An Environment for Protecting the Privacy of E-Shoppers

by

Dora Carmen Gálvez Cruz

Submitted in fulfilment

of the requirements for the degree of

Doctor of Philosophy

Department of Computing Science

Faculty of Information and Mathematical Sciences

University of Glasgow

November 2008

Declaration of Originality

I Dora Carmen Gálvez Cruz declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography.

Abstract

Privacy, an everyday topic with weekly media coverage of loss of personal records, faces its bigger risk during the uncontrolled, involuntary or inadvertent disclosure and collection of personal and sensitive information. Preserving one's privacy while e-shopping, especially when personalisation is involved, is a big challenge. Current initiatives only offer customers opt-out options. This research proposes a 'privacy-preserved' shopping environment (PPSE) which empowers customers to disclose information safely by facilitating a personalised e-shopping experience that protects their privacy. Evaluation delivered positive results which suggest that such a product would indeed have a market in a world where customers are increasingly concerned about their privacy.

Contents

List of Tables

List of Figures

Acknowledgements

Esta investigación fue financiada por el Consejo Nacional de Ciencia y Tecnología CONACYT. A la institución y administrativos, reciban mi más profundo agradecimiento.

Great appreciation to my much-respected supervisors; Dr. Karen Renaud and Dr. Richard Cooper for all their time and support. It has been invaluable.

A mi maravillosa familia, Dora Cruz Gutiérrez, Martha Cruz Gutiérrez, Rafael Gálvez Cruz, Manuel Gálvez Cruz, Angelina Lozoya de Gálvez y Xquenda Gálvez Lozoya. Gracias a ustedes los retos se convierten en aventuras y las adversidades en experiencias dignas de narrar. Su apoyo hace que pueda seguir sonriendo a pesar de cualquier cosa. Los amo con todo mi corazón. Muchísimas gracias.

To my dearest friends; the light that guides me when I'm lost in darkness: Dra. María de Ujué Moreno Zulategui, Ms. Samantha Alvarez Madrazo, Ms. Rosie Smith, Miss. Lorna McEwan Kirk, Mr. Tomas Brichta, Mrs. Chiara Taurino, Dra. Ma. Del Pilar Angeles, Dr. Victor González Castro, Dr. Horacio González-Velez, Miss. Mariam Al-Awadi, Mrs. Jan Mansfield and Dr. Nigel Mansfield.

I am deeply grateful with my proofreaders; Dra. María de Ujué Moreno Zulategui, Ms. Samantha Alvarez Madrazo, Ms. Rosie Smith, Ms. Lorna McEwan Kirk, Mr. Mark Shannon, Mr. Fearghas MacFhionnlaigh and Miss. Betty Green.

For all the assistance with the design of the questionnaires, my gratitude to Dr. Margaret Brown.

My gratitude to the participants of my experiment.

Finalmente, un pequeño homenaje a las maravillosas personas que vivirán por siempre en mi corazón. Dr. Antonio Gálvez Martínez de Escobar, Doña Carmen Gutiérrez García y Federico Cruz Gutiérrez (para quien yo siempre fui su doctora). Muchísimas gracias.

Publications

- D. Galvez-Cruz and K. Renaud. What E-Grocery Customers really want: Personalised Personalisation. Proceedings of the Fourth Latin American Web Congress (LA-WEB'06), pages 109-112, 2006. Puebla, Mexico.
- D. Galvez-Cruz and K. Renaud. Privacy by agreement. IADIS E-commerce, pages 338-341, 2006. Barcelona, Spain.
- D. Galvez-Cruz and K. Renaud. You Know My'Alter-Ego'-but You Don't Know Me! In Databases, 2007. BNCOD'07. 24th British National Conference on, pages 101-109, 2007.

Chapter 1

Introduction

In our society, privacy is perceived as a human right. However, there is no precise delimitation of its boundaries and therefore its control and regulation. One reason could be that the limits of privacy are as flexible or as strict as the prevailing culture dictates. For instance, what one culture perceives as a privacy violation, such as someone invading interpersonal space, other cultures would not consider this a violation. Considerable effort has been undertaken to determine a definition of privacy that will support its regulation and preservation. Privacy preservation has been subject to several initiatives from different perspectives from standardising organisations (i.e. Privacy International) to commercial initiatives (e.g. issuing security seals) to research projects (i.e. Privacy Bird using P3P).

However, despite of all these efforts, privacy issues arise daily. For instance, privacy issues arise in cases such as the inclusion of biometric identification in the UK identity cards [13], the growing UK - DNA database [14, 15], or the regular loss of control over disclosed information, such as that of the patient records of nine English NHS trusts in 2007 [98]. Whereas in some cases, the control over privacy resides in organisational policies (i.e. guidelines established by governmental bodies), it can also come under personal control.

In the Internet era, privacy issues affect a wide range of areas. The area of interest in this research is the personalisation of e-commerce. Since e-commerce started in 1995, it has experienced exponential growth. Techniques such as personalisation have allowed e-shoppers to tailor the shopping experience, thus giving the business a better chance of fulfilling the customers' needs and thereby increasing profits in an increasingly competitive market. However, this process requires the collection and analysis of a great deal of information, and in some cases this information is misused. This latent risk has, in some cases, alarmed potential shoppers who attempt to defend themselves by using methods such as de-activating cookies or abandoning e-commerce stores. The lack of a shared understanding of privacy, together with the increase in the number of commercial initiatives which advertise privacy protection, makes it very hard for concerned users to protect themselves properly.

With the aim of raising awareness of the importance of privacy and its preservation, as well as providing an easy-to-use environment that allows customers to have a privacy-preserving e-shopping experience, the PPSE, a privacy-preserved shopping environment, is proposed in the thesis statement:

> ***It is possible to develop a privacy preserving shopping environment (PPSE), which respects the customer's privacy needs while allowing the company to gather and use sufficient reliable customer-specified data to achieve a level of personalisation which can be used to encourage customer loyalty.***

Three different approaches have been identified in the current efforts towards preserving privacy: raising *awareness*, *regulation* and the use of *technology* (ART). Whereas other related efforts use one or at most two elements of the ART approach, the PPSE combines all three to provide an integral approach towards the preservation of privacy.

The support of the thesis statement was organised in two activities.

Firstly, a prototype of the PPSE environment was designed and implemented. The prototype contains a third party Web portal named Alter-Ego, which has the objective of facilitating and mediating the customer's disclosure of information and the e-tailer's user-specified data requirements while simultaneously providing a contract for asserting the privacy level which is called the Personal Level Agreement (PLA), that formalises the exchange of information (sensitive and preferences-related) between customers and e-tailers. Finally, to complete the test environment, an e-grocery shop was implemented.

Secondly, a user test was performed to evaluate the customers' satisfaction and potential loyalty.

1.1 Road map

This thesis is organised as follows:

- Chapter 2 explores e-commerce, its beginnings and current state by identifying historic milestones. It also analyses it from a business perspective. Finally, e-groceries are identified as one of the areas with the slowest growth within e-commerce. This chapter ends by introducing the elements that customers value most when e-shopping and privacy is singled out as the topic of research.

- Chapter 3 explores personalisation, its influence in the business process and the benefits that can be obtained from its use. The chapter ends by discussing privacy issues related to personalisation.

- Chapter 4 discusses privacy, its concepts and definitions and includes a proposal for a definition of privacy to be used within the rest of this dissertation. Privacy issues are further examined and privacy preservation initiatives are described.

- Chapter 5 presents a proposal to preserve privacy while e-shopping. The thesis statement, introducing the PPSE, is presented and elements required to test it are introduced.

- Chapter 6 discusses the design and implementation of the prototype PPSE.

- Chapter 7 presents details of the evaluation of the PPSE. This chapter ends by drawing conclusions about the validity of the thesis statement.

- Conclusions and future work are presented in Chapter 8.

Chapter 2

E-commerce

2.1 Introduction

In general, commerce is often one of the first technology adopters, with retailers constantly searching for ways of adapting technological and scientific discoveries to improve their profits. Within commerce, the Internet is perhaps the technology that has influenced it the most in the shortest time. With the outset, software applications, have been developed to exploit the full capacity of the Internet and the Web, opening an opportunity for the exploration of a new form of vending: e-commerce. The creation of faster computers with more powerful processing capacity, has allowed the Internet's developers to incorporate new features that make interaction easier and more attractive to users and potential customers, and this has been one of the key factors that marks the massive uptake of e-commerce [87, 55, 35, 114]. However, the Internet, as a technology, has been effectively present for fewer than 30 years, and e-commerce for no more than 15 years, and its influence on communications in general is still evolving. This chapter explores e-commerce, giving a brief historical introduction to the Internet and e-commerce in Section 2.2, and presenting e-commerce as a business in Section 2.3.

2.2 E-commerce

The origins of e-commerce are interlinked with the beginnings of the Internet. Without the infrastructure that the Internet provides to improve communications, e-commerce could not exist in its current form. The Internet and e-commerce's history are illustrated graphically in the time-line shown in Figure 2.1

2.2.1 Infrastructure - Origins

The origins of the Internet can be traced back to 1958, during the cold war. In 1958, the USA's Department of Defense created the Advanced Research Project Agency(ARPA) to research military related problems as a response to the Soviet technological success on launching Sputnik. Via ARPA, universities and corporations received funding for the creation of a computer network. The objective behind this computer network was to inter-connect computers and to share data and programmes remotely. The idea was that, if there was a failure or infiltration in one part of the network, the rest of the network would not crash and could still function [111].

In 1962, the idea of creating a "Galactic Network" was shared in a series of memos written by J.C.R. Licklider from MIT. That network would interconnect computers to share data and programs remotely. The implementation of these first ideas was made possible by using packet switching technology [80], the theory for which was published by Leonard Kleinrock in 1961. By 1965 Thomas Merrill and Lawrence G. Roberts created the first wide-area computer network by connecting computers in Massachusetts Institute of Technology (MIT) to computers in California using a low speed dial-up telephone line. In 1966, the plan for the ARPANET was developed by Roberts.

ARCANE's expansion process started in 1969 with the first node located in the University of California Los Angeles (UCLA). The second node rendered at the Stanford Research Institute (SRI). These two nodes were followed by nodes in University of California Santa Barbara (UCSB) and University of Utah. In Europe, England and Norway had the first international nodes in 1973, and by 1977 the ARPANET had 107 nodes. In 1972, and Following Kahn's[1] idea of open-architecture networking, the ARPANET started functioning with multiple independent networks. The term **Internet** was defined within a resolution issued from the Federal Networking Council (FNC) on the 24 of October 1995 [80]

2.2.2 First software developments

At the same time as the Internet's infrastructure was expanding, software development started for this environment. One of the first developments was the Network Control Proto-

[1] From Bolt Beranek and Newman technologies (BBN)

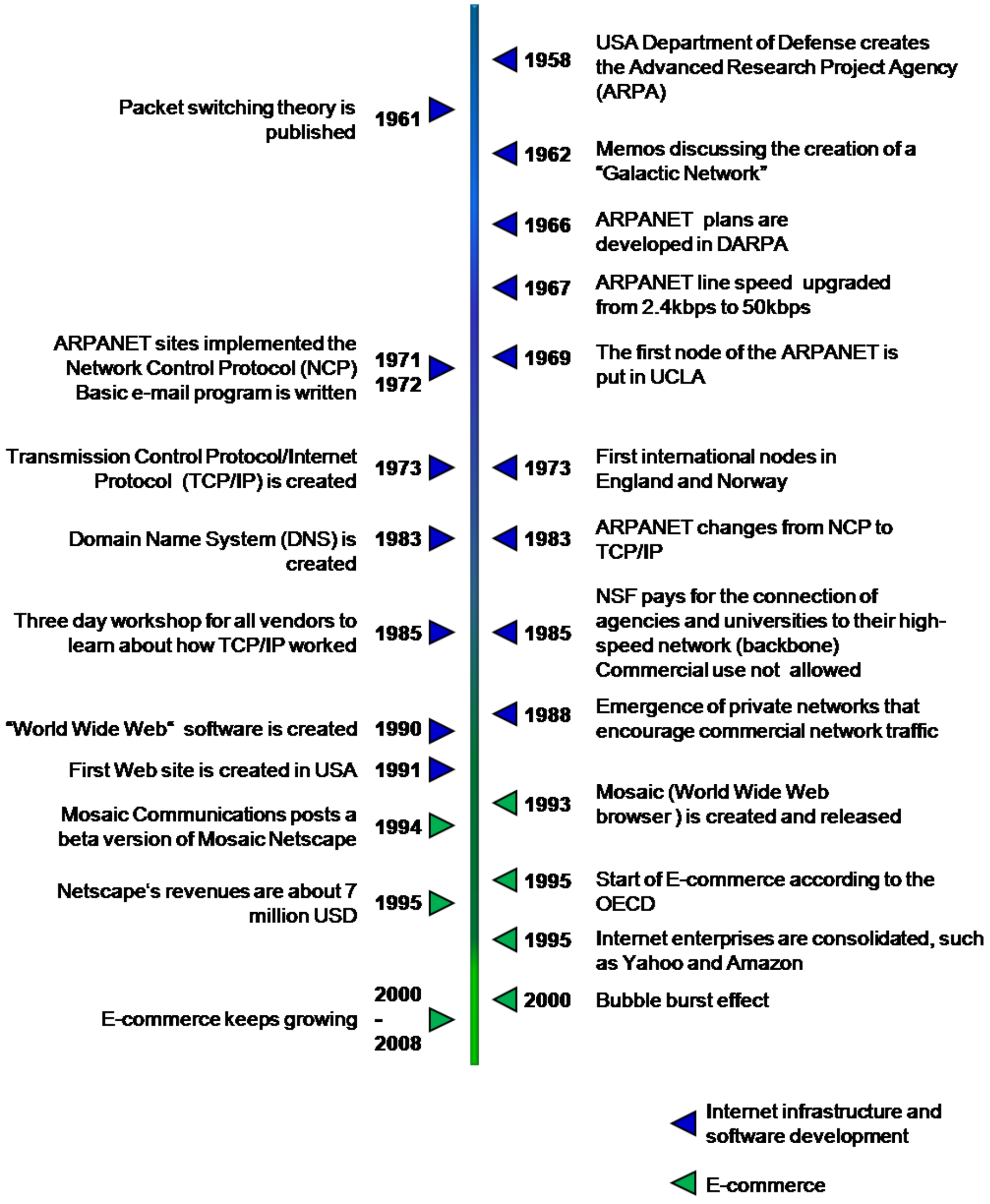

1958 USA Department of Defense creates the Advanced Research Project Agency (ARPA)
Packet switching theory is published 1961
1962 Memos discussing the creation of a "Galactic Network"
1966 ARPANET plans are developed in DARPA
1967 ARPANET line speed upgraded from 2.4kbps to 50kbps
ARPANET sites implemented the Network Control Protocol (NCP) 1971
Basic e-mail program is written 1972
1969 The first node of the ARPANET is put in UCLA
Transmission Control Protocol/Internet Protocol (TCP/IP) is created 1973
1973 First international nodes in England and Norway
Domain Name System (DNS) is created 1983
1983 ARPANET changes from NCP to TCP/IP
Three day workshop for all vendors to learn about how TCP/IP worked 1985
1985 NSF pays for the connection of agencies and universities to their high-speed network (backbone) Commercial use not allowed
1988 Emergence of private networks that encourage commercial network traffic
"World Wide Web" software is created 1990
First Web site is created in USA 1991
1993 Mosaic (World Wide Web browser) is created and released
Mosaic Communications posts a beta version of Mosaic Netscape 1994
1995 Start of E-commerce according to the OECD
Netscape's revenues are about 7 million USD 1995
1995 Internet enterprises are consolidated, such as Yahoo and Amazon
2000 Bubble burst effect
E-commerce keeps growing 2000 - 2008
Internet infrastructure and software development
E-commerce

Figure 2.1: Time line of the Internet and e-commerce origins.

col (NCP) released in the early 1970s. In 1972, Ray Tomlison at Bolt Beranek and Newman technologies (BBN) wrote the basic e-mail program. Later, the Transmission Control Protocol/Internet Protocol (TCP/IP) protocol was developed, and it was in 1983 that ARPANET changed from NCP to TCP/IP. The exponential increase in the size of the Internet required rapid adaptation to support its usage, and one of the biggest problems was keeping track of the myriad new sites. There were so many sites that, in fact, a simple list could not cover them all. This remained a problem until Paul Mockapetris invented the Domain Name System (DNS). The Internet was now bigger, more popular, faster and had easily locatable sites, but the users faced one particular problem: the Internet was still not easy to use. It was during this fraught expansion period, that in 1990, Tim Berners-Lee wrote software that he named the"World Wide Web" [80].

In principle, the World Wide Web operates on top of the Internet infrastructure, using its technology, protocols, computers and phone lines, focusing on three basic elements [111]:

- *Hypertext Markup Language* (HTML) - A computer language to format hypertext files,
- *Hypertext Transfer Protocol* (HTTP) - A communication protocol for the WWW to allow the downloading and formatting of Web pages, and
- *Uniform Resource Locator* (URL) - A unique address code attached to each file to provide an address for any file on the Web.

The World Wide Web (WWW) slowly infiltrated the Internet user's consciousness, but its use soon grew exponentially after the first USA Web site was created by Paul Kunz, a Stanford computer scientist [20]. In response to this growth, measures were taken to ensure that uniformity could be maintained across all users of the WWW, and so the World Wide Web Consortium (W3C) was created to publish a set of standards to be observed by software developers.

2.2.3 E-commerce appearance

The participation of the National Science Foundation (NSF), an agency dedicated to the support of scientific research, gave the Internet the boost it needed to continue expanding, so that it currently spans the globe and receives new members daily. The NSF paid for the connection of agencies and universities to their high-speed network (backbone), however, part of the conditions set by the NSF stipulated the mandatory use of TCP/IP protocol and the prohibition of commercial use of the NSFNET backbone. This prohibition, far from discouraging e-commerce, was the starting point, leading on to the creation of private networks that offered alternatives for commercial traffic.

Regardless of the incursion of commercial traffic, and even though the use of the WWW was spreading, browsing was still complex, and only scientist/computer programmers with access to the Internet participated. To assist non-experts computer-users with their Internet navigation, "Mosaic" was created and released by Marc Andreessen in 1993 [20]. This Web browser enjoyed immediate success due to ease of installation and its capacity to work on a variety of operating systems. Mosaic changed the nature of the Internet's traffic, as when it appeared, only one percent of the Internet's traffic was coded using the WWW. Two years after the uptake of Mosaic, 25% of the Internet's traffic was related to the WWW. Two years later, Microsoft released Internet Explorer.

The use of the Internet and the WWW was quickly adopted by the public in general. During the aforementioned expansion period, one of the main ideals ruling the Internet was that everything was provided freely. Knowledge and applications were shared without cost, and any attempt to commercialise software was immediately condemned by the users, who reacted aggressively to the suggestion.

Even allowing for this, the Internet soon entered a commercialisation period. A clear example is the case of Mosaic Communications who posted a commercial beta version of Mosaic Netscape in October 1994. The principle behind such commercialisation of software was to make it available free of charge for educational use while charging for private and commercial use. The cost of Mosaic was $39USD after a ninety-day free trial. Within hours thousands of computers around the world were downloading the software, the name "Mosaic" was changed to "Netscape" after a dispute with the University of Illinois[2], and by 1995, Netscape's revenues were reaching about 7 million USD [20]. With the availability of more user friendly software (based on the WWW) and of faster computers with more processing capacity and higher-speed Internet connections, navigating the Web overcame the computing-science-expert barrier and became available to a broader population of users. The potential outcome of reaching this larger audience attracted retailers' attentions, and

[2] Mosaic, developed by Marc Andreessen, had to change the name of the application since the original software under that name belonged to the University of Illinois

so they started exploring new ways of trading in what became *e-commerce*.

E-commerce now provided the opportunity to carry out business on three different levels; business to business (B2B), business to consumer (B2C), consumer to consumer (C2C).

2.2.4 Business to business - *B2B*

An organisation can be modelled as "a series of independent activities that deliver a product or service to a customer" [18]. This is illustrated by a generic model of an organisation called "value chain" and shown in Figure 2.2.

