of all data must be adequate, relevant and not excessive.
4. Data must be accurate and up-to-date.
5. No data should be kept for longer than necessary.
6. Data should only be processed in accordance with the individual's rights.
7. Data should be handled and processed in a secure way.
8. Data should not be transferred to countries outside the European Economic area unless the country has adequate personal information protection legislation.

The seven rights people have under DPA are [25]:

1. Right to access – people are allowed to find out what personal information the business has gathered about them.
2. Right to prevent processing – people can access that information a business has stored about them and request that it not be processed.
3. Right to prevent processing for direct marketing.

4. Right to object to automated decisions made about them based on their personal information.
5. Right to compensation for damages as a result of non compliance with these rules.
6. Right to rectify, block, or destroy personal information a business has stored which includes information that is inaccurate.
7. Right to ask the Commissioner to determine if a violation has been committed.

Of the eight best practices and seven rights defined in DPA, like PIPEDA, there are five practices/rights in DPA with which the IMA system can help a business improve its compliance. Since these practices/rights are so similar to the ones defined in PIPEDA, a comparison of .NET Passport, Liberty Alliance, Infocards, and IMA under each of these practices/rights would simply be a repeat of the PIPEDA comparison. Instead a brief summary of the similarities between PIPEDA and DPA is as follows.

PIPEDA	**DPA**
1.Consent must be obtained before data is collected or used.	1.Data must be fairly and lawfully processed (e.g., consent must be obtained).
2.Limiting collection of information only to what is required.	2. Processing of all data must be adequate, relevant and not excessive.
3.Limiting use, disclosure, and retention.	3.Data can only be used for limited purposes. No data should be kept for longer than necessary. Right to prevent processing for direct marketing.
4.Ensure accuracy of information used.	4.Data must be accurate and up-to-date.
5.Give individuals access to view their information upon request and correct any inaccurate information.	5.Right to access personal information a business has gathered about them.

Of the best practices and rights defined under DPA, there are three that DPA defines which are not covered under PIPEDA. These three practices are: the right to prevent processing of personal information for direct marketing, the right to object to automated decisions based on personal information gathered, and the right that for data to not be transferred to countries outside the European Economic area. The IMA system does not aid businesses with compliance with any of these three practices/rights. However, this is not important in terms of the IMA system, since these three practices/rights are areas which would not be covered by a SSO system.

5.3 Summary

The main goals of the IMA system were to improve a user's awareness of and access to his or her personal information and help businesses to more easily comply with some of the restrictions placed on businesses by privacy legislation. The information access evaluation which compared .NET Passport, Liberty Alliance, Infocards, and IMA showed that IMA system was able to meet its goal of being a SSO system that does not rely on the use of a trusted third party, yet still provides a greater ability to restrict, manage, and access personal information than other current systems do. Through the privacy legislation business compliance evaluation it was shown that the IMA system provided greater compliance for businesses by comparing the four identity management tools under privacy legislations such as PIPEDA and DPA. This is due in part to the IMA system not using a trusted third party, so no third party knowledge of storage of information is necessary, and due in part to improving users' access to their own personal information. This helps businesses more easily meet the requirements of PIPEDA and DPA.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

Even with the tools currently available for Internet users to protect their privacy on-line, there are few useful tools to protect privacy through personal information management. SSO systems such as Passport provide some basic management of personal information, but lack flexibility and the ability for users to remain anonymous. These systems also require people to disclose their personal information to a third party to be able to participate and do not allow users to keep track of businesses that have been given their personal information. P3P doesn't offer any real management of personal information and places the responsibility entirely on users to understand the P3P interface in order to protect their privacy. With P3P people must understand from their preferences what information is recorded and how it is used, without being able to view or control that information.

The IMA system is a useful addition to the areas of on-line privacy where current technologies are unable to offer more control for users over their personal information. This thesis focused on looking at problems with existing well known SSO identity management systems such as .NET Passport and Liberty Alliance and using this information as a the basis for the design of a new architecture. The goals for the design of this architecture were to help improve the privacy of an Internet user by eliminating the need for any party, other than the user and a business he or she wishes to access, from needing to possess the user's personal information, while at the same time providing the user with greater knowledge of and access to personal information stored by businesses.

Following these basic requirements the IMA system was developed as a SSO system that does not use a trusted third party, allows for the use of multiple identities from a single user account, and provides the user with the ability to view and update personal information a business has collected about the user.