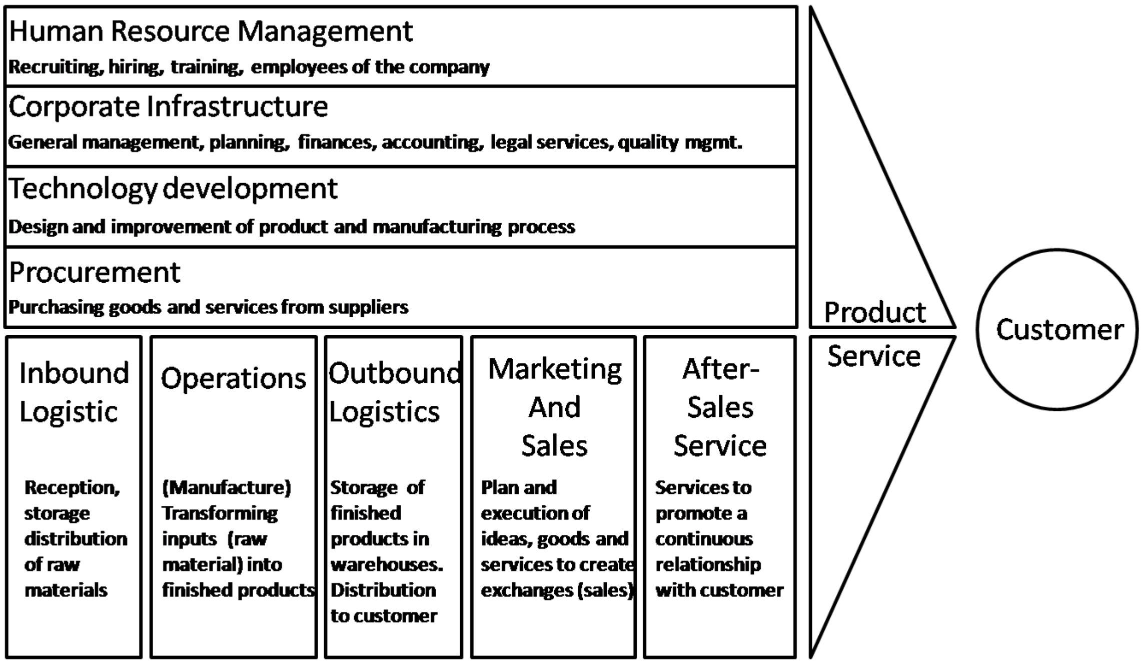

Human Resource Management
Recruiting, hiring, training, employees of the company
Corporate Infrastructure
General management, planning, finances, accounting, legal services, quality mgmt.
Technology development
Design and improvement of product and manufacturing process
Procurement
Purchasing goods and services from suppliers
Inbound Logistic
Reception, storage distribution of raw materials
Operations
(Manufacture) Transforming inputs (raw material) into finished products
Outbound Logistics
Storage of finished products in warehouses. Distribution to customer
Marketing And Sales
Plan and execution of ideas, goods and services to create exchanges (sales)
After-Sales Service
Services to promote a continuous relationship with customer
Product
Service
Customer

Figure 2.2: Generic model of organisation, known as ***value chain***. *Image adapted from [18].*

B2B e-commerce involves the sale of products and services between organisations and the automation of systems via a supply chain ("activities performed by an organisation in relation to its suppliers"[18]). Suppliers, distributors, manufacturers and stores all operate under the umbrella of this category of e-commerce [70]. The use of B2B, illustrated in Figure 2.3, affords the organisations lower purchasing costs due, among other reasons, to the reduction in the layers of processes involved. It also presents benefits to the business itself, allowing a reduction in inventory and production times. However, the development and maintenance of documentation standards, the security of data transmission and the secure access to extranets have all been identified as obstacles to B2B transactions [18]. A good example of B2B e-commerce is *OneSource*, an organisation that optimises information for fast interpretation, manipulation, analysis and reporting, according to the magazine *B2B marketing online* [7].

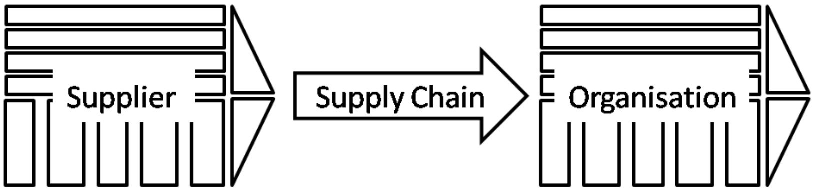


Figure 2.3: Business to business e-commerce (B2B). *Image adapted from [18].*

However, e-commerce was not the first technology-assisted process uptaken by the business sector. Before the expansion and popularisation of the Internet, technology was already assisting in commerce, such as in the case of Electronic Data Interchange (EDI). EDI's main objective is to provide a link between sender and receiver business applications with no human intervention at the receiving end. As illustrated in Figure 2.4, "EDI is the transmission of machine-readable data between trading partners' computers" [138], using a collection of standard message formats with which the transaction is carried out [138, 18]. EDI has mainly been used in business to business (B2B) transactions. From its inception, EDI was perceived as an answer to the problem of time delays and inaccuracies which paper-based business documents presented [138, 52]. The benefits of EDI include: reductions in the cost of handling business transactions, faster exchange and processing of information, reduction in the length of cycle from ordering to payment, and an improvement in the intra-company flow of information. However, the creation of "Extensive Markup Language" (XML) by the W3C, offered the possibility of allowing the definition of the content of a document, as well as the flexibility to specify standard templates for business documents. Both have been mentioned as key motivations for replacing EDI with XML-based transactions [18]. That said, B2B e-commerce is not the only exchange to be considered here; another new form of e-commerce also emerged: Consumer to consumer (C2C) trade.

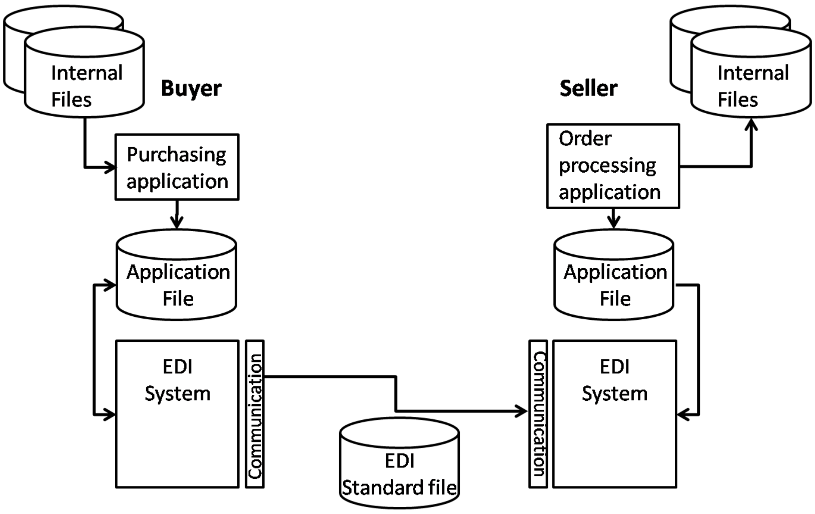


Figure 2.4: Buyer and seller flow of Electronic Data Interchange (EDI). *Image adapted from [138].*

2.2.5 Consumer to consumer - *C2C*

The Internet has supplied the infrastructure to allow consumers to be involved in the shopping experience to a greater degree than merely the conclusion of the transaction. C2C e-commerce, as shown in Figure 2.5, involves a customer's direct participation assisted by a community chain. The community chain "is based on informal social networks of individuals and is a major force underlying C2C e-commerce" [18]. An example of this kind of e-commerce can be found on eBay which provides the technology to facilitate C2C e-commerce [54]. However, it is in the third type of e-commerce, business to consumer (B2C), that focus of this work lies.

Customer
Community Chain
Customer

Figure 2.5: Consumer to consumer e-commerce (C2C). *Image adapted from [18].*

2.2.6 Business to consumer - *B2C*

B2C, as shown in Figure 2.6, is related to the interactions and transactions between an organisation and its consumers, and involves a customer chain ("chain of activities that an organisation performs in the service of its customers" [18])

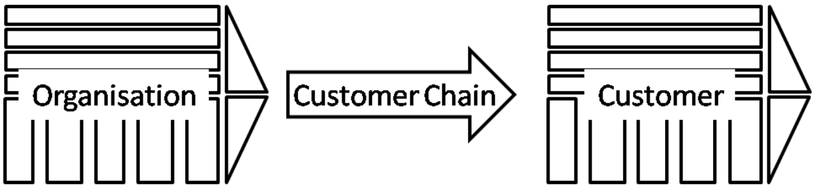


Figure 2.6: Business to consumer e-commerce (B2C). *Image adapted from [18].*

Using a structure such as the one shown in Figure 2.7, B2C e-commerce has facilitated a close relationship between organisations and customers. This type of e-commerce has allowed customers to have access to; merchandise from different locations (national or international), the possibility of comparing costs and quality in products and services, and the flexibility of permanent access to the store. Similarly, organisations benefit from this kind of commerce, as transaction costs associated with sales are reduced, and it presents saving opportunities in the storage of merchandise [70, 18]. Amazon and Dell are examples of these kinds of businesses.

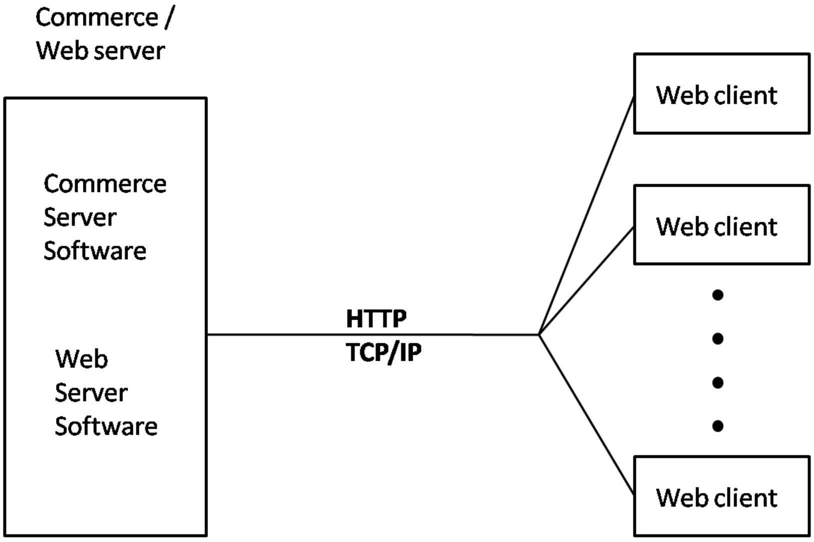
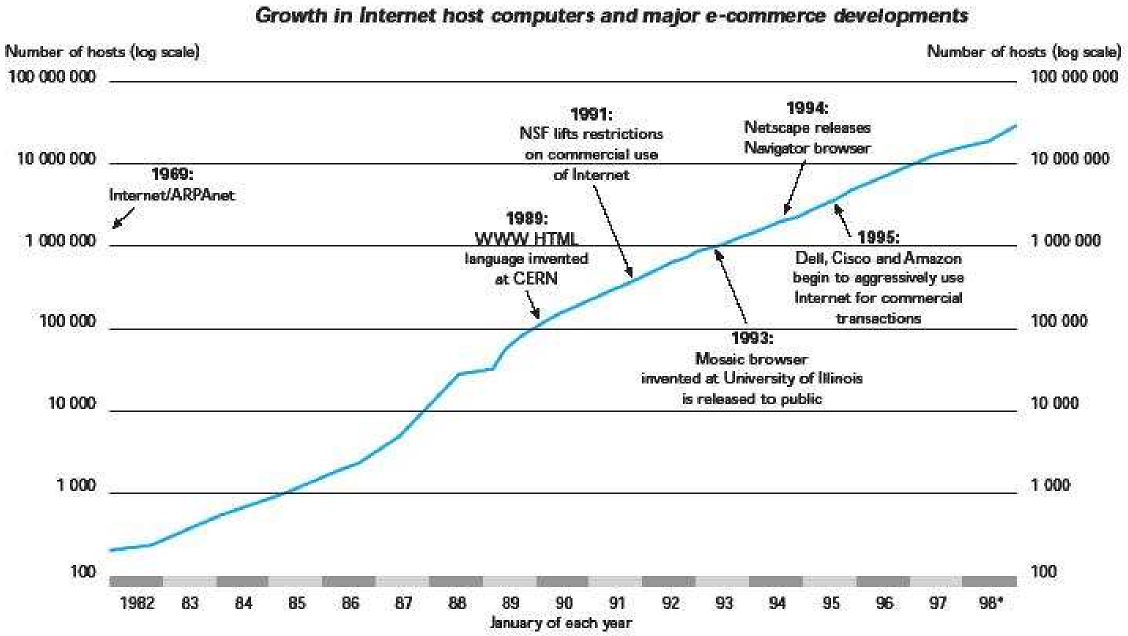


Figure 2.7: Business to consumer e-commerce (B2C). *Image adapted from [70].*

Please note; the term *e-commerce* used in this work refers to B2C e-commerce only.

The OECD reports 1995 as the start of B2C e-commerce. Even from this early stage the outlook was encouraging. Figure 2.8 [108], illustrates the OECD's graph of the growth in the

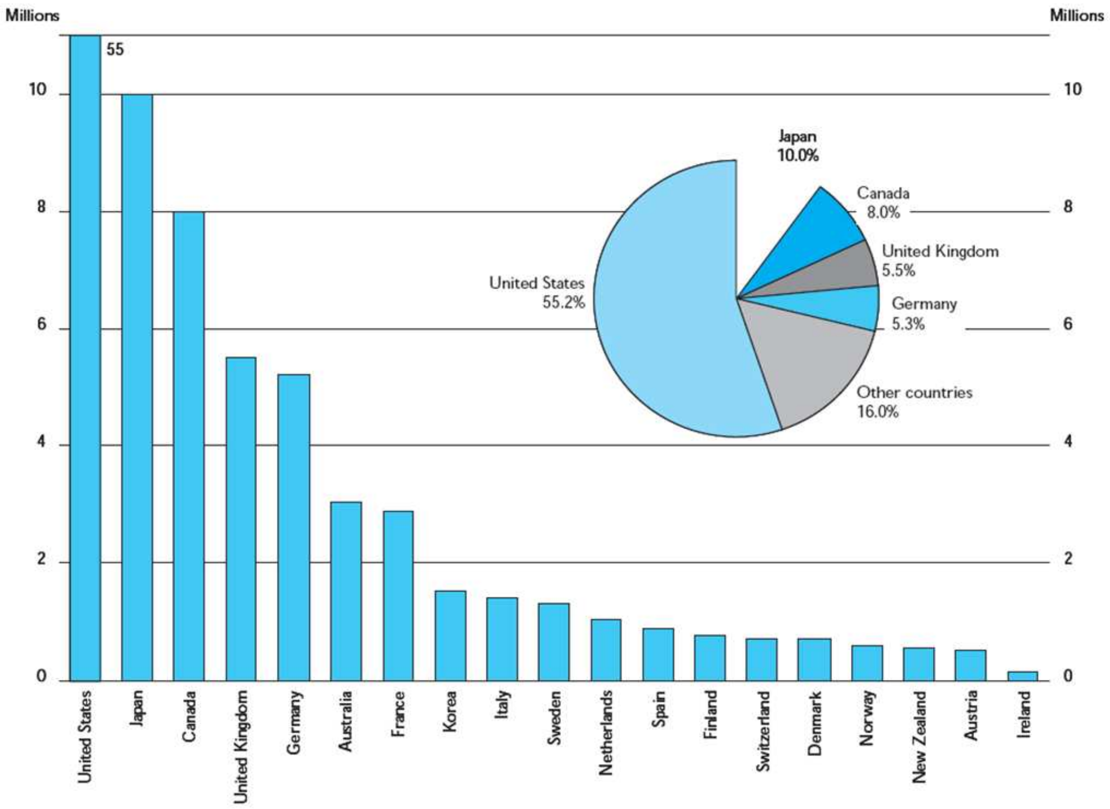
Internet host computers and major e-commerce developments. The impact that e-commerce had worldwide up to 1998 is shown in Figure 2.9 [108].



* New methodology used in January 1998.
Source: Network Wizards.

Figure 2.8: Growth in The Internet host computers and major e-commerce developments [108].

Figure 1.1. ***Adults accessing the Internet, selected OECD countries***[1]



1. Methodologies vary across countries. Data are provided as an indicator of the diffusion of Internet's use within countries. The numbers were measured between December 1997 and June 1998.

Source: OECD, based on data from http://www.headcount.com.

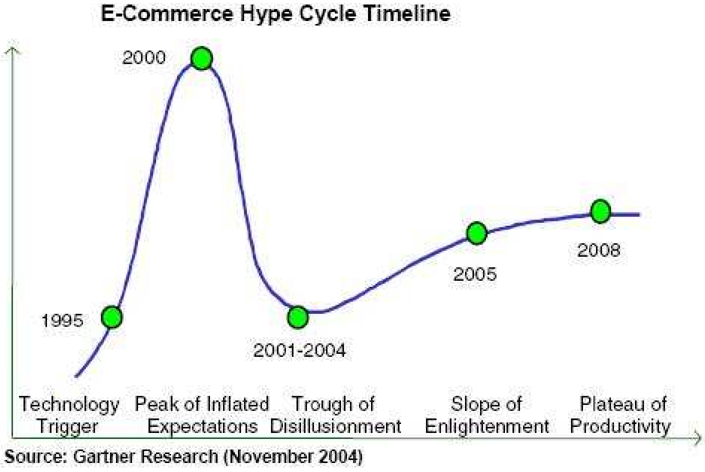
Figure 2.9: Adults accessing the Internet in selected OECD countries in 1998 [108].

The widespread acceptance of the WWW and the availability of Web browsers were complemented with the creation of search engines which facilitated searching for specific Web sites.

In 1994, David Filo and Jerry Yang wrote software that allowed them to group together their favourite sites. After posting the software on the Web, under the name Yahoo, it had immediate success. Yahoo, as an enterprise, was consolidated in 1995 allowing e-commerce sources and general Web sites to be easily located. A large number of Internet-based companies began emerging in the 1990's, and the subsequent rise in the stock market for Internet-based companies such as Amazon, Dell and eBay was exponential. Revenues were increasing exponentially in the stock market too, and successful stories were often presented in the media. Those factors, amongst others, accelerated the growth of e-commerce in a "bubble" effect.

However, at the end of 1999 and beginning of 2000, when the exponential growth could not be sustained, a contra effect known as the bubble burst occurred. During the bubble burst, the e-commerce-based economy collapsed, and the effects of this were immediately reflected in the stock market. In a one-month period, March-April 2000, the stock market value of the Internet companies suffered dramatic decline. In some cases, such as Akamai Technologies, the losses were close to 78 %, while in other cases, such as Amazon, the effect was less drastic, but with losses of 29.9 % were still incurred within the same month. The bubble burst effect in the UK was, according to Cassidy [20], similar to the experienced by the USA, but in a smaller proportion.

Despite the losses experimented during the bubble burst, a steady recovery of the profits obtained from e-commerce was projected by analysts. For example, in the USA, Gartner[3][127] illustrates in one of their "hype cycles" time lines their projection of the recovery of e-commerce, shown in Figure 2.10.



Source: Gartner Research (November 2004)

Figure 2.10: Gartner projection of e-commerce behaviour in the USA[127].

As projected by Gartner, the recovery of e-commerce after the bubble burst presented a slower but steady growth than before. A similarity in the curve of economy growth projection made by Gartner, Figure 2.10, and the behaviour of the e-commerce in the USA, Figure 2.11, can be found in the *e-commerce growth analysis* published by the U.S. "*Monthly Retail Trade Survey*". The "U.S. Census Bureau" publishes a quarterly estimate of e-commerce's

[3] Gartner is an independent IT research and advisory enterprise

growth, presenting data obtained from approximately 12,500 retail firms. The sample is selected from a pool of over two million retail firms using a stratified simple sample random method [154].

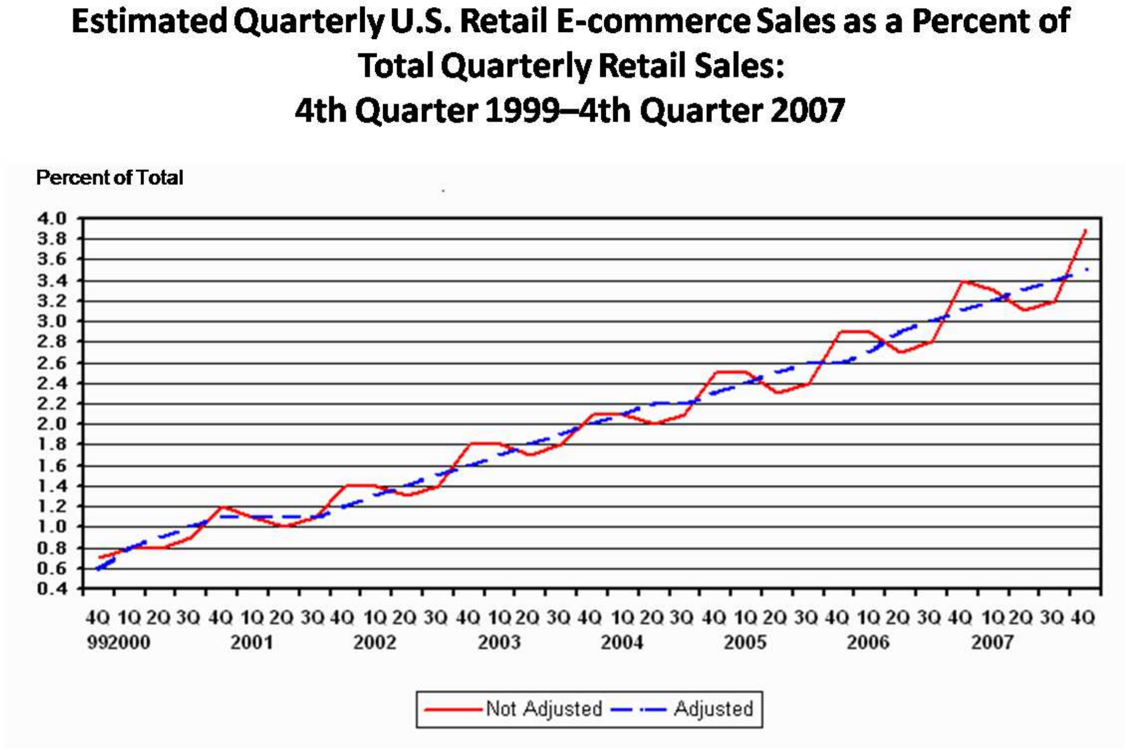


Figure 2.11: Information from U.S. Census Bureau 15 February 2008 [154].

However, e-grocery is one area of e-commerce that has not presented the same growing proportion as others, this is examined next.

2.2.7 E-grocery

In spite of the multiple advantages that e-grocery could bring to customers, such as; detailed catalogue information, storage of shopping lists and personalisation, customer preference for this area of e-commerce has not been so successful as the others. This slow growth in e-grocery can be seen in the reports presented from The European Interactive Advertising Association (EIAA)[4]. As Figure 2.12 shows, the e-grocery growth in 2004 and in 2006

[4] The EIAA is “A pan- European trade organisation for media companies focused on growing interactive business” [151]

(reported in 2007) had a slow pace [151].

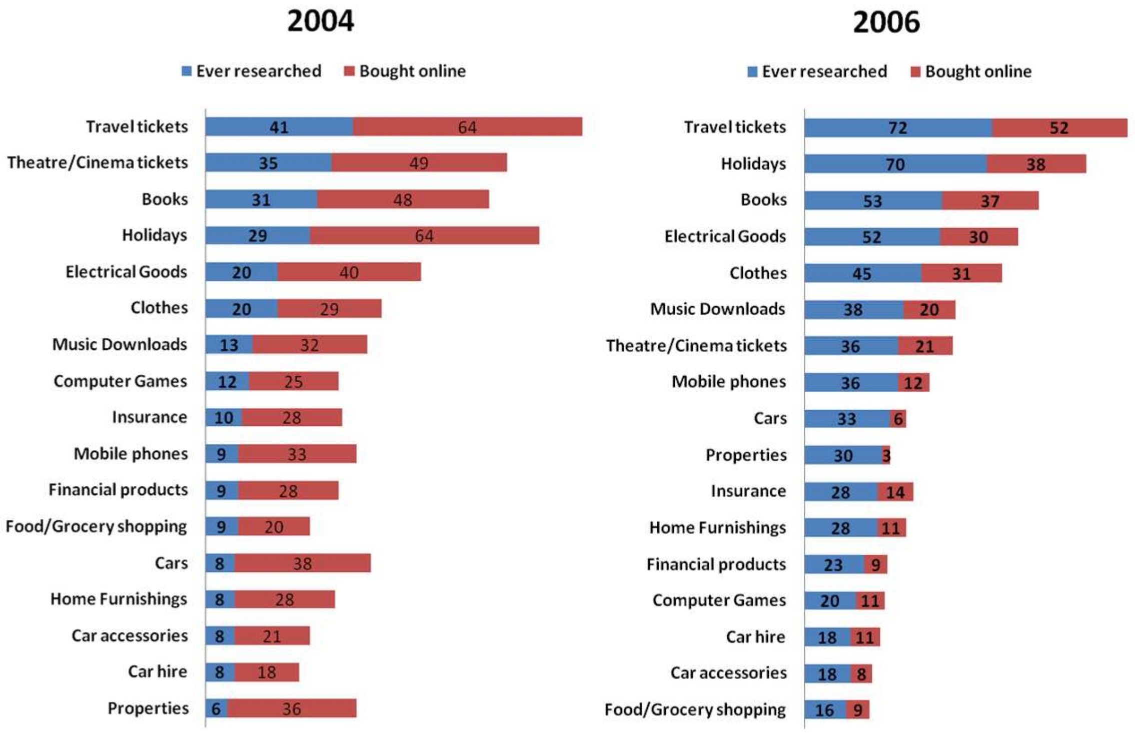


Figure 2.12: E-grocery growth in 2004 *Adapted from [135]* and 2006 *Adapted from [151].*

The reasons for the slow growth of this segment of e-commerce are still the subject of ongoing research. Regardless of the benefits of a direct relationship between the customer and e-grocery store can bring to both "e-tailers" and customers alike (such as better product information), only a minority of the customers consider e-groceries when they think about shopping on the Internet [48]. Issues so diverse as; sensory issues (such as the lack of "touch and feel" of the goods), shopping ambience, substitutions, correct packing of merchandise, temperature, the cost of delivery and management have been explored [130, 85, 68]. However, during the results of a survey presented by [48], a majority of respondents agreed that they would use an e-grocery site that provided them with results corresponding to their preferences and that respected their privacy.

A pilot study was undertaken during an early stage of this work. This study consisted of an online questionnaire primarily to explore the following question: *If customers were presented with an e-groceries site that had their own personalisation choices and reinforces their privacy, would that encourage them to buy e-groceries?* Secondary questions exploring the customer's perceptions about particulars on personalisation, e-grocery and privacy were raised as well. The opinions of 84 participants were collected and analysed after one month. The following significant results were retrieved (more detailed information can be found on [48]):

- The three topics most frequently mentioned by participants when talking about Internet shopping are: books, travel and electronics. These were also what users most frequently bought over the Internet.
- The three items least referred to are: cars, groceries and services.
- The three biggest reasons for Internet shopping (according to the participants' opinions), were: delivery to their door, laziness, and value for money.
- Only 11 bought groceries over the Internet, 90% (10 participants) considered it a successful experience, 8 participants received the goods they expected and which matched their expectations; however they were unsure whether they would return to the e-grocery store.
- The features participants valued most were: "data held about me will not be shared or sold" (privacy), "free delivery from the store" (store service), "I was able to view all the data recorded about me by the store" (privacy) and "regular delivery of items without having to keep going back to the Web site" (store service).
- Regarding e-loyalty, participants were asked if they would buy or recommend an e-groceries site that could provide a certain level of personalisation and preserve their privacy. 49 participants agreed to buy there (65%) and 30 participants (40%) responded that they would buy there on a regular basis.

From the analysed data, the participants' responses towards their shopping preferences online follow a similar pattern to that obtained from EIAA. The results obtained from the pilot questionnaire are shown in Figure 2.13. Understanding 'what customers want' is one of the main concepts in business, since knowledge of the customer's needs and desires can aid in planning processes. The next section explores e-commerce as a business.

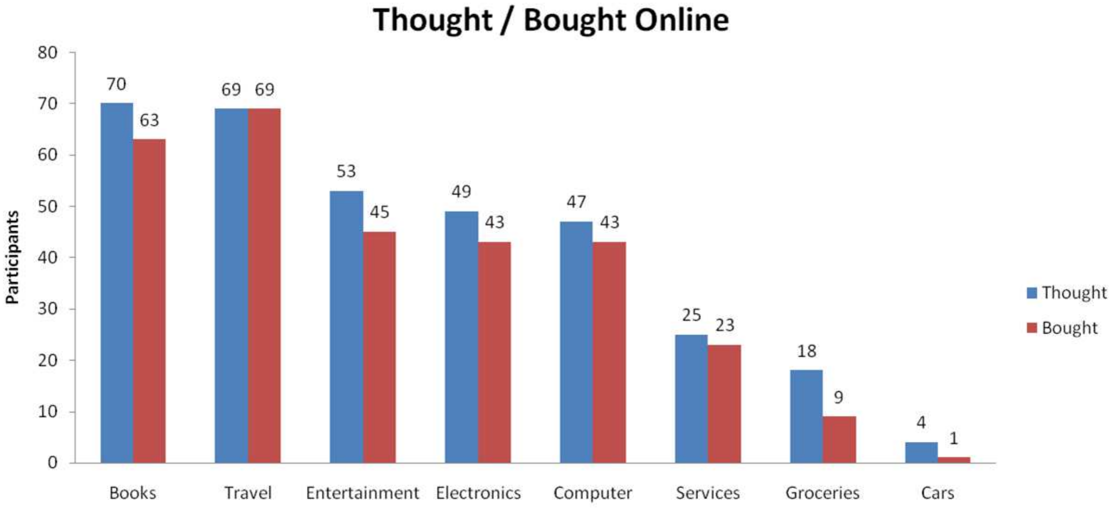


Figure 2.13: Results obtained from the pilot study about customer shopping online.

2.3 E-commerce - Business

From the business perspective, the presence of the Internet significantly expands the scope of the business model. A business model "specifies the structure and dynamics of a particular enterprise"[18] and includes entrepreneurship, strategy, economics, finance, operations and marketing [71]. In previous sections the different e-commerce approaches were introduced: B2B in Section 2.2.4, C2C in Section 2.2.5 and B2C in Section 2.2.6. These forms of e-commerce, shown in Figure 2.14, bring diverse benefits, such as cost savings (costs related to logistics, postage, storage, and employing and managing personnel) time savings (response time to markets, processing of payments), connection improvements (reduction of intermediaries), quality improvements and strategic improvements (efficient organisational forms of doing business) [18] to online business (e-commerce).

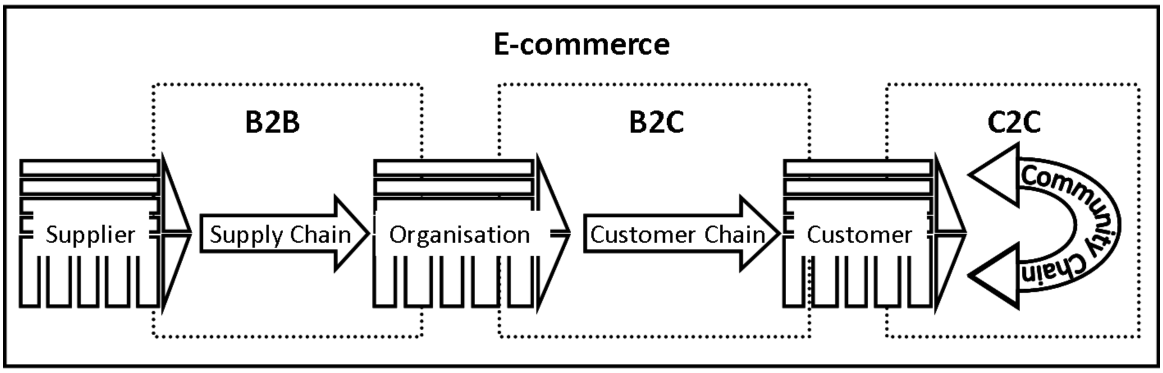


Figure 2.14: E-commerce, B2B, B2C and C2C. *Adapted from [18].*

Online commerce, similarly to conventional commerce, aims to foster a close relationship with the customer, and attempts to encourage loyalty to their e-commerce sites (e-loyalty). This is so that the customer returns to buy again later. Three generic trade cycles, as shown in Figure 2.15, can be identified according to the their frequency of occurrence; cash, credit and repeat. Cash occurs in an irregular frequency basis and involves one-off transactions between economic parties. Credit again involves irregular frequency transactions, since the processes of settlement and execution are separate. Repeat transactions have regular frequency in the transactions. E-commerce can be applied to all or different phases of the trade cycle.

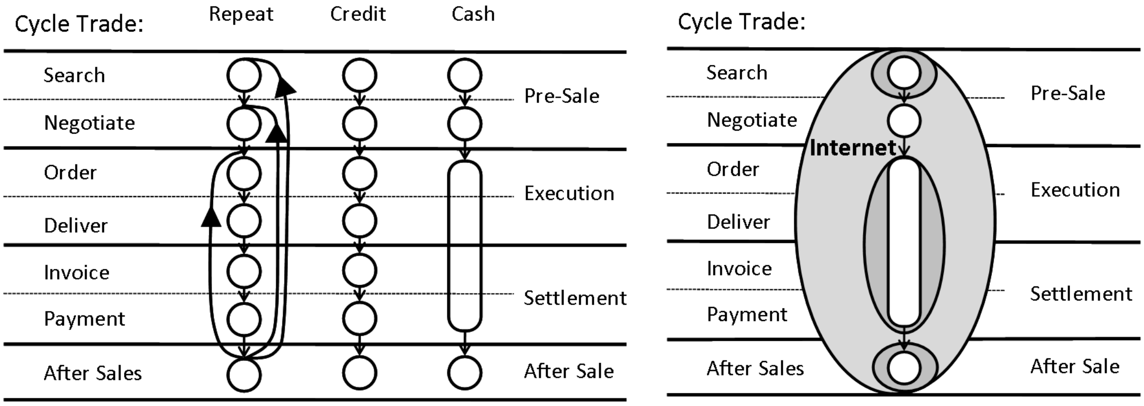


Figure 2.15: Left: Generic Trade Cycles. Right: Internet within generic trade. *Both figures Adapted from [159]*

.

Retailers have focused their attention on migrating their cycles from “cash” to “repeat” and to promote e-loyalty. Studies have been carried out to determine customer preferences and behaviours, such these can be matched by the store and so retain the customer’s attention. An example of these studies is presented by Paco Underhill, who defined and studied “The science of shopping” for traditional commerce [153]. In the science of shopping, a relationship between physical aspects (i.e. the ergonomics of a store), and the elements that guide the customer’s decision to buy certain items is one of the most important findings in Underhill’s work [153]. Another important finding is the positive reaction that customers exhibit to small changes in the store, such as the customer’s favourable reaction towards the location of goods on the shelves. The merchandise that is located at their eye-level is more likely to sell than those that require meticulous searching. This was particularly prevalent in the case of sweets purchased by children or old people. Old people and children increased the sales of sweets when they were located at eye-level as opposed to when they were on upper shelves. At the same time, Underhill, found that the merchandise situated in the first few metres inside the store (called the “landing zone”) is less likely to be sold than the rest of the goods in the store.

On the other hand, shopping is not only influenced by the location of goods. Customers also experience a range of different feelings while shopping; a mother might spend less money if the goods are for her, while spending more money if the goods are for her children [91]. When considering subjective factors related to shopping, Kasanoff makes a distinction between “necessity shopping” and “desire shopping”. In necessity shopping, customers buy the goods that they really need, whereas desire shopping occurs purely to satisfy a desire, fashion or mood. Retailers have found that they receive more income from desire shopping than from necessity shopping [67].

Retailers have explored ways of adapting the success factors identified by these studies and other practices of brick-and-mortar commerce into e-commerce. The advances in technology have provided the elements for adapting proven successful factors to e-commerce, hence, to compensate for the favourable location of goods within a store, e-commerce can attempt to attract the customers’ attention to certain offers. One of the techniques that can be used to assist and guide customers with the selection of goods is *personalisation*. Personalisation presents the customer with a tailored browsing environment, based on his or her previous browsing behaviour, expressed preferences, or previous purchases.

Another technique valued for business in general is “market segmentation”. Market segmentation can be defined as “the process of splitting customers, or potential customers, within a market into different groups, or segments, within which customers have the same or similar requirements satisfied by a distinct marketing mix[5]” [89].

With proper market segmentation in place, e-tailers can direct their efforts into more intelligently matching the customers’ needs. However, to be able to use these techniques (market segmentation and personalisation), customer information needs to be collected and analysed. To collect information, retailers most employ diverse methods. In traditional brick-and-mortar commerce, the collection practices take the form of loyalty cards or coupons in newspapers or magazines. The more detailed the requested information, the better the offers in order to obtain it. Hence, it is common practice to offer a free catalogue in exchange for information such as name, address, occupation and shopping habits. In e-commerce, however, the collection of information can be carried out without the customer’s knowledge. Since the customer’s online behaviour can be tracked continuously, and it is not uncommon for retailers to use tracking devices to gather as much information as they can. Unfortunately for consumers, this technological facility has been abused, leading to retailers collecting more information than they need to support their business planning. Problems arise when retailers misuse the collected information, use that information against the customers’ interests, or sell it to others, wherein the customer’s privacy and confidentiality are violated. If retailers adopt such improper practices, the advantages that e-commerce offers to customers are greatly diminished. Retailers employing improper practices always run the risk of being discovered. When this happens, customer trust is broken and, in the worst cases, is lost completely. The excessive collection of information and other privacy violations are explored in chapter 4.

The introduction of e-commerce into business has required development of techniques to match customer characteristics and behaviour. The new generation of customers is better informed, sometimes more so than the retailers themselves [77]. They seek value for money and make comparisons based on a large range of options before spending money [87, 152, 78]. These searches and comparisons have been assisted by third party Web sites, as shown in Figure 2.16, which facilitate the shopping experience.

[5] Marketing mix refers to the means available to improve the match between customer benefits and the store offers

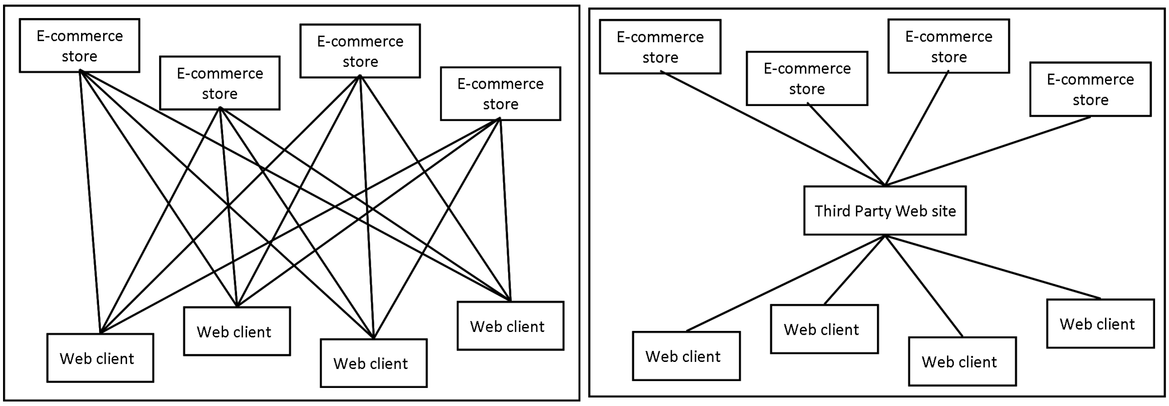


Figure 2.16: Left: Direct connection between customers and e-commerce sites. Right: Connection between customers and e-commerce sites assisted by Third party Web site

.