6.1 Contributions

The IMA system provides a design for a personal information management architecture that is offered as an alternative to .NET Passport and Liberty Alliance. It is hoped that this work will be the basis for more research into identity management systems (i.e., SSO systems) that provide users with more control over their own personal information, while allowing businesses to more easily comply with privacy legislation. The main contributions of this project are:

1. Lack of reliance on a third party for account storage and management
    a. The IMA system is designed around the basic business-client relationship and does not require the participation of any other party. This is accomplished by allowing client side management and storage of identities, which allows for information to need to exist only in two places: the client's computer and the participating business the user is accessing.
2. The use of multiple discrete identities all managed from a single user account
    a. In the IMA system the user has the ability to create and manage more than one identity with each containing a different set of personal information, (i.e., personal identity, work identity, etc.) while at the same time allowing

all identities to be stored in a single user account, and to all be accessed by a single username and password.

3. The ability to create identity-to-business associations
    a. The IMA system allows a user to associate an identity to a business and have that identity used for all future visits, yet still allow the user to at any time be able to change to another identity when the purpose for accessing the business's website changes. These associations are all stored in the user's account which only the user can access. This prevents anyone, other than the account owner, from knowing about the range of identities the user may have or which businesses are associated with which identity.
4. Disclosure, correction, and removal of personal information stored by a business
    a. The IMA system gives the user the ability to have access to the identity information the user has provided to a business, and also provides access to the information the business has associated with a user's identity. The IMA system allows a user to request that any participating business provide a list of information the business has associated with an identity (i.e., such as the user's preferences for certain products) and allows the user to update, correct, add, or remove information as appropriate.
5. Compliance for businesses with privacy legislation disclosure requirements
    a. The IMA system also helps provide participating businesses with improved compliance with information disclosure rules placed on businesses by privacy legislation. Under privacy legislation such as PIPEDA a customer has the right to request that a business disclose what

information the business has collected about the customer, and must allow the customer to see this information within a reasonable amount of time.

b. The business must allow the customer to remove or correct any inaccurate information. Since the IMA system allows the user to manage and update identity information, this helps businesses to insure this identity information is up to date and correct. Also, since the IMA system allows users to retrieve and update information associated with an identity, this helps businesses more easily deal with information disclosure rules placed on them by privacy legislation.

6.2 Future Work

The next step for the IMA system will be to build a larger, more complete implementation. Before that can be accomplished there are several areas that will need to be explored further in order to make a larger scale implementation of the IMA system more practical. Issues that will need to be addressed further are security issues with both the client side and business side of the system; implementation of features discussed including personas, account access from multiple locations, and lost password retrieval; and finally a more comprehensive evaluation of the IMA system's design,.

Security of a user's personal information needs to be handled carefully in a system such as this one. A user's account information could fall into the hands of a malicious user who would not only be able to gather a great deal of information about the user, but would also be able to access participating businesses while pretending to be someone else. The first security issue that will need to be addressed is to determine the best way to store the encrypted XML user account (i.e., on the user's machine, on a

remote storage system, etc.). The current implementation uses the .NET XML encryption libraries provided, which is fairly secure, however more research into secure storage of a user's account should be explored. The second security issue to look at is a better way for each business to securely identify each identity to prevent unauthorized access to a user's identity and profile information. At present the unique key each identity uses is just a simple GUID; however, something more secure than this may need to be used. If someone else was able to get this GUID he or she could create a new identity with this GUID and use it to access a business as that person. The third security issue that will need to be solved will be coming up with a way to ensure a participating business is actually participating and is not posing as a participating business. A possible way to prevent fake businesses from being setup for the purpose of leeching user information is to have a third party who governs the use of the IMA system and authenticates participating businesses. This third party would handle no user personal information of any kind, but would rather provide a service for an IMA client to check a business for participation when the business's site is visited.

One feature that may be implemented in future versions of the IMA system is Personas. A Persona is basically a subset of personal information that can be shared among several identities. For example, a user may have a "Personal" identity to use at a site for shopping for products for him or herself, but may also create a "Shopping" identity for when the user is shopping for gifts for friends and family. Both of these identities will have different profile information associated with them by the participating business, but will still have the same set of personal information. A Persona is basically a way to store a set of personal information that can be used by more than one identity, but

only needs to be entered once and can be updated in one place. Some information about the design of Personas was provided in this thesis, but was not implemented in the current implementation of the IMA system.