However, some negative practices such as customers abandoning the “shopping trolley” just before the purchase, or window shopping have transferred to e-commerce and remain in operation and indeed some are even easier to carry-out there [50]. Reasons for shoppers to use e-commerce vary. Some customers shop using the Internet because they perceive it as a status symbol [152], others use it because of necessity [48], however, whatever the reason, customers make use of e-commerce sites. Attracting and maintaining e-loyalty is becoming one of the more important tasks for retailers, especially since the goods acquired using the Internet are not immediately obtained [87]. A series of surveys has been the selected as the most efficient methodology for collecting customer’s opinions in order to explore what customers value the most when using shopping online.

The following list presents a summary of the responses obtained from diverse surveys. Customers most value the following when they use e-commerce shopping sites[6]: [87, 55, 35, 114].

- Customer satisfaction
- Information content
- Security / Security of payment

[6] The ordering is not significant.

- Ability to remember returning customers (customers do not wish to re-enter name, address and payment information)
- Download times
- Online communities (Such as chat, immediate assistance from the store, etc)
- Privacy policy regarding my personal details
- Ease of ordering online
- Cost of delivery
- Ease of finding out about the product
- Low price
- Previous experience with site
- Recommendation by friend or colleague
- Retailer's off-line presence

According to this list, customers concerns can be categorised in three groups using the following criteria:

1. E-store - Technology
    - Download times of information from Internet
    - Online communities (Such as chat, immediate assistance from the store, etc)
    - Usability - Ease of ordering online
    - Security - Security of payment
    - E-store - Personalisation
        - Information content
        - Ability to remember returning customers
        - Easy of finding out about the product
2. The store's business plan

- Cost of delivery
- Low price
- E-loyalty - Previous experience with site / supplier
- Recommendation by friend/colleague
- Retailer's off-line presence

3. Privacy policy regarding personal details

Therefore, while a majority of the customer's needs can be satisfied either technologically (group 1) or with a better business plan (group 2), privacy (and its preservation) can be singled out as a highly important issue aligned to the customer's values. Thus, armed with this information, e-tailers can make e-commerce a mutually beneficial relationship with customers. The avoidance of improper practices should be encouraged and a different approach to allow e-tailers to encourage e-loyalty should be explored. A trustworthy shopping environment could reinforce customer trust and reassure them that their privacy is protected. They may then feel more free to enjoy the benefits that e-commerce provides such as personalised shopping.

Personalisation can be perceived as a double-edged sword. On one hand, it allows a more focused, time-saving and recommendation-assisted shopping. On the other hand, the indiscriminate collection of information required to provide personalised recommendations, and the risk of improper inference of data, can lead to privacy issues. Personalisation is further explored in chapter 3.

2.4 Conclusion

With the Internet's supporting infrastructure, easier-to-use Web browsers, supportive software and series of mechanisms to ensure secure transactions, e-commerce has a fertile soil in which to flourish. The ubiquitous nature of the Internet has provided a perfect working environment for the introduction of new goods and shopping for traditional or exotic goods that were, up to now, unavailable. In the ideal case, the use of e-commerce represents a mutually beneficial relationship between retailers and customers. However, the technology employed to assist the shopping experience in e-commerce can be abused.

Privacy and incorrectly-used personalisation techniques are practices that diminish the benefits that e-commerce can provide. E-customers who become aware of these unethical practices often react defensively, and attempt to protect themselves by; abandoning the selected goods at the last moment, giving false information when the e-commerce site asks them to register before browsing, and, in the worst case scenario, avoiding shopping in e-stores altogether.

Due to the importance that customers give to the preservation of their privacy while shopping online, privacy can be considered as an essential aspect of e-commerce [87]; therefore there is a clear need to provide them with an environment within which they can control disclosure of their private data. At the same time the ability to enjoy the features that e-commerce can provide, such as personalisation, is essential.

E-grocery, an area of e-commerce that has not presented the same growth as others, offer many potential benefits to customers, but carries with it a potential privacy risk. The data inferred from an indiscriminate collection of customer information can result in privacy violations and, in the case of disclosure to third parties, results in confidentiality violations. Therefore, this research focuses on the e-commerce approach to privacy-preserved shopping within an e-grocery store.

Before we can explore this matter further, the next chapter presents a more detailed discussion of personalisation and its potential privacy problems.

Chapter 3

Personalisation

3.1 Introduction

Technological progress has allowed e-commerce to give back to e-tailers what a mass consumption market has limited: the capability of matching the client's individual needs and preferences with a personal approach in an automated fashion. It was in 1852, when department stores were introduced in Paris by Aristide Boucicaut [126], that customers were presented with shelf-located, price-marked and readily-accessible merchandises, making commerce impersonal. However, technological advances have now allow e-commerce customers to have a Web experience tailored to a particular user or set of users, by means of a process called personalisation. This in turn delivers an impression of a more individualised service from the e-tailer.

Due to its importance to the business, several related areas and uses have grown to be associated with personalisation. For instance, it is a contributor to business strategies, an influential factor in the customer's shopping experience, and a prominent input in marketing analysis studies (such as market segmentation). Unfortunately, personalisation, by its very nature, presents one major drawback: privacy.

Section 3.2 of this chapter presents e-commerce's uses and profits derived from employing personalisation (from a business perspective). The technical process followed to personalise, including its different phases, is introduced in Section 3.3. Applications of personalisation results, such as recommendation lists, are discussed in Section 3.4, and finally Section 3.5 examines the one major downside of the personalisation process; the threat to the privacy of customers.

3.2 Personalising the business

In Section 2.3, the different trade cycles, *cash*, *credit* and *repeat*, were discussed. E-commerce's trade cycle corresponds to the cash cycle which occurs on an irregular frequency basis and involves one-off type transactions between economic parties. Therefore, the importance of nurturing customer loyalty and promoting customer returns to the e-store is evident. However, in order to promote customer loyalty, the visitors needs to first be transformed into a buyer. It has been suggested that personalisation is an influencing factor not only in converting visitors into buyers (the "stickiness" process), but in influencing customers' choices too [109]. This section will discuss the role that personalisation plays in business and its influence on the stickiness process, the elements influencing customers' shopping and choices, and finally, exploring the financial success achieved by e-commerce by using personalisation techniques.

3.2.1 The influence of personalisation

In a universe brimming with different e-commerce Web sites and acknowledging the consequent competition, e-tailers have only a short span of time in which to engage the visitors' attention and to convince them to buy; therefore, finding ways of obtaining and increasing "stickiness" in e-commerce has been much explored. The term "sticky" is used when describing Web sites that "engage prospects and compel them to become purchasers" [109][p.400], and a sticky customer is "a consumer who has developed an affection, affinity or addiction to a site that compels him or her to return there often" [109][p.401]. Stickiness is therefore the keystone in the creation of a customer-business relationship, and a high level of stickiness can prospectively cause the trade cycle to evolve from *cash* to *repeat*.

Due to the importance of stickiness, the factors that identify sticky Web sites have undergone much research. Results obtained from surveys in relation to e-commerce Web sites (such as Amazon.co.uk, Dell.co.uk and expedia.co.uk) show that stickiness relies on multiple factors (behavioural and attitudinal), which is contrary to the previously held belief that considered the duration of the customer's visit to be the main metric for evaluating Web sites' success [109]. While behavioural factors, such as speed of transaction execution, can be directly obtained from customers' browsing sessions, attitudinal factors, such as 'whether the content is provided in an interesting manner', are abstract factors requiring the execution of a customer survey in order that they be taken into consideration as well.

Attitudinal factors, also called intrinsic motives, have a big influence on the customer's decisions while shopping. Surveys have shown that the major reasons for shopping on-line were resultant from intrinsic motivations, such as the playfulness of the Web site, and that extrinsic motivations were not significant [131]. These surveys showed that participants with cognitive absorption experience (a state of deep involvement with the software, that includes elements such as focused immersion, heightened enjoyment, control and curiosity) were more likely to shop online [131]. Therefore, it can be concluded that it is important to attract and maintain the customer's attention on the Web site and also to increase the customer's familiarity with the e-commerce site. The more thorough the user's knowledge of the e-commerce Web site, the greater their cognitive absorption. The permanence of the customer in an e-commerce Web site can be influenced by the presentation of different options that the customer might consider to be related to his shopping [131]. These options can be provided by employing personalisation techniques.

Since personalisation techniques deliver outcomes that can be used to influence the customer, they have, from the business perspective, an important role in persuasion strategy. Therefore it is important to observe the customer's reactions to Web personalisation in order to formulate a business strategy. Tam *et al.* [147] identified three persuasive factors: *level of preference matching* (also called "quality of content"), *recommendation set size* and *sorting cue.* The results of a series of surveys showed that customers are more likely to be persuaded to purchase when their preferences and needs are understood and matched by promotions and sales efforts. Matching preferences refers to "the extent to which the Web content generated by the personalisation agent appeals to users"[147][p276] and it is used by Tam *et al.* [147] as a measure of the Web personalisation's quality. Hence, it can be observed that an important factor which customers place a high value on is the quality (matching preference) of personalised Web content. The results obtained from Tam *et al.* [147] also showed that the provision of a sorting cue (specific cues used to direct users, such as Amazon.com's sales rank that shows the popularity of the product), was highly related. Finally, the size of the recommendation set was an effective attractor of users' attention.

Therefore, a personalisation-assisted strategy to influence the customer should focus on:

- constructing a detailed profile of the customer, so that the e-commerce store can match their offers to the customer's preferences and needs;

- paying special attention to the quality of the content (high level of preference matching) they offer;
- providing sorting cues to guide the customer's choices; and
- providing a large recommendation set.

3.2.2 Personalisation: a business process

It is important to note that careful planning is required before performing the data mining. The better the understanding of factors needed to support marketing analysis, the better the results that can be obtained from personalisation. For instance, a car rental company might wish to know the models of the most rented cars by males and females of between 30 - 35 years of age, when travelling to Germany on a weekend. They can then use those results to direct a personalised media campaign or a personalised rental offer.

Personalisation is a process where the results obtained from this first use of information obtained by personalisation techniques are not the end of the matter. On the contrary, the provided information becomes part of an iterative process that analyses the current situation and delivers elements which in turn allow future projections to be calculated for the business. The results obtained after applying the plans for such projections, are analysed to determine whether it was indeed a good business decision or not, and so begins the personalisation cycle all over again. It can therefore be stated that personalisation in e-commerce is not a single action, but rather part of an iterative process, shown in Figure 3.1.

Adomavicius *et al.* [2] identify three stages in the personalisation process that illustrate this iterative nature: *understand*, *deliver* and *measure. Understand* refers to the collection of customer data and the 'pattern-obtaining' achieved by analysis of the collected data. Two sub-groups form this understanding stage, *data collection* and *building customer profile. Delivery*, the second stage, is the action of tailoring the results obtained from the personalisation phases to the customers, and delivery sub-groups are observed to be: *matchmaking* (the process of matching or tailoring appropriate content and services to individual consumers), and *delivery and presentation.* Delivery and presentation refers to the way the information is presented to the customer, such as filtered content or recommendation lists. Finally, the *measure* stage evaluates the effectiveness of the implemented personalisation. As Figure 3.1 shows, timely feedback applied to each stage of the personalisation process, should improve performance.

Regardless of the process that each company uses for adapting personalisation to their particular business, personalisation can be implemented in e-commerce yielding positive results.

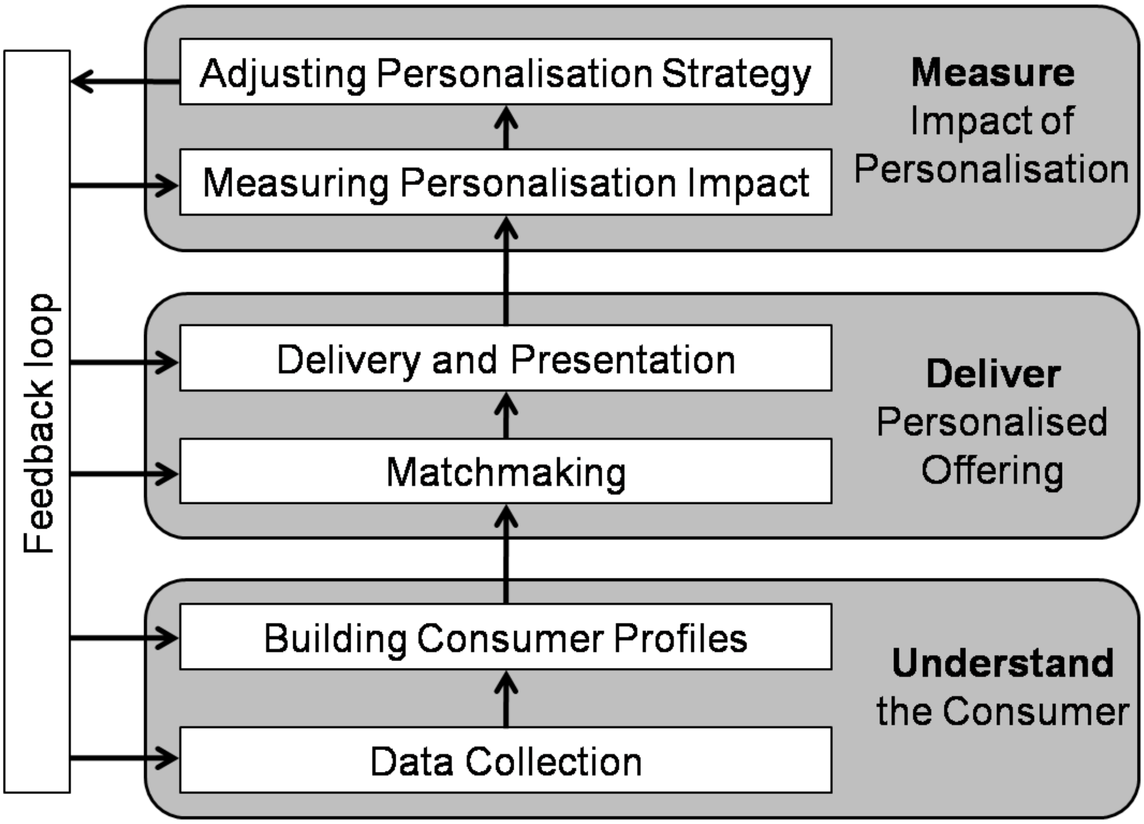

Adjusting Personalisation Strategy
Measuring Personalisation Impact
Measure Impact of Personalisation
Delivery and Presentation
Matchmaking
Deliver Personalised Offering
Building Consumer Profiles
Data Collection
Understand the Consumer
Feedback loop

Figure 3.1: Personalisation process. *Adapted from [2].*

3.2.3 Personalisation: economic results

Personalisation has been widely used in e-commerce. The value that personalisation techniques bring to the business has long been acknowledged by companies such as Amazon, where the founder Jeff Bezos attributes a large part of Amazon's success to the implementation of personalisation in their e-commerce store [72].

The use of personalisation to guide customers and suggest related elements has had a positive impact on the customer [147] and has also been proved profitable, as shown in the analysis made by *"Contact Center World"* [29], concerning the increase of personalisation-attributable revenues by region, as shown in Figure 3.2.

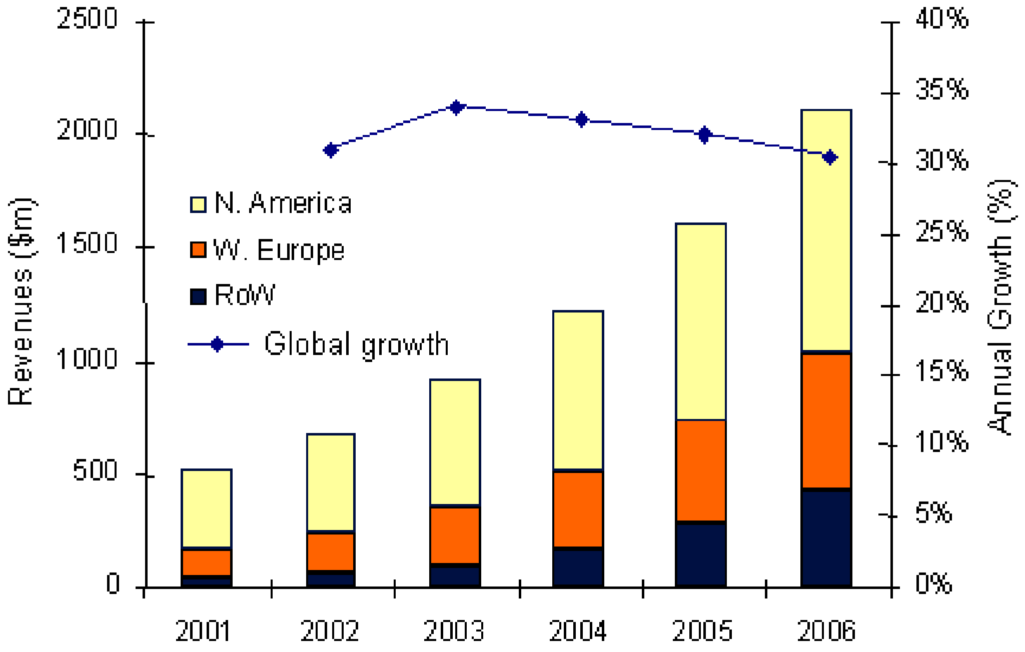


Figure 3.2: Global personalisation revenues by region ($m) and growth, 2001-2006. Note. RoW = Rest of World [29].

One of the main advantages that personalisation offers business is its dynamism and adaptability. Customers' preferences change with different seasons and fashions, and an iterative personalisation process is capable of adapting to match the changing demands [95, 67]. E-tailers can benefit from the dynamism that personalisation provides whilst still allowing the detailed collection of information related to the customer's shopping behaviour. Furthermore, providing dynamic content to customers has been one of the primary reasons for its adoption and, in some cases, its popularisation [42, 63, 122]. The success of Amazon, combined with Jeff Bezos' conferences about the benefits of personalisation [21], has been another motivating factor for e-tailers to include personalisation as part of their business structure, even allowing for the monetary investment and implementation time that doing so within an e-commerce site requires. Therefore, e-tailers desiring to enjoy the benefits that personalisation can deliver must balance their requirement for improved knowledge of their market with the investments required to personalise it [37, 57, 63].

In spite of the business advantages that collecting information from the customer provides, or the benefits that tailoring can offer towards improving the business shopping strategy, or even the influence that personalisation can have in increasing customer stickiness, the major drawback of personalisation is its impact in the customer's privacy. Privacy issues relating to personalisation will be discussed at length in Section 3.5.

3.3 Personalisation: a technical perspective

Personalisation can be defined as "any action that tailors the Web experience to a particular user, or set of users"[92][p.43] and it "[takes] advantage of the knowledge acquired from the analysis of the user's navigational behaviour (user data)" [38][p.1]. Personalisation therefore tailors to a user (or users) based on the results of the analysis done on that user's (or other users) navigational behaviour. As a useful technique, personalisation has been employed in relation to a number of Web applications, including those involved in e-commerce. For example, Amazon presents their customers with a list of recommendations based on their previous purchases (tailoring the Web experience to that particular user), but at the same time, they also provide recommendation lists with choices popular with other customers who selected similar items (taking advantage of the knowledge acquired from the analysis of the user's navigational behaviour).

Using personalisation techniques, a user's Web navigation can be tailored to present a set of results extracted from other users that share similar Web browsing behaviours. The presentation of such personalised information can assist users in attaining their objectives in a faster and easier way. While in this way the results of personalisation can be presented to guide a more focused navigation, recommendation lists can conversely broaden the customers' variety of choices.

Customisation, which is often confused with personalisation, is also related to the tailoring of the Web experience to the particular user, but unlike personalisation, in customisation human participation is required to define the parameters to be used in customising the Web page [31]. Hence, while personalisation is achieved by analysing browsing information collected from the users' interactions with the Web site, customisation takes a more direct and more static approach, presenting users with a set of parameters to choose from.

Personalisation, as discussed, requires little, if any, human participation and generally includes the following phases:

- Collection of information,

- Analysis of the collected information, and
- Tailoring the analysed information to the user,

These phases, shown in Figure 3.3, will be examined next.

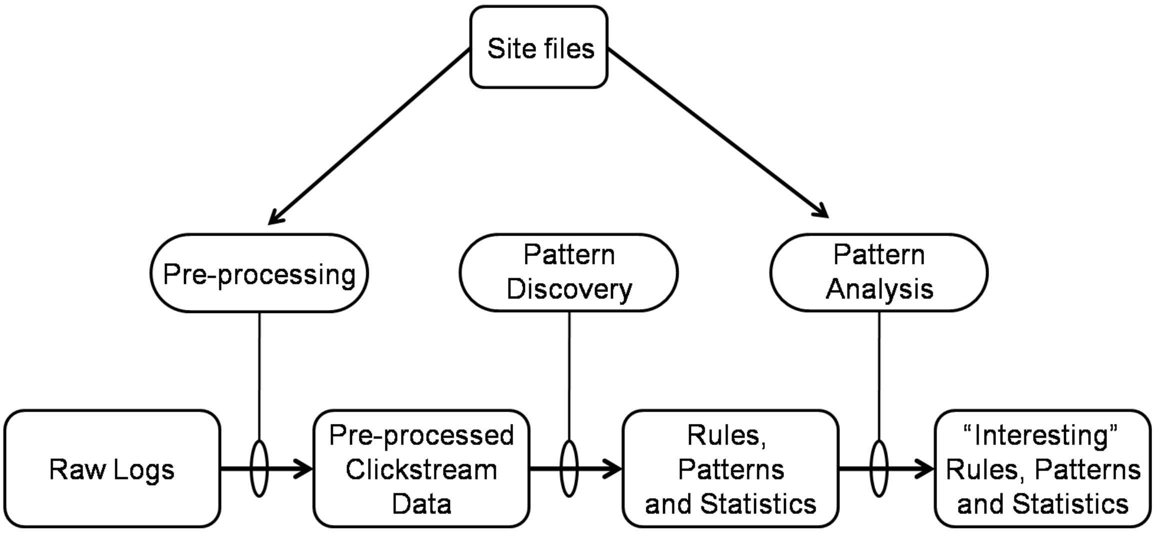
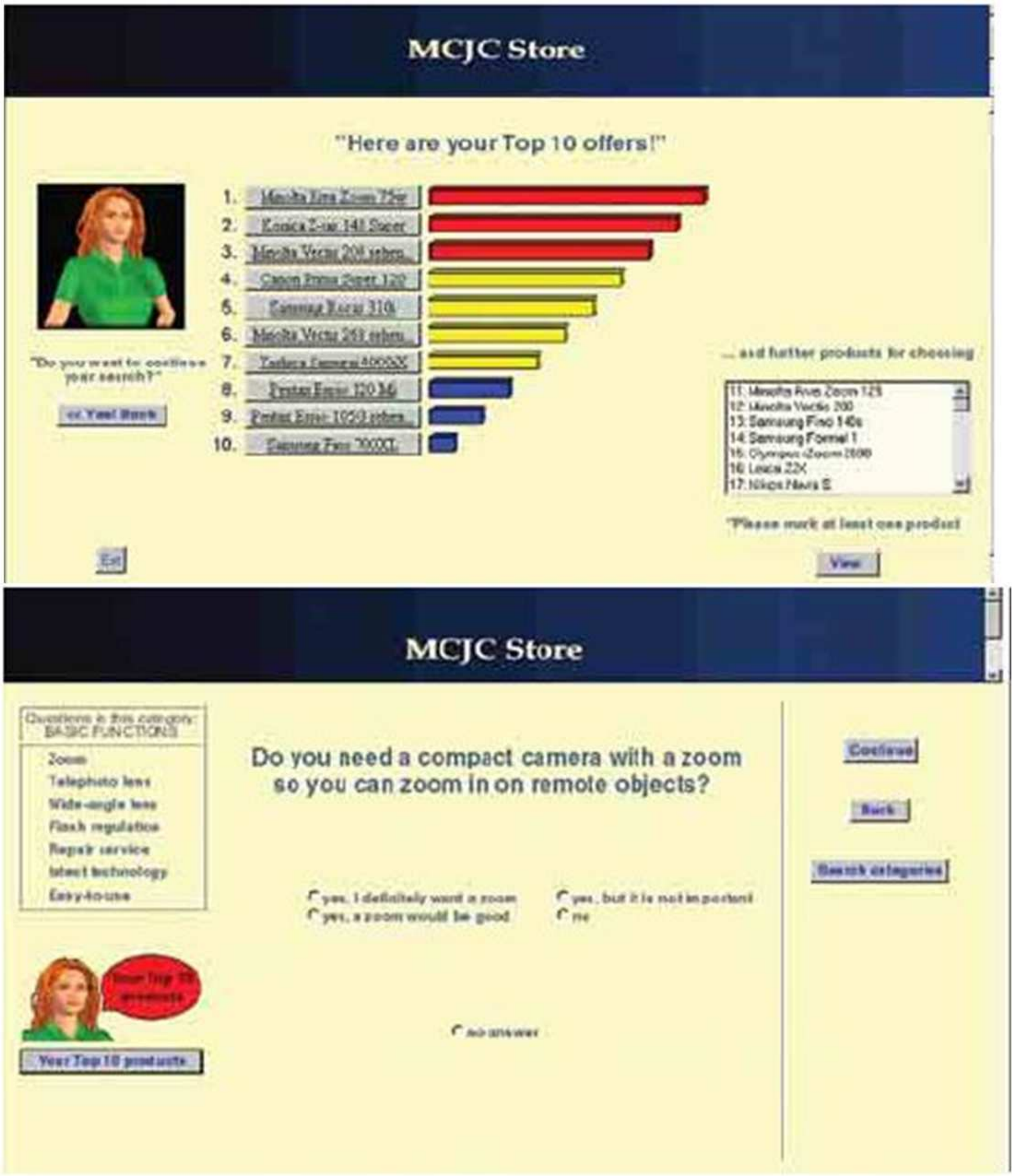

Site files
Pre-processing
Pattern Discovery
Pattern Analysis
Raw Logs
Pre-processed Clickstream Data
Rules, Patterns and Statistics
"Interesting" Rules, Patterns and Statistics

Figure 3.3: Personalisation — Web usage phases (also belongs to the *understand* phase [2]). *Adapted from [142].*

3.3.1 Collection of information

The collection of information comprises two main approaches, namely *explicit* and *implicit*. In the *explicit* approach, which is commonly used for customisation, customers are asked to provide their own preferences either by completing forms or selecting from interactive elements on the screen. Software agents are commonly used to assist the interactive collection of information, such as the one shown in Figure 3.4 [12, 125, 66].



MCJC Store

"Here are your Top 10 offers!"

1. Minolta Riva Zoom 75w
2. Konica Z-up 141 Super
3. Minolta Vectis 20[illegible]
4. Canon Prima Super 120
5. Samsung [illegible] 310i
6. Minolta Vectis 2[illegible]
7. Yashica Samurai 4000IX
8. Pentax Espio 120 Mi
9. Pentax Espio 105G [illegible]
10. Samsung Fino 700XL

"Do you want to continue your search?"

... and further products for choosing

11: Minolta Riva Zoom 125
12: Minolta Vectis 200
13: Samsung Fino 140s
14: Samsung Formel 1
15: Olympus iZoom 2000
16: Leica 22X
17: Nikon Nuvis S

"Please mark at least one product

MCJC Store

Questions in this category:
BASIC FUNCTIONS

Zoom
Telephoto lens
Wide-angle lens
Flash regulation
Repair service
latest technology
Easy-to-use

Do you need a compact camera with a zoom so you can zoom in on remote objects?

○ yes, I definitely want a zoom
○ yes, a zoom would be good
○ yes, but it is not important
○ no

○ no answer

Figure 3.4: Software agents assisting interactive collection of information. [16].

One of the advantages of the explicit approach is evidently that customers can state their own preferences in a clear, swift and uncomplicated manner. This means of collecting information is also relatively inexpensive and less prone to misinterpretation. Users feel assured that their opinion is taken into account and that they remain in control of the navigation. The main disadvantage of the *explicit* approach is however, its lack of dynamism. For instance, users must be asked to select their options each time a new feature is updated, and so even if, at the beginning, the user felt in control, after the recurrent completion of forms or selecting of options on the screen, an unwanted feeling of incursion into their activities arises and time is wasted [31].

The second approach to collecting information for personalisation, the *implicit* approach, is a dynamic process requiring minimal, if any, participation from the customer, as it involves storing the information generated by the user's activities during their Web browsing. The information is collected using Web logs, packet sniffers (which extract the Web server's usage data directly from TCP/IP packets) and cookies [142, 38].

The collected information, as shown in Figure 3.5, includes [142]:

- Page views — consisting of all the files which contribute to the "visual rendering of a Web page in a specific client environment at a specific point time" [76], such as images, scripts and frames etc. Page views are associated with a single user action such as a mouse-click.
- Click-stream — a sequence of page view requests.
- User sessions — a single user's click-stream sequence throughout his or her Web browsing session, across the entire Web.
- Server sessions — a set of page views within a user session for a particular Web site (also called a *visit*).

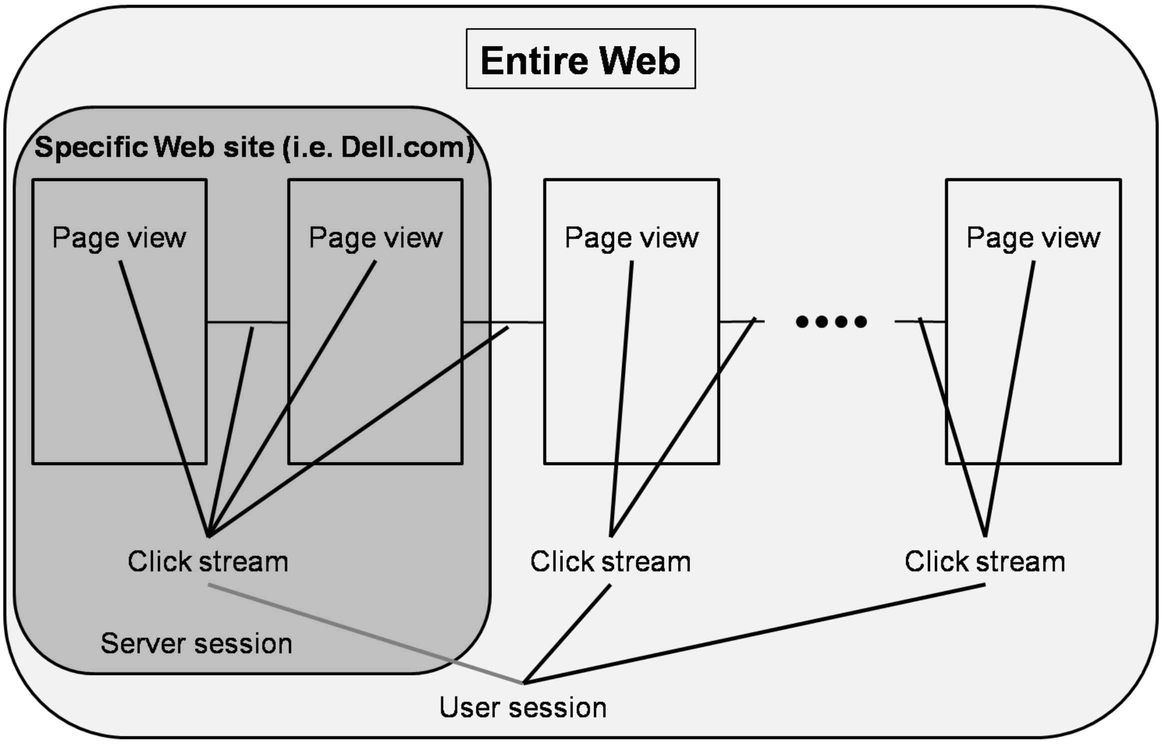


Figure 3.5: Types of information collected for personalisation purposes.

The collection of information, first step in personalisation, is followed by a series of data analyses which detect patterns within the information. The analysis of data which is carried out to find these patterns will be subject to review within the subsequent section.

3.3.2 Analysis of data and pattern finding

To be able to extract useful information from the collected raw data, various methods of analysis are used with the purpose of discovering pertinent patterns. Web mining is one of the extraction methods most commonly mentioned in the literature.

Before obtaining a pattern from the collected data, Web mining categorises and models the data in a system called pre-process. This pre-processing can divide information into the following categories: *usage-based*, *content-based* or *structure-based.*

Usage-based pre-processing involves data that is collected using IP addresses, software agents, and server-side click streams. The usage-based categorisation is related both to the users and their Web page pattern of usage, i.e. the date and time of access of the Web page and IP addresses are collected.

In content-based pre-processing, the content is extracted, and information such as text and graphics are included in this classification. Finally, structure-based pre-processing concerns information such as the data contained in hypertext links between page views and the arrangement of HTML or XML tags within the page [23, 142].

After the data is pre-processed, many diverse methods are used to obtain patterns, including [142, 38, 92]:

- *Descriptive statistical analysis* — The patterns obtained using *descriptive statistical analysis* show elements gained by applying frequency, median, etc. to the pre-processed data. An example of a pattern obtained by this method is the average time a user spends on a specific Web page.

- *Association rule generation* — The patterns derived using *association rule generation*, reveal Web pages that are referenced together and "capture the relationships among items based on their patterns of co-occurrence" [92][p.150]. For instance, by using a rule obtained from a user's navigation patterns it can be determined that a user who visits page *A.html* and *B.html* has a high likelihood (75%) of visiting page *C.html* [92].

- *Clustering* — When *clustering* is used to obtain patterns, a set of items with similar characteristics are grouped. Srivastava [142] divides clustering into *usage* and *page* clusters. *Usage clustering* groups together users exhibiting similar browsing patterns, such as demographics, and the patterns revealed by usage clusters can be applied for the purpose of market segmentation. *Page clustering* associates pages with related content.

- *Classification* — The method of discovering patterns by defining classes and mapping data into them is called *classification.* For instance, customers buying from a specific section (such as /books/programming) were in the 20-25 age group and lived in Glasgow. This kind of information can be used as input for demographic studies in market segmentation.

- *Sequential patterns* — The objective of sequential pattern analysis is to trace how items were followed during a Web navigation session. This variety of information would be valuable for presenting specific advertisements to select groups.

- *Dependency modelling* — Developing a model which represents significant dependencies among various variables in the Web domain is the ultimate product of this method. The modelling of Web usage patterns "will not only provide a theoretical framework for analysing the behaviour of users, but is potentially useful for predicting future Web resource consumption" [142][p17]. For example, a model can be built that represents the different stages that a visitor underwent whilst shopping in an online store based on the items he chose. This data could then be applied to analyse the Web site's stickiness.

The results obtained from employing the aforementioned analysis techniques can be applied to determine whether a visitor to an e-commerce site is a potential buyer based on the analysis of the different stages of their visit. The ability to determine the likelihood of a visitor becoming a purchaser is of major importance for the business, since providing options that match the customers' needs can influence their shopping, enhancing the e-commerce site's stickiness.

The results obtained from the patterns allow the creation of rules for the future tailoring of Web experiences, and in the case of e-commerce, may also represent interesting findings to the business. The constructed rules can be used differently according to the objective of personalisation. In the case of e-commerce, this means focussing the customer's options by tailoring the way the information is presented. Another use of these rules is to broaden the customer's options by offering recommendation lists.

3.4 Tailoring the analysed information to the user

Personalisation, as a technique for assisting business, facilitates the acquisition of specific results which feed marketing analysis that, in turn, supports business planning. As mentioned in Section 3.3.2, the personalisation phases gather information, also called raw data, pre-process the information and analyse it to derive certain patterns. These patterns can be used for assisting the business own marketing analysis, be sent to external companies and be utilised to focus customer attention by providing dynamic navigation. A customers' shopping focus can be narrowed by displaying merchandise for them in a specific order or equally it can be broadened by providing recommendation lists. Figure 3.6 shows the process of personalisation and its uses in filtering information and providing recommendation lists. It also shows the patterns obtained when the personalisation phases are used to create recommendation lists, presenting the information either by *filtering* it or *pointing* at it, and by storing the information to be used later as a repository to match the profile generated by other customers with similar browsing activities.

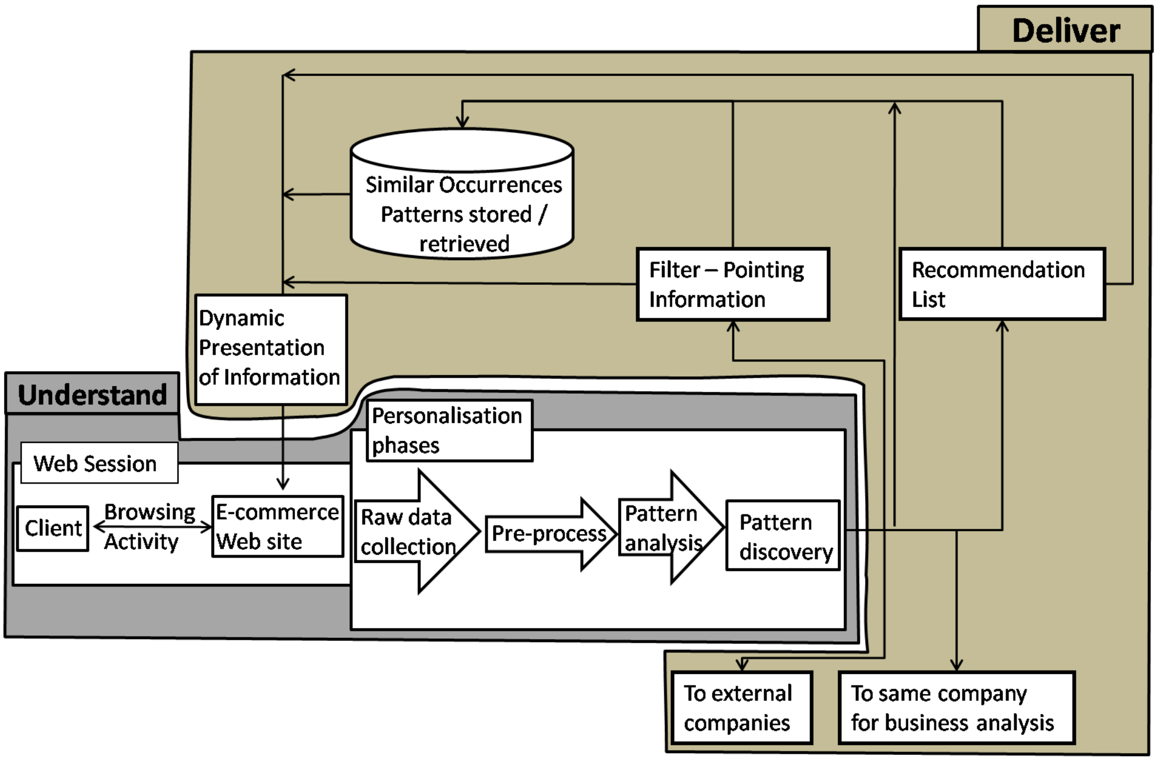


Figure 3.6: Personalisation uses — recommendation list and filtered information are illustrated as part of the personalisation process shown in Figure 3.1.

Two preference-matching mechanisms can be offered to customers: a historical match based on that particular customer's previous browsing activity (or shopping), and a general recommendation profile based on database registers, which contain the patterns obtained from noting the personalisation processes of other customers with similar browsing activities (or shopping).

Section 3.2.2 discussed the different stages in the personalisation process from a business perspective: *understanding*, *delivery* and *measurement*. Figure 3.6 shows those elements of the business perspective and how the personalisation phases, explained above, are contained within the *understanding* stage. It also shows how the results obtained from personalisation methods are tailored to the visitors/customers during the *delivery* stage, however, the process of tailoring results for the customer does not necessarily occur in the same instant they are produced.

Interaction with patterns stored in databases gives rise to a disadvantage when the components have to be updated to incorporate new-found patterns (known as asynchronous cooperation). To solve this problem, Baraglia *et al.* [9] proposed the use of off-line components to forecast the customer's future movements, avoid user intervention on the model-building module, and to provide more dynamism in personalisation.

The rules constructed in the process of personalisation can be applied to the formulation of recommendation lists, which in turn lead to greater dynamism within the browsing and purchasing process.

3.4.1 Recommendation lists

Recommendation lists can be divided into two categories: *content-based* recommendations and *collaborative* recommendations. Balabanović [8] explains the difference by stating that "In *content-based* recommendation one tries to recommend items similar to those a given user has liked in the past, whereas in *collaborative* recommendation one identifies users whose tastes are similar to those of the given user and recommends items *they* have liked"[8][p.66]. An ideal recommendation list would employ both delivery methods, giving the benefits of both. In cases where similar items cannot be found because users have unusual tastes, profiles collected from other users with similar tastes can improve the effectiveness of recommendations [8].