The current implementation of the IMA system makes no attempt to make it easy for a user to use accounts from more than one computer. Offering this feature will need to be addressed so that a user will be able to more easily access accounts from multiple locations. Since the IMA system does not rely on a third party system for account access and storage, multi-location access will require a different approach. There are several options available. One option is to store the user's encrypted XML account file with a single business the user trusts and from which the account may be retrieved and used at different locations. Another option is to have a third party system that stores encrypted XML account files, but has no knowledge of the contents of the file, but simply provides a storage and retrieval system for accounts.

In a typical SSO system such as .NET Passport, retrieval of a lost password is handled by a service provided by the third party that authenticates the user based on some other information. Since the IMA system has no third party to rely on for password retrieval, this can be handled through the client application. When a user creates a user account, there could be a simple password retrieval question (e.g., "What is my dog's name?") the user enters, associated with an answer only the user knows (e.g., "Spike"), which can then be used to retrieve a password.

The evaluation done for the IMA system was scaled back significantly from what was originally planned, mainly due to lack of time and resources. Once a more complete implementation of the system has been built and security issues have been addressed, a

more extensive evaluation of the design will need to take place. This evaluation will have to take place in two parts: the client evaluation and the business evaluation. For the client evaluation a user study will need to take place. The intent of the IMA system is not to win over people, who are not interested in identity management SSO systems, but to target users of current SSO systems such as .NET Passport. A group of users of other SSO identity management systems will need to be assembled to provide a basis for the client evaluation study. The study would take place by having each user use the IMA client application with a series of example participating businesses and have each user follow a series of tasks to perform, such as associating identities to businesses, retrieving profile information, removing profile information, and updating identity information. Following these exercises, each user will answer questions such as how intuitive were the IMA client toolbar and manager interfaces; how easy is the system to use when managing identities, associations, and profile information; and how useful did the user find the IMA client in managing personal information as compared to the SSO systems each user is currently using. To evaluate the business side of the IMA system businesses will need to be consulted. Preferably businesses who already participate in some SSO type system would be used. The business would need to be consulted about how difficult it would be for the business to implement the IMA Web Service with its present system, and about the business's concerns about security and management of the IMA system. Finally, businesses would need to be questioned about disclosing personal information to users and what the businesses like and dislike about this system and the potential problems that may arise.

At the moment, the current design of the IMA system mainly looks at business web sites that manage user profiles. Future work may include looking at ways to combine P3P and IMA into a single privacy policy compliance and identity management SSO system. Also a topic for a future paper may be to look at how existing systems, such as .NET Passport and Liberty Alliance, may be modified to incorporate features such as disclosure on demand and the use of multiple identities.

6.3 Conclusions

The IMA system offers a design for a personal information management system that is provided as an alternative to .NET Passport and Liberty Alliance. The main purpose of this research work has been to demonstrate the benefits of a SSO system that offers increased access for users to their personal information while allowing users to maintain more than one identity. The main contribution of this project has been the design of a SSO identity management system that does not use a third-party or require businesses to transfer identity information from one business to another. It is hoped that this work will be the basis for more research into identity management systems (i.e., SSO systems) that provide users with more control over their personal information, that allow businesses to increase their compliance with privacy legislation, and thus improve the privacy of Internet users.

REFERENCES

[1] M. Ackerman and L. Cranor, "Privacy critics: UI components to safeguard users' privacy" Conference on Human Factors in Computing Systems (CHI '99), ACM Press, 1999, pp. 258-259.

[2] M. Ackerman, L. Cranor, and J. Reagle, "Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences" Proceedings of the ACM Conference on Electronic Commerce, 1999, pp. 1-8.

[3] Amazon.ca, "Privacy Notice" 2003; http://www.amazon.ca/exec/obidos/tg/browse/-/918814/702-9416203-2498452.

[4] K. Aschenbrenner, "Implement Secure .NET Web Services with WS-Security" May 2003; http://www.devx.com/security/Article/15634/0/page/1.

[5] P. Ashley et al., "Enterprise Privacy Authorization Language (EPAL 1.1)" 2003; http://www.zurich.ibm.com/security/enterprise-privacy/epal/Specification/index.html.

[6] D. Blum, "Federated identity gets a boost" Oct. 2004; http://www.nwfusion.com/columnists/2004/101104blum.html.

[7] J. Borking, "Privacy Incorporated Software Agent (PISA) System Architecture" 2001; [1]http://www.pet-pisa.nl/dscgi/ds.py/Get/File-237/Report_PISA_Architecture_12122002-v7.pdf.