Recommendations can be sorted according to the way the user interacts with the interface into the following [128]:

- *Browsing* refers to users navigating the Web site in search of particular items. For instance in traditional stores, a shopper can look for a specific book aided by shop assistants but also be attracted to other nearby books.

- *Similar item* shows articles that might not have been sought previously or of which customers were unaware. These items are similar to the searched items, one example being Amazon's "Customers who bought".

- *E-mail* keeps the customer informed about the arrival of new merchandise, so they can be the first to buy it. An example can be found in Amazon's "Eye".

- *Text comments* query customer's opinions and subsequently list them next to the item to be purchased. Text can also be used to assist the recommendation list. An example can be found in E-bay's feedback system.

- *Average rating* requests that customers give a numerical ranking to the merchandise. This feature can be used in collaboration with other features, such as text comments, as is the case of Amazon's 'suggestions'.

- *Top-N* presents an ordered list of the preferred, unrated items that a customer may be interested in. This is compiled after the Web site has recorded the likes and dislikes of a customer.

- *Ordered search results* refer to the way in which the merchandise is presented to the customer and will be analysed in greater detail in Section 3.4.2.

Schafer *et al.* [128] present a two-dimensional taxonomy for mapping applications to recommendation techniques. The dimensions are *degree of automation* and *degree of persistence* in recommendations. *Degree of automation* refers to how much information the Web site requires in order to recognise the customer and so provide recommendations, and contains two metrics *ephemeral* and *persistent. Ephemeral* recommendations do not require customers to have had previous sessions these recommendations can be provided in a single session (i.e. recommendations would be presented based on similar items chosen by other visitors instead of his own previous visits or purchases). On the other hand *persistent* recommendations require the customer's identification by the Web site in order to provide recommendations. The recommendations are based on that particular customer's likes and dislikes accrued by logging his previous sessions (including his activities and purchases).

Degree of persistence refers to the explicit effort required on the part of the customer to define his preferences. The taxonomy calls the references that are generated with no customer participation *automatic* and the recommendations that involve the customer's direct participation (customisation) *manual* (for instance the creation of a *wish list*).

From a business perspective, it is of vital importance to turn a visitor into a buyer (stickiness) and to create a relationship with the customer so they will return and continue to buy from that e-commerce site, and recognising this, Schafer *et al.* [128] suggested that the most effective course of action is to provide *persistent* systems which require *manual* effort, since the site would then likely remain in the customer's preference list due to their investment of effort in that Web site. The other means of tailoring information to the customer is by the leading presentation of selected information.

3.4.2 Presentation of information

In general, two forms of presenting merchandise to customers arise during the personalisation of rules: *filtered information* and *pointing information.*

Filtering information involves the presentation of a limited selection of information according to rules constructed from the patterns. From the whole set of information that results from a search, only that which matches the selected personalisation criteria is presented [11, 12, 125]. This method helps customers to locate goods more quickly, and to better focus their shopping. However, from the business perspective, one disadvantage of this method is the restricted presentation of information, especially when it fails to correspond with the customer's expectations [31]. This option also presents a privacy risk, and alarmingly, the amount of money spent on previous purchases could potentially be used as a filter resulting in customers seeing only the most expensive merchandise (if they are perceived to be affluent), a practice which would yield negative commercial results.

Pointing information, on the other hand, presents all the information resultant from customers' searches, and highlights the information conforming to the rules obtained from the customer's profile [66]. This method gives customers an indication of their expected preferred merchandise while still presenting all of the goods. From a business perspective, customers can, at any point in their shopping, narrow their search or add a new item to their shopping basket. That said, however, this method can be taxing for the customer as it necessitates scrolling through all the presented elements in order to locate items matching their specific requirements or needs.

Personalisation techniques provide the means of attracting the customer's attention by offering diverse options that they may not even have realised existed. Commercially, personalisation benefits a business by providing customers with a more engaging and interactive Web site experience, and by also giving the business input data for their supporting market analysis. However, business must be aware that in an age of growing privacy awareness and of trepidation about an impeding ‘big-brother’ state, to obtain customer profiling and the potential misuse of collected information present a persistent threat to the privacy of their customers.

3.5 Privacy Issues

During 2000, just after the e-commerce bubble burst, the unstable nature of the benefits of investing in personalisation became apparent, and at the same time, the risks to privacy were revealed. A statement made by Richard Smith, chief technology officer of the privacy foundation and published by the American Federal Trade Commission concluded; “At e-commerce Web sites, snooping goes by the name of “personalization”” [136].

The collection of information with the intent of building a better profile of the customer forms the basis of a number of techniques used to improve traditional business, as mentioned in Section 3.2.1, however, the information collected and patterns obtained from the personalisation phases, might easily be misused. Examples of this include the false inference of preferences (or “personalised” manipulation of information), such as in the case of dynamic pricing. The subsequent section introduces privacy issues whilst emphasising the consequences that the misuse of techniques such as personalisation can lead to with regard to customer privacy.

One privacy issue involving both personalisation and e-commerce can be found in relation to the matching of customer preferences with the customer. In order to facilitate the tailoring of a Web experience to a particular user (or set of users) those users need to be identified. The user’s identification does not necessarily require the inclusion of personal details (information than can identify a living individual), or sensitive information (information particular to the individual which, if shared, could harm or embarrass them, information such as religious beliefs, physical or mental health condition or sexual orientation). However in cases where customers have previously been identified, the profiles obtained by personalisation techniques can be linked back to them. In this situation, the following privacy issues arise; the *permanence of data*, *false inference* and *manipulation*.

3.5.1 Permanence of data

In a technologically adept environment, where the cost of data storage does significantly impact the business, the permanence of the stored data generates what Stajano calls “denied oblivion” [143]. The privacy risk involved with *denied oblivion* resides in the information remaining stored for an unlimited time. Information such as customer records obtained from loyalty cards, will remain in the company databases until a change in policy or administration occurs (if it occurs at all), and this stored data, which is in itself not necessarily harmful at the time of collection, might become harmful in the future when the use of further-evolved technologies allows for analyses that generates potentially sensitive issues. At the same time, there is a risk of companies having third party associates that store the results of the analysed data for a long period and use them later for different purposes than the original collection purpose. Regulation has been placed to limit the collection and uses of data. However, regulation is subject to *interpretation* which represents a potential risk itself. For instance, it can be said that the patterns obtained after a personalisation process is not the original collected data, and therefore the company is free in its use and disclosure. Regulation related to the preservation of privacy is discussed in more depth in Section 4.4.2.

3.5.2 False inference

The second privacy issue in relation to the identification of customers is *false inference.* Cranor [31] and Kobsa *et al.* [69] identify a privacy risk with computers “figuring things out”. A customer who does online shopping on behalf of another person might generate a profile that does not correspond to their preferences. Furthermore, customers experience difficulties in correcting their profiles, such as the case of the Amazon e-mail list, with recommendations based on previous purchases. Whether this might be seen merely as an annoyance, the “figuring things out” could represent privacy problems especially when the generated information could be misused or could cause embarrassment. For instance, if a customer record shows high purchases of alcohol, it can be deduced that he drinks and enjoys alcohol and would appreciate being presented with special alcohol offers, even if a medical condition forbade him to drink alcohol. The existence of information matching him and a high consumption of alcohol represents a latent privacy risk that would become damaging in the eventuality that the information is disclosed to associated third parties, such as the insurance company that covers his case.

3.5.3 Manipulation

Another privacy issue in relation to personalisation and e-commerce arises when the obtained information is used to manipulate presented information or to mislead customers into false offers. For instance, as part of a marketing experiment, Amazon charged different prices for the same DVD to the same customer, depending on which browser the person was using to access the Web site [123]. Another misuse of personalisation can occur with technology capable of presenting customers with different prices for merchandise based on their own previous purchases ("*dynamic pricing*"). However, dynamic pricing is not the only way of manipulating customers. Images can be used in manipulation as well. Images have a positive effect in the customer's perceptions [160, 48]. Moreover, it has been observed that when users interact with software agents with human characteristics, their reaction becomes overly trusting [56]. Hence, customers can be misled into disclosing information or follow suggestions of human-modelled (*anthropomorphic*) software agents against their better judgement.

3.5.4 The value of information

The misuse or manipulation of information poses a question about the value of the information that e-tailers collect, and the value that customers place on their own information. From the customer perspective, the results of a survey that explored metrics to measure stickiness factors, presented by Oxley *et al.* [109], found that participants classified as important the fact that "my personal information is kept private". However it had little relationship to the stickiness metric. Furthermore, participants valued more highly the amount of information about products or services shown on the Web site. Therefore, information about products or services, and not privacy, were prized factors that would make a visitor want to purchase.

On the other hand, from the business perspective, the Internet provides an intangible media where customers have the opportunity to abandon a particular e-commerce Web site with a single click. Therefore, detailed information about customers, including their preferences and needs, represents an advantage to an e-tailer over his competitors. The more information that is stored about the customer's requirements, the better opportunities the e-tailer has to propose merchandise or services to match those requirements.

To identify what e-tailers would be willing to trade-off in exchange for customer information, Taylor explored the customer interaction focusing on two privacy regimens: *open* and *closed.* In the *open privacy regimen*, the firm have the right to collect and sell customer information including identity and purchasing habits, and companies have the opportunity to charge higher "experimental prices", while customers remain blind to this practice. In the *closed privacy regimen*, customers have the right to remain anonymous. The findings of the exploration of both regimens showed that firms did better when they had committed to keep the customer's data private, specifically in the cases where customers were aware of privacy issues and policies (closed privacy regimen). On the other hand, firms did well and customers did badly when customers were unaware of privacy issues and policies (open privacy regimen) [149].

Therefore to protect customer's privacy, it is important to promote the creation of awareness in the customers, so that they can obtain the biggest benefits from personalised e-commerce shopping. Privacy and the efforts to preserve and protect it are discussed in Chapter 4.

3.6 Conclusion

The use of personalisation has proved profitable for e-commerce. It provides extra assistance in supporting business analysis. Personalisation delivers the benefits of processing and of the analysis of large amounts of information, and provides the business with detailed customer-information to feed into its marketing strategy. These results can also be used to add dynamism to e-commerce Web sites, to increase the stickiness factor, to provide influential elements to the customers' choices and shopping, and finally, it can be used as part of the business strategy to encourage shopping. At the same time, personalisation can provide elements to engage customers in a more participatory role [131, 9].

Regardless of the potential benefits that personalisation can deliver in e-commerce, privacy loss is considered to be a major drawback. The next chapter explores privacy as a concept, while Chapter 5 presents the concept of a privacy protecting environment where e-commerce personalisation can be implemented while respecting the customer's need for privacy.

Chapter 4

Privacy

4.1 Introduction

Privacy issues arise on an everyday basis information about people is being collected, stored and analysed without their knowledge. Its indiscriminate retrieval from the Internet represents a risk to privacy, such as the case where a search on the Web resulted in the disclosure of 2600 CIA employees' identities, including the location of several of the agency's covert workplaces within USA [150]. This risk is more extreme when users trust that their information is being protected and are not aware that collection has taken place. For example, consider the case of *AOL searcher No. 4417749* [10]. During a breach of information, the number used to protect the user's identity while browsing, was linked to a 62 year-old widow, Thelma Arnold, who lives in Lilburn (Georgia in the USA). Three months of search-related data was accessible to the public. Ms. Arnold was astonished to see that her browsing habits, in topics as variant as "numb fingers" to "60 single men" to "dog that urinates on everything", were observed and stored. After this revelation concerning lack of privacy, she reported being left disillusioned and planned to drop her AOL subscription [10].

This chapter sets a context for privacy and related concepts, such as confidentiality and trust, in Section 4.2. Section 4.3 divides privacy issues into voluntary, involuntary and inadvertent disclosure. Finally, approaches taken to preserve privacy, and each approach's advantages and disadvantages, are presented in Section 4.4.

4.2 Related concepts and definitions

4.2.1 Privacy

The concept of privacy is as malleable as is the way humans perceive it, and so is its definition. Several authors have proposed definitions focusing on particular aspects of privacy, but there are also studies which claim that defining privacy is, as yet, an unresolved issue [45]. Hence, it is the aim of this section to explore the facets of privacy and to try to provide a foundation for this research by proposing a definition which will serve for the rest of the document.

Definitions of privacy can be found in standard dictionaries, encyclopaedias and the literature of a number of organisations. From the legal perspective, the importance of defining privacy is related mainly to the need to regulate it. It is not uncommon that due to the lack of a proper privacy definition, privacy issues that go to court result in an unfavourable outcome for the affected claimant [45]. However, for the purpose of this research, defining privacy is focused on delineating its significance and extent. This will help to contextualise privacy and to understand its impact on and relevance to e-commerce.

The Oxford dictionary online (OED), defines privacy as:

Definition 4.1 ***"The state or condition of being withdrawn from the society of others, or from public interest, seclusion."****[101]*

This definition agrees with other dictionary definitions, for example the definition given by Princeton University states that privacy is:

Definition 4.2 ***"The quality of being secluded from the presence or view of others"****[115]*

Definition 4.3 ***"The condition of being concealed or hidden"****[115]*

These definitions identify two main aspects of privacy; the first refers to the affected person and the right to establish a separate space; and the second refers to the society and the limitations of others' access to the person's space. Hence, these definitions work together to formulate an idea of a frontier between a person and the surrounding environment. Therefore, it can be said that the context given by the previous definitions is related primarily to the delimitation of the person's boundaries.

From a different perspective, organisations such as "*Privacy International*" consider privacy as a fundamental human right, linked with human dignity. They define privacy as:.

Definition 4.4 ***"the desire by each of us for physical space where we can be free of interruption, intrusion, embarrassment, or accountability and the attempt to control the time and manner of disclosures of personal information about ourselves."*** *by Robert Ellis Smith, editor of the Privacy Journal[116]*

Ellis' definition goes beyond the OED and Princeton definitions as it specifies privacy in terms of a physical space. He also specifies the activities that can be protected within that space. The definition concludes by giving control over any disclosure of personal information to the person.

For The Calcutt Committee in the United Kingdom [116], privacy's definition is considered a right and focuses on protection against intrusion. This definition has a legal orientation:

Definition 4.5 ***"The right of the individual to be protected against intrusion into his personal life or affairs, or those of his family, by direct physical means or by publication of information"****[116]*

In the same (legal) context, efforts to define privacy can be traced back to 1890 as evidenced by the Harvard Law Review publication "The right to privacy" [157]. This publication raises the issue of photographers taking 'instantaneous photographs' without previous consent, and considers it a clear invasion of the person's privacy.

Definition 4.6 *Judge Cooley refers to privacy as* ***"the right to be let alone".****[157]*

Definitions 4.5 and 4.6 focus on the protection of the individual's space.

While organisations define privacy by focusing on the concept itself, others delineate privacy based on related terms and contexts where privacy can be found. In this way, privacy is associated with autonomy, dignity, anonymity, freedom, liberty, control and consent [45], as well as the determination of a boundary.

Definition 4.7 ***"Invasion of privacy is the transgression of that boundary"****[45].*

Finally, according to Privacy International, privacy can be associated with four main concepts [116]:

- **"Information privacy"**, also called data protection, refers to the withholding of the information collected about a person and the regulation of that collection. Any records such as bank account, health or government records fit into this category.

- **“Bodily privacy”**, is concerned with physical tests, including any medical sample taken from the person’s body, i.e. blood samples, DNA and any genetic or medical tests.
- **“Privacy of communications”**. All communication media is included in this category, regardless of the technology. Mail, e-mail, telephone, fall into this category.
- **“Territorial privacy”** deals with the limits of intrusion. These limits can be domestic, work, surveillance cameras, etc.

There are some central ideas about privacy that can be distilled from the definitions above, which are:

- A *physical space* can be defined, in which the person can:
  - set boundaries;
  - be concealed from society; and
  - be protected against unauthorised intrusion.
- The subject should have control over the *disclosure* of personal information.
- The person should be left alone, and receive the same protection for their family.
- Privacy can be related to the following terms:

  - Autonomy - Dignity - Anonymity - Freedom
  - Liberty - Control - Consent
- Finally, privacy can be related to the following contexts

  - Data protection - Bodily - Communications - Territorial

Therefore, it can be seen that, given the variety of perspectives, defining privacy is not a trivial task. To ground this work, the following privacy definition will be used throughout:

Definition 4.8 ***Privacy is the faculty and right that a person has to define, preserve and control the boundaries that limit the extent to which the rest of society can interact with or intrude upon. At the same time, he or she retains full control over information generated by and related to him or her.***

Definition 4.8 proposes privacy as a human right and gives the person the control and responsibility over the delimitation of the boundaries that society can access or intrude. The definition also proposes that people, by the mere fact of existence, possess information that defines them, and the disclosure of this information should remain in the person's control. Finally, although this work focuses only on online privacy, the definition also covers the importance of the control over body information and any related information that can be extracted or deduced from it, putting the person in control of that information and its disclosure.

To summarise, this work proposes that privacy should keep the person in control of three categories; "*control over disclosure*", "*control over body / personal information*" and "*the right to be left alone (boundaries)*". These categories are illustrated in Figure 4.1.

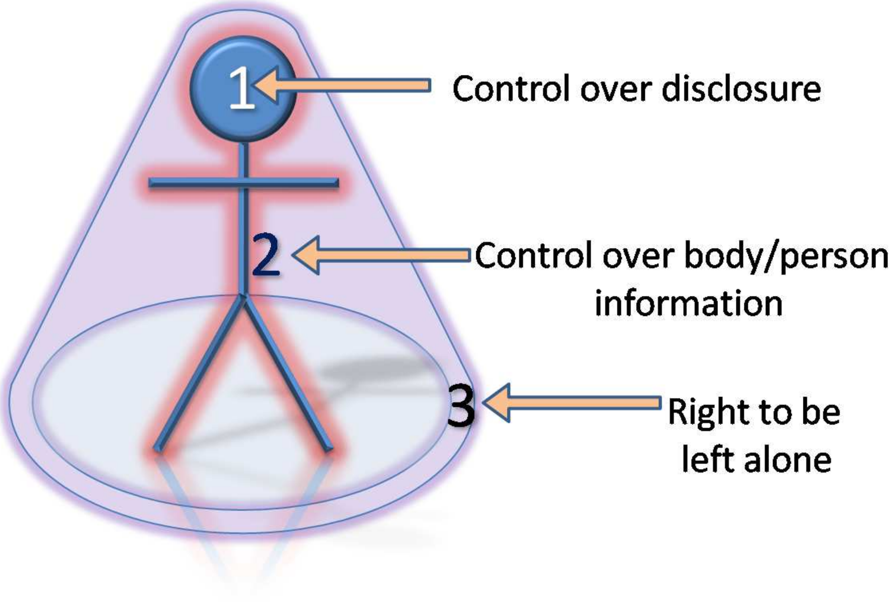


Figure 4.1: Privacy categories

However, even if privacy is jealously guarded, and there is careful and limited disclosure of information, as soon as the information escapes the owner of the information, that person is no longer in control of the information and can only trust that the disclosed information will be used for the correct purpose. The use or misuse of disclosed information by others involves a different concept, *confidentiality*, which is explained next.

4.2.2 Confidentiality

Privacy and confidentiality are related concepts, and are often confused. Alexander [4] explains the difference between confidentiality and privacy as follows:

> Privacy *"denotes a zone of inaccessibility of mind or body, the right to be left alone and to maintain individual autonomy, solitude, intimacy, and control over information about oneself"*

while confidentiality:

> *"concerns the communication of private and personal information from one person to another"*

These concepts coincide with those given by the British Standard 7799 [34], which states:

> *"Confidentiality: ensuring that information is accessible only to those authorized to have access"*

Therefore, hereafter within this work,

Definition 4.9 *Confidentiality will be associated with* ***the preservation of the secrecy of personal data disclosed by another person****.*

Together with privacy and confidentiality, another concept that can be related is trust.

4.2.3 Trust

The definition of trust is as variable as the perspectives of the research concerned; therefore a general concept of trust is difficult to define [90]. McKnight *et al.* [90] explore an "interdisciplinary model of high-level trust concepts", and divides them into; *dispositional trust*, *institutional trust* and *interpersonal trust.*

Dispositional trust comes from psychology, and states that "actions are moulded by certain childhood-derived attributes that become more or less stable over time" [90][p41] and "means that one trusts other generally"[90][p42]. *Institutional trust* comes from sociology, and states that "behaviours are situationally constructed"[90][p41] and "means that one trusts the situation or structures" [90][p42]. And finally, *Interpersonal trust* reflects "the idea that interactions between people and cognitive-emotional reactions to such interactions determine behaviour" [90][p42], and "the direct object is the specific other individual one trusts" [90][p42].

On the other hand, within an e-commerce context, the analysis of literature related to trust performed by Chen *et al.* [22] and Harrison *et al.* [90] have concurred on the following facets of trust:

Overall trust, "general trust which is not related to a specific behaviour of the other party, or any component of trust" [22][p305],

Competence, companies fulfilling their promises to the consumers, and having sufficient safeguards in place to fulfil them,

Integrity, companies acting consistently, reliably and honestly when fulfilling their promises and

Benevolence, "the probability a company holds consumers interests ahead of its own self-interest and indicates sincere concern for the welfare of the customers" [22][p305].

McKnight *et al.* [90] have integrated these facets (*overall trust*, *competence*, *integrity* and *benevolence*) into the "interdisciplinary model of high level trust concepts" (*dispositional*, *institutional* and *interpersonal* trust) and have adapted it to e-commerce. The integration is shown in Figure 4.2.

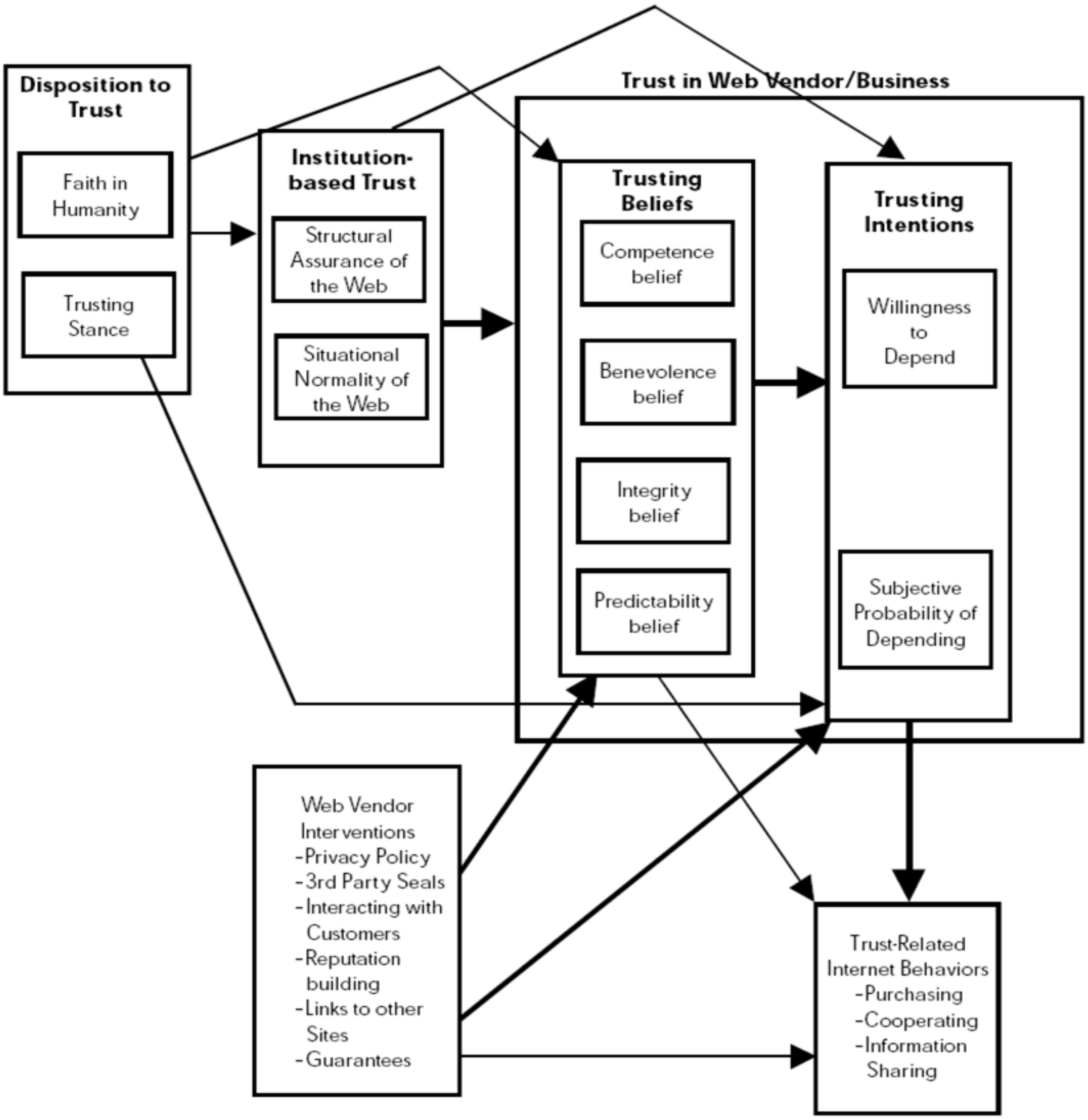


Figure 4.2: A model of e-commerce customer relationships trust constructs. *Adapted from [90]*

Since overall trust, competence, integrity and benevolence are part of the interdisciplinary model of high level trust concepts (*dispositional trust*, *institutional trust* and *interpersonal trust*), a relation between privacy, confidentiality and trust and the interdisciplinary model of high-level trust concepts can be established. Therefore, it can be said that trust and privacy are related to *dispositional trust* when the decision to disclose information is taken. Trust and confidentiality are related to *interpersonal* and *institutional trust*, due to the expectation that the institution, structures and persons will respect the explicit and tacit agreements and the information will not be disclosed to unwanted parties. This relationship is illustrated in Figure 4.3.

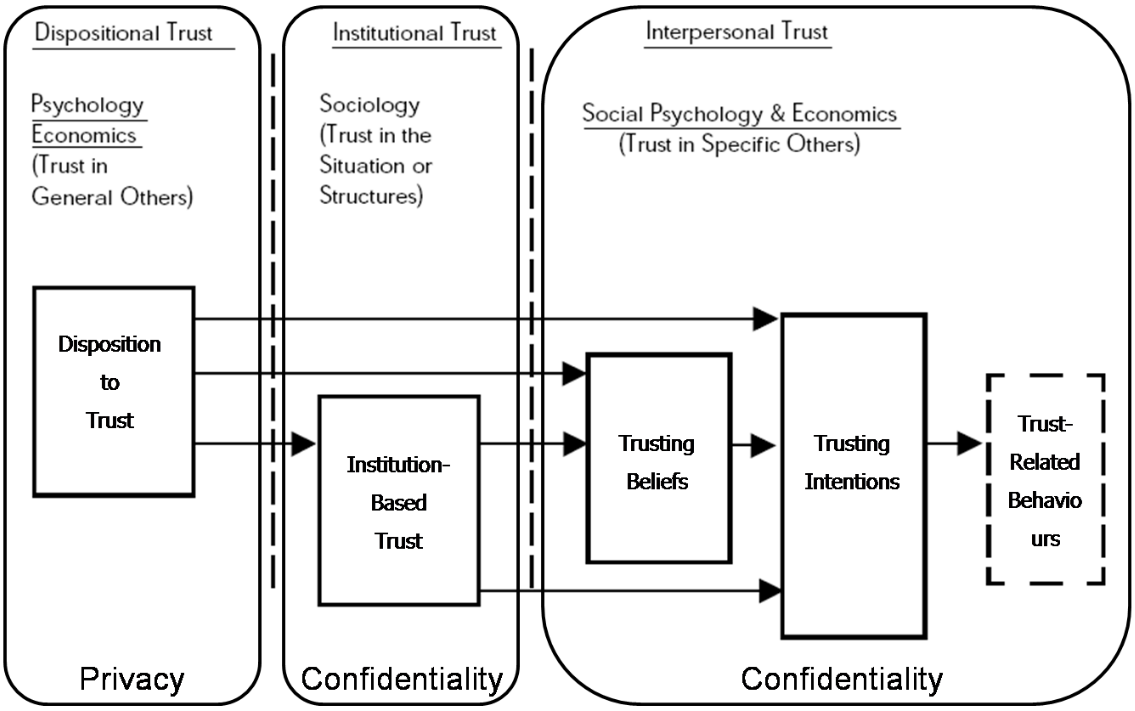


Figure 4.3: An interdisciplinary model of high-level trust concepts and its relation with privacy and confidentiality. *Adapted from [90]*

The process of clarifying the concepts within this work has as main purpose: to identify the issues related either to privacy or confidentiality. By carefully separating them, privacy issues can be isolated and analysed and solutions proposed. The next section presents an overview of privacy issues.

4.3 Privacy - Issues

There are many privacy problems present in everyday life, and technology is just one more factor which can put people's privacy at risk. People have become so used to technology that its presence is mostly unnoticed. The ubiquitous presence of technology in daily life allows the recording of data related to everyday activities. Furthermore, this also is seldom noticeable and users are unaware of the potential misuse of that collected information.

As mentioned in the Definition 4.8, a person has the right to define, preserve and control the boundaries between him and the rest of society. However the information generated by, or related, to a person can either be voluntarily disclosed, for instance when people share their profile on “MySpace” or “Facebook”, involuntarily disclosed, for instance fingerprints collected for admissions in theme parks, or inadvertently (and involuntarily) disclosed, such as the information collected by the use of surveillance methods such as CCTV cameras. These cases present privacy risks, and are discussed next.

4.3.1 Voluntary disclosure

In the voluntary disclosure of information, the disclosure is made to fulfil some expectations, within a certain level of control and with a certain level of awareness of the consequences of that disclosure. The information remains under the person’s control until he or she decides to disclose it. For instance, with the expectation of recovering or preserving health, patients disclose the most personal, private and sensitive information to a medical practitioner in the understanding that the disclosed information will remain under the direct control of the practitioner and the health services and will be used only for the agreed terms (to recover or preserve health). However, it is when the control over that information is lost that the privacy of the patient’s data is violated. For instance, at the end of 2007, nine English NHS trusts admitted losing patient records [98]. In relation to that case, Joyce Robins, from the patient support group Patient Care, said “records can have anything from your ex-directory phone number to your HIV status”. Patients that experienced the loss of their records experienced diminished trust. That decrease of trust concerned practitioners, as expressed by Dr Richard Vautrey, of the British Medical Association, who said “it would be damaging if patients became reluctant to be fully open with their doctors” [98].

On the other hand, on the Internet, Web sites like *Facebook, MySpace, Bebo, Friendster, Flickr* or *Picasa* encourage users to share their information, pictures and experiences with a social network. Such disclosure sometimes happens without the realisation of the risk that it represents to the participant’s privacy. For example, a human-resources manager could read through ideas posted in social network forums looking for the job applicants’ political opinions as the first filter of the hiring process. Political, religious or sexual opinions could be taken into consideration when hiring and could affect one’s reputation [156]. However, users retain a certain level of control over their disclosure. Other ways of disclosing information are by keeping a *blog*, participating in virtual worlds such as *Second life, Habbo hotel* or by using instant messaging such as *MS Messenger, ICQ*, etc.

Unfortunately, once information is disseminated via the Internet, retaining control becomes an impossible task. The facility of generating multiple copies of the original information exacerbates the problem. Once the information is part of the public domain, its storage and diffusion can pass into the control of a number of people, not only one trusted person or institution. For instance, the case in which a senator from Alaska asked his staff to gather information as if they were attempting to steal his identity, just to test the ease or difficulty of the process. The collection of information was not only easy, but the results of this search also came up with details about his close family [146]. It is clear that, during the inadvertent disclosure of information, the risks of privacy violations increase. If privacy risks are not perceived, no protection is sought.

4.3.2 Involuntary disclosure

In the involuntary disclosure of information, the reasons for the disclosure are not necessarily clear there is no significant control during the disclosure and, in some cases, there is no awareness of the consequences of that disclosure. This type of disclosure can be done under circumstances where the person faces few other options but to disclose the information. For example, photographs of passengers travelling by plane are taken at boarding time (between checking the ticket and boarding the plane) in some of the UK airports (such as Gatwick, Manchester and Edinburgh) [97]. These photographs are taken in order to verify the passenger's identification at boarding time. While a spokeswoman from BAA said "we introduced the photo-taking as a security measure even before 11 September. The photo is later destroyed" [97]. Passengers felt forced to have their photo taken. It was after £4,000 was paid as compensation to Tim Hedgley for having his photograph taken without his consent at Manchester Airport, that the airport started posting communications to their passengers telling them that they were not obliged to have their photo taken [97]. However, the use of photographs and fingerprint-scanners are proposed as a new security measure to be implemented during 2009 in UK airports with high-traffic national and international terminals (such as Gatwick and Manchester) [81].

4.3.3 Inadvertent disclosure

During inadvertent disclosure, information about people, and, in particular, Internet users, is collected without the person's knowledge of it happening. This collection of information happens with no possibility of control by the person whose behaviour is being tracked and there is thus no opportunity of limiting or preventing it. Different surveillance methods are used to obtain this kind of information. For instance, Radio-Frequency Identification (RFID) tags can be used to track cars or people's movements within stores. The use of RFID tags in a store allows the store to track the paths that customers took during their visit to the store. RFID tags can also speed up inventories [19]. In cars, RFID tags could automate the payment of tolls, 'congestion' charges or the issuing of tickets for violations of speed limits, but could also be used to keep a track of addresses and duration of visits by drivers. RFID is proposed to be used even in passports and the passport information (including detailed personal and biometric information) could be remotely accessed by the reader within a distance of 10 metres [129]. Several privacy risks have been associated with the use of RFID. Razaq *et al.* [118] divides them into; *disclosure* ("dissemination of tag information to any reader that should not read this information" [118][p23]), *denial of service* (tags blocked by unauthorised readers, as a result of malicious attack), *integrity* (unauthorised change of information on the tag or during transmission), and finally, *cloning* ("an unauthorised tag's malicious action results in an alternative device that spoofs a reader into believing that the tag is correctly prompting the reader to exchange information" [118][p23])

Commerce has used diverse strategies to obtain information about the customer's shopping habits. For instance, supermarkets provide their customers with loyalty cards, which assign customers "points" related to the amount and characteristics of purchases. However, loyalty cards are primarily used to collect information about purchases, and they can also be used to match the customer's demographics with their shopping habits, making them easily identifiable. A second generation of loyalty card is a device contained in the trolley, which the customer can use to scan in the bar code of each item to be purchased [41]. During the selection of items, the customer can request the system to check how "healthy" the items are, and if they are not "healthy", there is an alarm that indicates this. While older customers commented on their resistance to the use of this new feature, younger customers, from 18-34 years old, who were less concerned about the disclosure of their information, were more enthusiastic about adopting this technology [41].

In relation to loyalty cards, a privacy issue arises when the collected information is used to identify the customer and the stored records are used against him. For instance, there is the case of Mr. Rivera in Los Angeles, USA. When he began an action to sue Vons store for a kneecap injury due to slipping on spilt yoghurt, he was told that his high alcohol consumption, stored on his records, was going to be shown in court. Mr. Rivera's complaint was not successful [155, 158].

Technology facilitates the easy analysis of the collected information. Chapter 3 discussed personalisation techniques and how information about the customers is collected and analysed to fulfil marketing studies. It also described the possible misuses of the analysed information. Often, customers do not know that their personal information is being collected and are not aware of the risks that this represents [94, 140]. As soon as they realise the collection of their information is taking place or have the feeling that their activities are being tracked, they tend to avoid being in contact with that site as they have lost trust in it [94, 124].

However, among the problems that concern some researchers is that e-customers are apparently eager to give up their private data in exchange for a few benefits [145, 67]. At the same time, regardless of the multiple efforts to preserve privacy and the regulations put in place for this purpose, there are cases when the control over the handling of information relies on third parties, such as the case of outsourcing. For instance, in Pakistan, a woman working for an outsourcing firm tried to blackmail her employer, by threatening to make available to the public the data that she was working with, if she did not receive a higher salary [27]. To prevent these problems, when sensitive data is transmitted to another country, the data processing rules of the origin country are applied with special vigilance [49], or data is made anonymous. Multiple approaches have been proposed to protect the user's privacy. The next section discusses these approaches.

4.4 Related work

Based on the means used by non-profitable organisations and business in their attempts to preserve privacy, this work has identified three different approaches: raising *awareness*, *regulation* and the use of *technology* (ART). Strategies for preserving privacy, within the aforementioned categories, are presented next.

4.4.1 Raising awareness

There is a growing awareness of the problems that the loss of privacy can bring to users in general. The media has publicised diverse privacy issues such as *identity theft*, and this publicity has had the effect of creating and disseminating awareness. Therefore, it can be said that the creation of awareness is a privacy preserving mechanism. Organisations such as "*Privacy International*" bring together a number of privacy experts sharing, among other things, the aim of raising the level of privacy-awareness [40]. They also work towards establishing privacy measures throughout the world and facilitate the flow of information about privacy outside the group. Their effort is oriented towards monitoring the effectiveness of the privacy protecting measures, assessing the impact of technology in privacy, and monitoring the nature and extent of privacy violations country by country (among others).

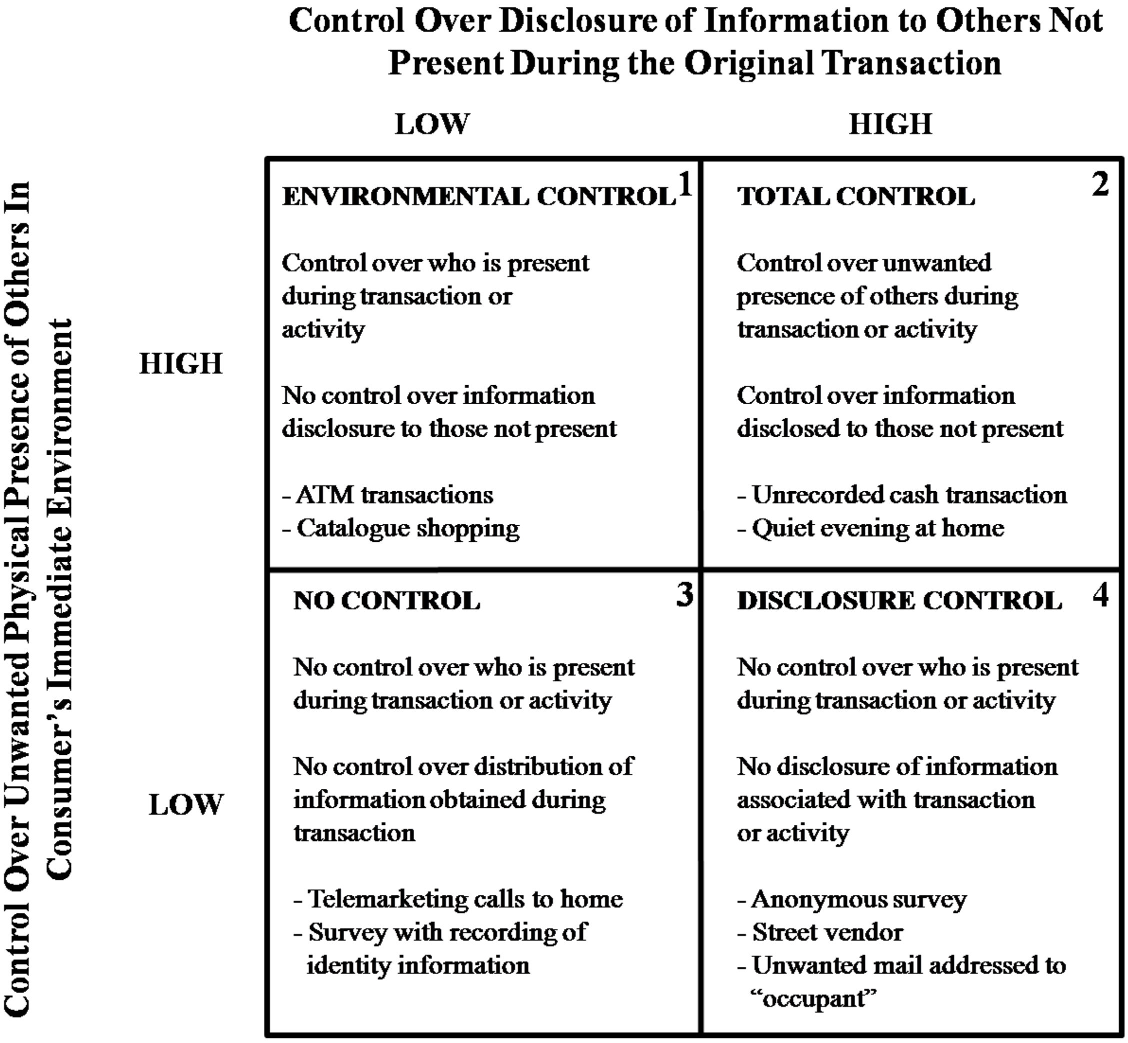
On the other hand, risk awareness has been linked to a reduction in the level of trust and an increased demand for control, especially in relation to consumer privacy [106]. In relation to consumer privacy, four control states have been identified; *total control*, *environmental control*, *disclosure control* and *no control* [53, 106]. Whereas consumer privacy, defined as "*the consumer's ability to control (a) presence of other people in the environment during a market transaction or consumption behaviour and (b) dissemination of information related to or provided during such transactions or behaviours to those who were not present* [53][p152]", typifies *consumer control* over information disclosure and the environment in which a consumer transaction occurs. *Environmental control* can include the use of data mining for personalisation activities and *information control* concerns to information being used for purposes other than those originally agreed. Therefore, these four levels of control, shown in Figure 4.4, are: [53, 106].