[8] P. Boutin, "Just How Trusty Is Truste?" Apr. 2002; http://www.wired.com/news/exec/0,1370,51624,00.html.

[9] S. Cantor et al., "Liberty ID-FF Architecture Overview" 2003; http://www.projectliberty.org/specs/liberty-idff-arch-overview-v1.2.pdf.

[10] D. Carr, "What's Federated Identity Management?" Nov. 2003; http://www.eweek.com/article2/0,4149,1378436,00.asp.

[11] CBC News Online, "New privacy law requires client's consent" Mar. 2001; http://www.cbc.ca/stories/2000/12/14/Consumers/privacy_law001214.

[12] Chapters.Indigo.ca. Privacy Policy. Available online at http://www.projectliberty.org/specs/liberty-idff-arch-overview-v1.2.pdf.

[13] P. Connolly, "Who are you? Multiple Personalities are a reality that identity management schemes must address" July 2002; http://www.infoworld.com/article/02/07/26/020729opsecurity_1.html.

[14] Consumer Privacy Guide, "What is TRUSTe and how does it work to protect my privacy?"; http://www.consumerprivacyguide.org/faq/truste.shtml.

[15] Cover Pages, "Security Assertion Markup Language (SAML)" Apr. 2005; http://xml.coverpages.org/saml.html.

[16] J. Evers, "EBay Cancels Its Passport" Jan. 2005; http://www.pcworld.com/news/article/0,aid,119137,00.asp.

[17] Federal Trade Commission, "Privacy Online: Fair Information Practices In the Electronic Marketplace" May 2000; http://www.ftc.gov/os/2000/05/testimonyprivacy.htm.

[18] FIDIS, "Future of Identity in the Information Society"; http://www.fidis.net/.

[19] GetNetWise, "Guide to Internet Terms: Privacy Policy"; http://www.getnetwise.org/glossary.php#P.

[20] G. Gross, "Privacy: Internet users want to have cake and eat it shock" May 2003; http://www.techworld.com/news/index.cfm?fuseaction=displaynews&NewsID=85.

[21] R. Guth, "Microsoft Tests Software to Fight Identity Theft on Web" Mar. 2005; http://everybreathdeathdefying.com/blog/archives/000266.html.

[22] IBM, "Specification: Web Services Security (WS-Security)" Apr. 2002; http://www-106.ibm.com/developerworks/webservices/library/ws-secure/.

[23] IBM, "Enterprise privacy architecture (EPA)" 2004; http://www.zurich.ibm.com/pri/projects/epa.html.

[24] Information Commissioner, "Data Protection Act 1998" 1998; http://www.opsi.gov.uk/acts/acts1998/19980029.htm.

[25] Information Commissioner, "Data Protection Act Factsheet" 1998; http://www.informationcommissioner.gov.uk/cms/DocumentUploads/Data%20Protection%20Act%20Fact%20V2.pdf.

[26] S. Johnston, "Pondering Passport: Do You Trust Microsoft With Your Data?" Sept.

2001; http://www.pcworld.com/news/article/0,aid,63244,00.asp.

[27] J. Kobielus, "SAML promises Web services security" July 2002; http://www.nwfusion.com/news/tech/2002/0701tech.html.

[28] L. Korba and C. Beaudoin, "Privacy Incorporated Software Agent (PISA)" Apr. 2005; http://iit-iti.nrc-cnrc.gc.ca/projects-projets/pisa_e.html.

[29] Liberty Alliance Project, "Liberty Alliance Project"; http://www.projectliberty.org/.

[30] S. McClure and J. Scambray, "Platform for Privacy Preferences strives for user control, but its sanctions lack teeth" Oct. 1999; http://www.infoworld.com/cgi-bin/displayArchive.pl?/99/40/o03-40.58.htm.

[31] D. McCullagh, "Is TRUSTe Trustworthy?" Nov. 1999; http://www.wired.com/news/politics/0,1283,32329,00.html.

[32] B. McWilliams, "Stealing MS Passport's Wallet" Nov. 2001; http://www.wired.com/news/print/0,1294,48105,00.html.

[33] J. Menn, "Microsoft's Passport fails to travel far as Web strategy" Dec. 2004 ; http://seattletimes.nwsource.com/html/businesstechnology/2002136272_passport31.html.

[34] Microsoft, "Microsoft .NET Framework Developer Center"; http://msdn.microsoft.com/netframework/.

[35] Microsoft, ".NET Passport"; http://www.passport.net/.

[36] Microsoft, "Web Services Developer Center"; http://msdn.microsoft.com/webservices/building/wse/.

[37] Microsoft, "Microsoft .NET Passport Privacy Statement" 2003; http://www.projectliberty.org/specs/liberty-idff-arch-overview-v1.2.pdf.

[38] Microsoft, "INFO: Post-Service Pack 2 .NET Framework Core CLR Hotfix Package" June 2004; http://support.microsoft.com/kb/328544/EN-US/.

[39] Microsoft, "BUG: The BeforeNavigate2 Event of WebBrowser Control Does Not Fire" July 2004; http://support.microsoft.com/kb/327135/EN-US/.

[40] E. Norlin and A. Durand, "Towards Federated Identity Managment" Aug. 2004; http://discuss.andredurand.com/stories/storyReader$320.

[41] PRIME Project, "Privacy and Identity Management for Europe"; http://www.prime-project.eu.org/.

[42] Privacy Commissioner of Canada, "Findings on Air Canada's Aeroplan Frequent Flyer Program under the Personal Information Protection and Electronic Documents Act" Mar. 2002; http://www.privcom.gc.ca/media/nr-c/02_05_b_020320_e.asp.

[43] Privacy Commissioner of Canada, "A Guide for Individuals" Nov. 2003; http://www.privcom.gc.ca/information/02_05_d_08_e.asp.

[44] Privacy Commissioner of Canada, "The Personal Information and Electronic Documents Act: A Primer On its Privacy Provisions" Dec. 2003; http://e-com.ic.gc.ca/epic/internet/inecic-ceac.nsf/en/gv00213e.html.

[45] Privacy Commissioner of Canada, "A Guide for Businesses and Organizations" Apr. 2004; http://www.privcom.gc.ca/information/guide_e.asp.

[46] B. Richardson and J. Greer, "An Architecture for Identity Management" 2nd Annual Conference on Privacy, Security and Trust (PST04), Privacy, Security and Trust 2004 , 2004, pp. 103-108.

[47] P. Roberts, "Liberty Alliance Explains Its Sign-On Services" Feb. 2003; http://www.pcworld.com/news/article/0,aid,109277,00.asp.

[48] S. Spiekermann, J. Grossklags, and B. Berendt, "E-privacy in 2nd Generation ECommerce: Privacy Preferences versus actual Behavior In Electronic Commerce (EC'01)" 3rd ACM Conference on Electronic Commerce, vol. EC '01, 2002, p. 38--47.

[49] Synomos, "IBM Announces an Enterprise Privacy Architecture Protecting Brand, Building Customer Trust and Adding Business Value Through Privacy Protection" June 2001; http://www.synomos.com/html/news/pr01-06-29.html.

[50] D. Ticoll, "Companies ignore privacy laws at their peril" Aug. 2003; http://www.globetechnology.com/servlet/ArticleNews/TPStory/LAC/20030814/TWTICO14/TPTechInvestor/.

[51] TRUSTe, "TRUSTe"; http://www.truste.org/.

[52] USA Today, "New U.S. privacy laws needed to protect Internet consumers" June 2003; http://www.usatoday.com/tech/news/internetprivacy/2003-06-25-privacy-policy_x.htm.

[53] W3C, "SOAP Version 1.2 Part 1: Messaging Framework" June 2003; http://www.w3.org/TR/soap12-part1/.

[54] WebServices.Org, "IBM Introduces New Language to Automate Privacy Compliance" July 2003; http://www.webservices.org/index.php/ws/content/view/full/3181.

[55] World Wide Web Consortium (W3C), "P3P 1.0: A New Standard in Online Privacy" 2003; http://www.w3.org/P3P/brochure.html.

[56] World Wide Web Consortium (W3C), "XML Schema Requirements" 1999; http://www.w3.org/TR/1999/NOTE-xml-schema-req-19990215.

[57] World Wide Web Consortium (W3C), "Extensible Markup Language (XML) 1.0 (Third Edition)" Feb. 2004; http://www.w3.org/TR/REC-xml/.

[58] WP 14.1, "Framework V1" Mar. 2005; http://www.prime-project.eu.org/public/prime_products/deliverables/fmwk/pub_del_D14.1.a_ec_wp14.1_v1_final.pdf.

[59] WP3, "Structured Overview on Prototypes and Concepts of Identity Management Systems" 2004; http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.1.overview_on_IMS.pdf.

[60] R. Yasin, "What is Identity Management?" Apr. 2002; http://infosecuritymag.techtarget.com/2002/apr/cover_casestudy.shtml.

[61] Yee, G. and Korba, L. Privacy Policy Complaince for Web Services. In Proceedings of the IEEE International Conference on Web Services. San Diego, California, USA.

[62] P. Zolnikov, "Extending Explorer with Band Objects using .NET and Windows Forms" Apr. 2002; http://www.codeproject.com/csharp/dotnetbandobjects.asp?print=true.

APPENDIX A

EXAMPLE IE TOOLBAR SETUP

Setting up a simple Hello World example IE Toolbar:

Download an example project and follow the instructions at:

http://www.codeproject.com/csharp/dotnetbandobjects.asp?print=true

The example solution available from codeproject.com has two projects:

SampleBars

--Has an "Hello World" example ToolBar and ExplorerBar

Register

--Registers the project with "Gacutil" and "Regasm"

--Remember that this project must be run last in the build process

--Also the Visual Studio solution configuration must be set to "Release"

Build → Configuration Manager → Release

NOTE: The following changes will need to be made to the Windows PATH parameter so that the "Register" project can find the "gacutil.exe" and "regasm.exe" utilities:

Need to set the system path for the "gacutil.exe" and "regasm.exe" to allow the "Register" project in the code downloaded to register the toolbar with IE.

Go to System→ Environment Variables

Append to PATH (for gacutil.exe)

C:\Program Files\Microsoft Visual Studio .NET 2003\SDK\v1.1\Bin

Append to PATH (for regasm.exe)

C:\WINDOWS\Microsoft.NET\Framework\v1.1.4322

Open the project in Visual Studio and “Rebuild Solution”

To View the newly created toolbar:

--Restart Internet Explorer

--Go to the menu options

View → Toolbars → Hello World Bar

(Starts the example toolbar)

View → Explorer Bar → Hello World Bar

(Starts the example toolbar)

APPENDIX B

IE TOOLBAR LIBRARIES SETUP

In order to create a toolbar that can be used with IE, it is necessary to setup the required libraries. The following is a set of instructions provided to explain how the required libraries are created. Also a known bug with the .NET Framework that affects the IMA client application is explained.

Create the dependent libraries with a strong name:

1. Create the signed key
    a. Sn –k myKey.snk
2. Create both required libraries with strong names (Produces the AxSHDocVw and SHDocVw dll files with a strongly typed name)
    a. Aximp c:\windows\system32\shdocvw.dll /keyfile:mykey
3. Put the assemblies in the Global Assembly Cache (Gac) using GacUtil
    a. gacutil /if Interop.SHDocVw.dll

References:

http://www.windowsforms.net/Forums/ShowPost.aspx?tabIndex=1&tabId=41&PostID=16003

http://www.csharpfriends.com/quickstart/howto/doc/Interop/Building_Samples_NET2COM.aspx

BeforeNavigate Bug in .Net Framework

This bug affects the BeforeNavigate event handler in the .NET Framework that listens for this event being raised in IE when the URL in the address bar changes.

Here is a link to an example with the fix:

http://support.microsoft.com/default.aspx?scid=kb;EN-US;325079

APPENDIX C

USER ACCOUNT SCHEMA

The user account schema defines all items required for a user account which includes the basic account information such as username and password, stores all identities the user has created, and stores all associations the user has made between his or her identities and participating businesses.

```
<?xml version="1.0" ?>
<xs:schema targetNamespace="http://tempuri.org/user.xsd">
   <xs:element name="user" type="userType" />
   <xs:complexType name="userType">
      <xs:sequence>
         <xs:element name="username" type="xs:string" />
         <xs:element name="password" type="xs:string" />
               <xs:element name="defaultid" type="xs:string" />
               <xs:element name="identities" type="identitiesType" />
         <xs:element name="associations" type="associationsType" />
      </xs:sequence>
   </xs:complexType>
   <xs:complexType name="identitiesType">
      <xs:sequence>
         <xs:element name="identity" type="identityType"
                     maxOccurs="unbounded" />
      </xs:sequence>
   </xs:complexType>
   <xs:complexType name="identityType">
      <xs:sequence>
            <xs:element name="key" type="xs:string" />
            <xs:element name="id" type="xs:string" />
            <xs:element name="firstname" type="xs:string" minOccurs="0" />
            <xs:element name="lastname" type="xs:string" minOccurs="0" />
            <xs:element name="email" type="xs:string" minOccurs="0" />
            <xs:element name="company" type="xs:string" minOccurs="0" />
            <xs:element name="address" type="xs:string" minOccurs="0" />
            <xs:element name="city" type="xs:string" minOccurs="0" />
            <xs:element name="state-prov" type="xs:string" minOccurs="0" />
            <xs:element name="postalcode" type="xs:string" minOccurs="0" />
            <xs:element name="phone" type="xs:string" minOccurs="0" />
      </xs:sequence>
   </xs:complexType>
   <xs:complexType name="associationsType">
      <xs:sequence>
         <xs:element name="association" type="associationType" minOccurs="0"
                     maxOccurs="unbounded" />
```

```
      </xs:sequence>
    </xs:complexType>
    <xs:complexType name="associationType">
      <xs:sequence>
          <xs:element name="url" type="xs:string" />
          <xs:element name="activeid" type="xs:string" />
          <xs:element name="id" type="xs:string"
                      minOccurs="0" maxOccurs="unbounded" />
       </xs:sequence>
   </xs:complexType>
</xs:schema>
```

APPENDIX D

EXAMPLE USER ACCOUNT XML

Provided here is an example account XML file that would be produced by the IMA Manager for a user based on the user account schema.

```
<?xml version="1.0" encoding="utf-8" ?>
<user xmlns="http://tempur_u46 ?org/user.xsd">
    <username>jsmith</username>
     <password>abc123</password>
      <defaultid>Anonymous</defaultid>
       <identities>
          <identity>
               <key>Generated by some hash function</key>
                <id>Anonymous</id>
         </identity>
         <identity>
                <key>Generated by some hash function</key>
                 <id>Personal</id>
                 <firstname>John</firstname>
                 <lastname>Smith</lastname>
                 <email>jsmith@hotmail.com</email>
                  <company>none</company>
                  <address>234 Queen St.</address>
                  <city>Toronto</city>
                   <state-prov>Ontario</state-prov>
                   <postalcode>e5t3f5</postalcode>
                   <phone>434-344-2344</phone>
            </identity>
            <identity>
                   <key>Generated by some hash function</key>
                   <id>Work</id>
                   <firstname>John</firstname>
                   <lastname>Smith</lastname>
                   <email>jsmith@mycompany.com</email>
                   <company>mycompany</company>
                   <address>2313 York St.</address>
                   <city>Toronto</city>
                   <state-prov>Ontario</state-prov>
                   <postalcode>e3r6t4</postalcode>
                   <phone>434-756-8767</phone>
             </identity>
        </identities>
        <associations>
          <association>
                <url>www.amazon.ca</url>
                <activeid>Personal</activeid>
                <id>Anonymous</id>
```

```
            </association>
             <association>
                    <url>www.chapters.indigo.ca</url>
                    <activeid>Personal</activeid>
             </association>
             <association>
                     <url>www.ncix.com</url>
                     <activeid>Work</activeid>
                      <id>Anonymous</id>
              </association>
              <association>
                      <url>www.dell.ca</url>
                      <activeid>Work</activeid>
               </association>
           </associations>
</user>
```

APPENDIX E

REPORT SCHEMA

The report schema defines the XML document that would be returned by a participating business when a user has made a request for information the business has associated with an identity the user has used at the business's web site. The report schema contains information about the identity used, the business the identity has been used at, and the information the business has associated with that identity.

```
<?xml version="1.0" ?>
<xs:schema id="report" targetNamespace="http://tempuri.org/report.xsd">
   <xs:element name="report" type="reportType"/>
      <xs:complexType name="reportType">
         <xs:sequence>
        <xs:element name="business" type="businessType"/>
        <xs:element name="identity" type="identityType"/>
         </xs:sequence>
      </xs:complexType>
      <xs:complexType name="businessType">
         <xs:sequence>
        <xs:element name="name" type="xs:string"/>
        <xs:element name="url" type="xs:string"/>
        <xs:element name="email" type="xs:string"/>
        <xs:element name="phone" type="xs:string"/>
        <xs:element name="disclaimer" type="xs:string"/>
         </xs:sequence>
       </xs:complexType>
        <xs:complexType name="identityType">
          <xs:sequence>
               <xs:element name="key" type="xs:string"/>
               <xs:element name="id" type="xs:string"/>
               <xs:element name="profile" type="profileType"/>
          </xs:sequence>
        </xs:complexType>
        <xs:complexType name="profileType">
        <xs:sequence>
              <xs:element name="dataitem" type="dataItemType" minOccurs="0"
              maxOccurs="unbounded"/>
        </xs:sequence>
         </xs:complexType>
         <xs:complexType name="dataItemType">
        <xs:sequence>
           <xs:element name="id" type="xs:int"/>
           <xs:element name="media" type="xs:string"/>
```

```
            <xs:element name="title" type="xs:string"/>
            <xs:element name="category" type="xs:string"/>
            <xs:element name="subject" type="xs:string"/>
            <xs:element name="association" type="xs:string"/>
                  <xs:element name="remove" type="xs:boolean" />
         </xs:sequence>
         </xs:complexType>
</xs:schema>
```

APPENDIX F

EXAMPLE REPORT XML

Provided here is an example report XML file that would be returned from a participating business when a request for information associated with an identity had been made by the user.

```
<?xml version="1.0" encoding="utf-8"?>
<report xmlns="http://tempur_u46 ?org/report.xsd">
  <business>
    <name>Programming Books</name>
    <url>www.programmingbooks.com</url>
    <email>customersupport@programmingbooks.com</email>
    <phone>555-555-1234</phone>
    <disclaimer>
        If you have any questions or concerns about this information
        please contact us directly by using the email provided with the
         subject "Privacy Question".
     </disclaimer>
  </business>
  <identity>
      <key>Generated by some hash function</key>
      <id>Personal</id>
      <profile>
          <dataitem>
              <id>1232</id>
              <media>book</media>
              <title>Programming C#</title>
              <category>programming</category>
              <subject>C#</subject>
              <association>purchased</association>
                  <remove>false</remove>
          </dataitem>
          <dataitem>
               <id>3212</id>
               <media>book</media>
                    <title>Thinking in Java</title>
               <category>programming</category>
               <subject>Java</subject>
               <association>viewed</association>
                    <remove>false</remove>
          </dataitem>
           <dataitem>
                    <id>2343</id>
                    <media>software</media>
                    <title>JBuilder 9.0</title>
```

```
                    <category>IDE</category>
                    <subject>Java programming environment</subject>
                    <association>purchased</association>
                    <remove>false</remove>
          </dataitem>
           </profile>
    </identity>
</report>
```

APPENDIX G

IDENTITY SCHEMA

The identity schema defines the structure of a single identity that a user creates and adds to his or her account and provides to a participating business. A user account stores a list of identities.

```
<?xml version="1.0" encoding="utf-8" ?>
<xs:schema targetNamespace="http://tempuri.org/XMLSchema.xsd" >
    <xs:element name="identity" type="identityType" />
    <xs:complexType name="identityType">
          <xs:sequence>
        <xs:element name="key" type="xs:string" />
        <xs:element name="id" type="xs:string" />
        <xs:element name="firstname" type="xs:string" minOccurs="0" />
        <xs:element name="lastname" type="xs:string" minOccurs="0" />
        <xs:element name="email" type="xs:string" minOccurs="0" />
        <xs:element name="company" type="xs:string" minOccurs="0" />
        <xs:element name="address" type="xs:string" minOccurs="0" />
        <xs:element name="city" type="xs:string" minOccurs="0" />
        <xs:element name="state-prov" type="xs:string" minOccurs="0" />
        <xs:element name="postalcode" type="xs:string" minOccurs="0" />
        <xs:element name="phone" type="xs:string" minOccurs="0" />
          </xs:sequence>
    </xs:complexType>
</xs:schema>
```

APPENDIX H

EXAMPLE IDENTITY XML

Provided here is an example identity XML file that would be provided to a participating business when a user decides to associate an identity with a business.

```
<?xml version="1.0" encoding="utf-8" ?>
<identity xmlns="http://tempuri.org/user.xsd">
        <key>Generated by some hash function</key>
        <id>Personal</id>
        <firstname>John</firstname>
        <lastname>Smith</lastname>
        <email>jsmith@hotmail.com</email>
        <company>none</company>
        <address>234 Queen St.</address>
        <city>Toronto</city>
        <state-prov>Ontario</state-prov>
        <postalcode>e5t3f5</postalcode>
        <phone>434-344-2344</phone>
</identity>
```