Total control, those that have full/total control over their disclosed information and environment.

Environmental control, those that have little control over their disclosed information, but full control over the environment.

Disclosure control, those that have full control over their disclosed information, but no control over the environment, and

No control, those that have no control over their information, or the environment.



Control Over Disclosure of Information to Others Not Present During the Original Transaction

Control Over Unwanted Physical Presence of Others In Consumer's Immediate Environment

	LOW	HIGH
HIGH	**ENVIRONMENTAL CONTROL** 1 Control over who is present during transaction or activity No control over information disclosure to those not present - ATM transactions - Catalogue shopping	**TOTAL CONTROL** 2 Control over unwanted presence of others during transaction or activity Control over information disclosed to those not present - Unrecorded cash transaction - Quiet evening at home
LOW	**NO CONTROL** 3 No control over who is present during transaction or activity No control over distribution of information obtained during transaction - Telemarketing calls to home - Survey with recording of identity information	**DISCLOSURE CONTROL** 4 No control over who is present during transaction or activity No disclosure of information associated with transaction or activity - Anonymous survey - Street vendor - Unwanted mail addressed to "occupant"

Figure 4.4: Taxonomy of privacy states. *Adapted from [53].*

On the other hand, a series of privacy concerns have led Westin to create "*privacy indices*" [75]. These indices, obtained from a series of surveys that aimed to explore privacy concerns, were obtained by dividing participants into three main groups: *Fundamentalist*, *Pragmatic* and *Unconcerned*.

- The *Fundamentalist* group consists of people who distrust organisations asking for their personal information, are worried about computerised-gathered information and its uses, and favour regulations (revised and new measures) to protect their privacy. Members of this group actively use controls to protect their privacy.

- The *Pragmatic* group weigh the benefits of protection and regulation against the amount of information they are prepared to disclose, believing that trust should not be freely given but "earned", and seek to have opt-out options against the indiscriminate collection of information.

- Finally, the *Unconcerned* group trust the collection of information by organisations, are not in favour of new privacy regulations and do not use controls to protect their privacy.

Over time, a change in the privacy perceptions has been observed. The number of participants falling into the *Unconcerned* category has decreased, the *Fundamentalists* group has maintained its numbers, while the number of *Pragmaticists* has increased. Westin attributes this change to the increase of knowledge about technology and the awareness of protection methods [75].

Based on Westin's observations, the creation of awareness is an important factor which changes the user's privacy perceptions. Hence, an approach to e-commerce is needed in which customers can have access to elements which raise privacy awareness and at the same time gives them control over what to disclose and under which circumstances the disclosure should occur. On the other hand, it has been suggested by Olivero *et al.* [106] that customers who know that their information has a value for marketing purposes, should be empowered with the capability of a trade-off between their information in exchange for some benefits. Therefore, to have a privacy-protecting approach to e-commerce, in which the customer has the knowledge of the value of their information, and is in full control of the disclosure, would be a valuable asset for the privacy concerned customers and would provide protection to the "unconcerned" group. At the same time, the proposal of empowering customers and creating awareness has been envisaged as a necessary step towards the preservation of privacy [53]. This is confirmed by Olivero *et al.* [106] who found, following analysis of literature and interviews, that the increase of risk awareness in the customers reduced their level of trust, and that their demands for controlling their information and its disclosure increased.

The second approach towards preserving or protecting privacy is related to regulation, and this is discussed next.

4.4.2 Regulation

Privacy issues have been present throughout history, a new aspect of which is *identity theft*. The U.K. home office defines *identity crime* as "a generic term for *identity theft*, creating a *false identity* or committing *identity fraud*" where a *false identity* can be a fictitious or altered identity.

Identity theft "occurs when sufficient information about an identity is obtained to facilitate *identity fraud*, irrespective of whether, in the case of an individual, the victim is alive or dead." And "*identity fraud* occurs when a false identity or someone else's identity details are used to support unlawful activity, or when someone avoids obligation/liability by falsely claiming that he/she was the victim of identity fraud".

However, according to Clarke [25], "human identity is a delicate notion which requires consideration at the levels of philosophy and psychology. Human *identification*, on the other hand, is a practical matter". This approach is used by Sproule *et al.* [141] to define identity theft using a conceptual model, shown in Figure 4.5, which makes a clear division between identity theft and identity fraud.

Identity theft, according to the U.K. home office, includes the activities related to the collection of personal information and the development of a false identity, while *identity fraud* is the use of a false identity to commit crimes. Therefore, according to Sproule *et al.* [141], identity theft is "the unauthorised collection, possession, transfer, replication or other manipulation of another person's personal information for the purpose of committing fraud or other crimes that involve the use of a false identity" and identity fraud is "the gaining of money, goods, services, other benefits, or the avoidance of obligations, through the use of a false identity".

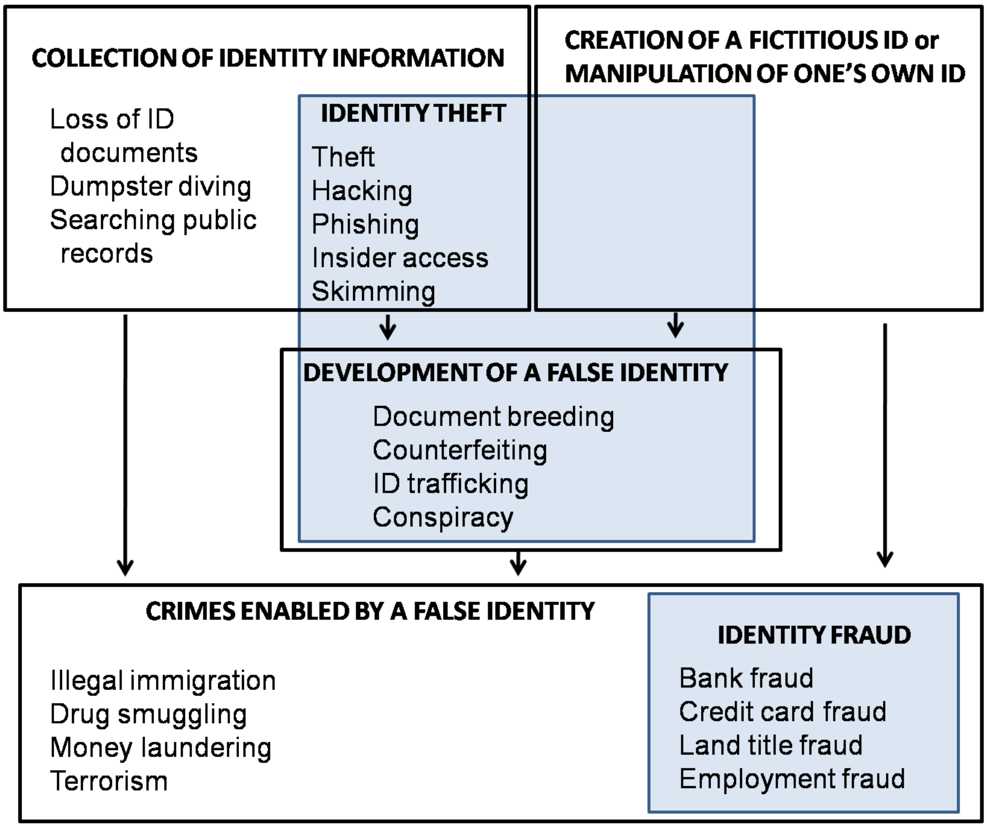


Figure 4.5: Identity theft and identity fraud within the conceptual model. *Adapted from [141].*

The repercussions of identity theft and identity fraud are a myriad. For instance, according to the Home office, the impact of identity theft in the UK economy represented a loss of £1.2bn in 2006 [102]. On the other hand, CIFAS ("a not for profit membership association dedicated to the prevention of financial crime and staff fraud" [24]) reports 77,500 cases of identity theft in 2007 alone [24]. Unaware customers are at risk of falling for phishing or spam attacks when using online services, especially with the increase in the use of online services. For example, Internet banking has increased 505% since 2000 [5]. From the diverse identity theft attacks, a particular one emerges concerning child identity theft. While adults can detect identity theft via credit card reports, or unusual activities in a relatively short time, child identity theft can take years to be noticed and resolved. These attacks involve the theft of a social security number or the name and date of birth of a child. The misuse of this information may pass unnoticed for many years and lead to the challenge of proving that the child is indeed the correct owner of that identity [162].

However, identity theft is not the only privacy-related issue. Any unauthorised invasion of a person's moral, intellectual or physical space can constitute a violation of their privacy. Reading somebody else's diaries, opening somebody else's mail or taking unwanted or unauthorised photos all represent privacy violations. Regulation has been attempted throughout history. Milestones in the history of privacy regulation will now be covered. To illustrate the evolution of privacy, examples of cases where privacy issues arose in the UK and the USA are presented, along with implemented regulations. In this history milestones, the USA is used as an example of the evolution of privacy and its regulation in a country with similar culture to the UK, Figure 4.6 shows the time-line of these historic landmarks.

An early aspiration to regulate privacy is evident in the use of the phrase 'The house is one's castle', during a legal case in the United States of America (USA) in 1604. Since then, there have been many privacy related cases, mainly in the USA [117]. During the last decades of the 18th century, the USA's biggest privacy concerns were related to the unauthorised opening of mail. These actions caused the creation of a law which, in 1782, forbade that practice. In 1877 this was extended to forbid government officials from opening mail without a warrant. In 1790, the USA held their first census. The census results were publicly posted. This practice allowed people to verify the correctness of the census content. However, concerns with confidentiality violations resulted in this practice being abolished in 1870.

In 1890, an article was published under the title of the 'Right to privacy', which highlighted the right of privacy related to topics such as reading somebody else's mail, but specially emphasised what they called 'instantaneous photographs' [157]. The article proposed that obtaining and using photographs taken without the previous consent of the person, was a clear invasion of their privacy. Finally, the same publication refers to privacy as the right "to be let alone" [157].

War changed the way privacy issues were perceived. During the First World War, the UK established the use of an identity card derived from the first national register. The purpose of this initiative was to determine the number of men capable of fighting. The use of this first identity card was discontinued in 1919 [3]. In the USA, the Social Security System, which maintained a national register, was established during 1935. In 1939, during the Second World War, the UK returned to using identity cards. On this occasion, there were three particular goals in the issuing of identity cards. The first was the coordination of national

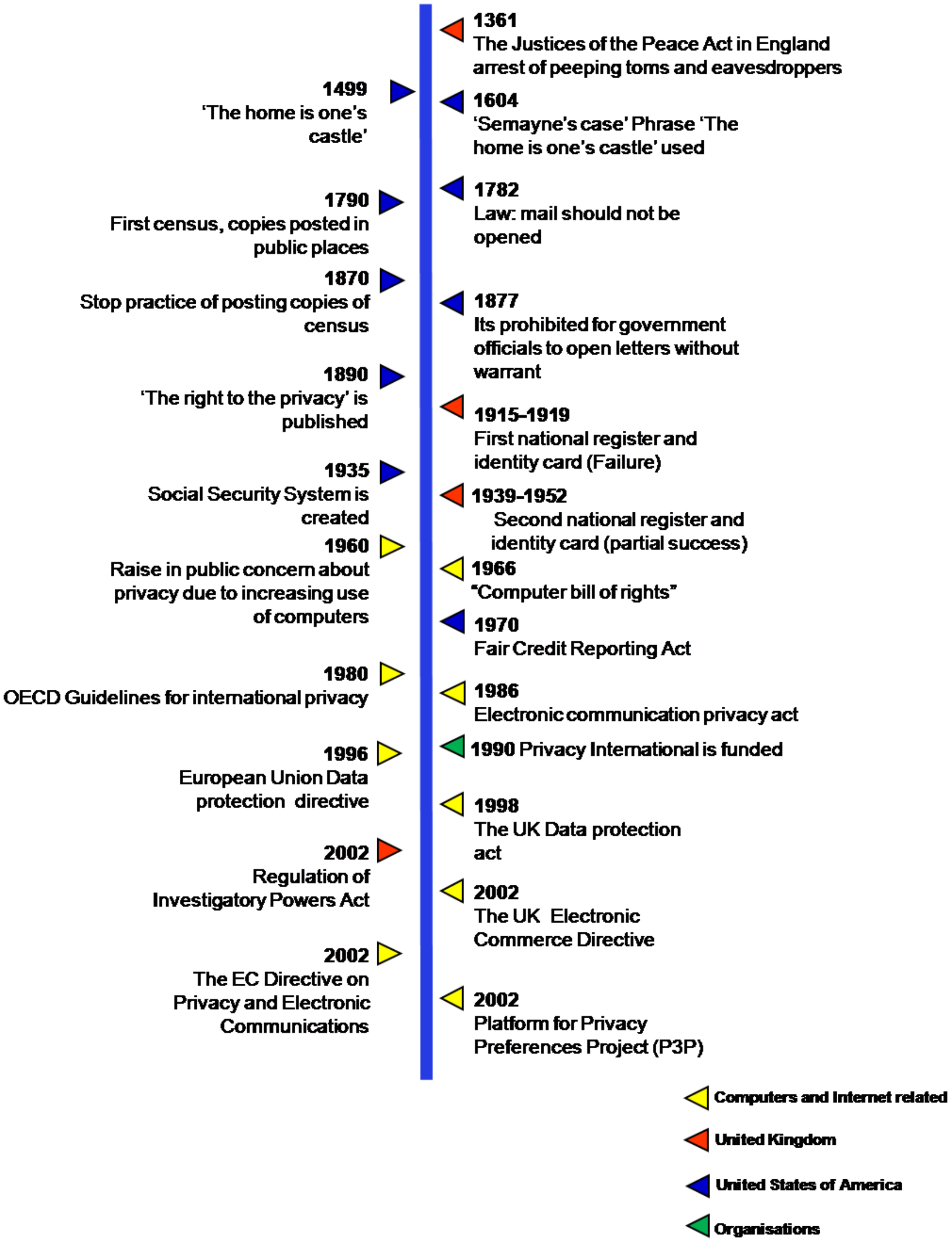

1361
The Justices of the Peace Act in England arrest of peeping toms and eavesdroppers
1499
'The home is one's castle'
1604
'Semayne's case' Phrase 'The home is one's castle' used
1790
First census, copies posted in public places
1782
Law: mail should not be opened
1870
Stop practice of posting copies of census
1877
Its prohibited for government officials to open letters without warrant
1890
'The right to the privacy' is published
1915-1919
First national register and identity card (Failure)
1935
Social Security System is created
1939-1952
Second national register and identity card (partial success)
1960
Raise in public concern about privacy due to increasing use of computers
1966
"Computer bill of rights"
1970
Fair Credit Reporting Act
1980
OECD Guidelines for international privacy
1986
Electronic communication privacy act
1990 Privacy International is funded
1996
European Union Data protection directive
1998
The UK Data protection act
2002
Regulation of Investigatory Powers Act
2002
The UK Electronic Commerce Directive
2002
The EC Directive on Privacy and Electronic Communications
2002
Platform for Privacy Preferences Project (P3P)
Computers and Internet related
United Kingdom
United States of America
Organisations

Figure 4.6: Privacy Timeline.

service, the second was for national security and finally, it was used to implement food rationing. This second identity card system was more widespread than the first. However, it was discontinued in 1950, after Clarence Willcock's refusal to show his identity card to a policeman. The case was sent to court and it was decided that the use of an id card, appropriate in wartime, was an "annoyance" in peacetime [36]. After the war, there were attempts to protect privacy and to create legislation to control it. In one attempt to legislate the collected information, the Fair Credit Report Act was created in 1970 in the USA, which allowed individuals to check and amend any inaccuracy within their credit history.

From the computing science perspective, privacy issues related to computers were openly addressed in the 1960's, when research produced publications related to the weaknesses of the forms in which the information was stored in computers and how it could be misused e.g. in the lack of control over data access. During 1966 a "computer bill of rights" proposed guidelines to the storage and access to the data, i.e. to maintain records of when the data was accessed and by whom [58]. Computer privacy was addressed again in 1980, when the Organisation for Economic Co-operation and Development (OECD) published their first guidelines for international privacy [107].

During the 1990's several efforts to enforce the protection of privacy were made. The organisation *Privacy International* was created during the course of 1990 with the purpose of bringing privacy issues into open discussion [40]. The European Union produced a directive for data protection in 1996. The USA's Federal Trade Commission (FTC) has, since 1998, brought action against companies that violate their own privacy policies. Also during 1998, in the UK, the European Convention of Human Right's human rights act was incorporated into the UK law, along with the Data Protection Act [105], establishing regulations dictating rules regarding the collection and usage of data. The UK's Regulation of Investigatory Powers Act, 2002, establishes under what conditions communications can be intercepted [103]. An electronic commerce directive, issued by the UK during 2002, regulates the commercial activity carried on over the Web [104] and a Directive on privacy and electronic communications was issued by the EC during 2002 to protect the user's privacy [104].

In 2002, as a result of privacy workshops, the "*Platform for Privacy Preferences (P3P) Project*" was created with the main aim of expressing privacy practices in a machine readable way. The World Wide Web consortium — W3C ("an international consortium where Member organisations, a full-time staff, and the public work together to develop Web standards”), has closely related projects such as the P3P project, PRIME (“explores the future of privacy enabled Identity Management”) and TAMI (creates “technical, legal, and policy foundations for transparency and accountability in large-scale aggregation and inferencing across heterogeneous information systems”) [33, 84]

In spite of all these efforts producing regulations to protect privacy, violations to privacy remain ever present. The use of regulation as a privacy preserving approach has two main disadvantages. The first is that the penalty for noncompliance can be applied only after the privacy violation has occurred, and the second is that the regulation, and the appropriate penalty, is subject to interpretation. At the same time, to be able to protect privacy by using any of these regulations, Web users still need an understanding not only of the existence of the laws and regulations, but also how they are exercised. In addition, while these regulations consolidate the efforts of several countries, their use is by no means global.

Finally, constant vigilance of users over the information collected about them, even if they are aware of the regulations in place, would be frustrated in cases of covert collection of information, where users are unaware of the extent of information that has been collected about them and they therefore have no control over the disclosure.

On the other hand, current violations of privacy and the indiscriminate disclosure of information suggest that Web users are not aware of the risks of disclosing their information and the protection that these regulations represent. Therefore, the best approach might be a controlled environment which applies the relevant regulations and penalties and provides the users with elements to control and preserve their privacy. The next section explores the technology approach to the preservation of privacy.

4.4.3 Technology

The Internet can be perceived as an intangible medium within which information is one of the most valuable assets. As mentioned in Section 3.5, detailed information about customer’s preferences and shopping habits provides e-tailers with tools to perform their marketing studies and to encourage sales by offering the customers a match to their requirements. The amount of information available on the Internet, with data mining available to filter that information, makes the misuse of the obtained information easy. The ease with which a user’s browsing can be traced and the ubiquitous nature of the Internet, are among the reasons for the particular privacy concerns related to the Internet.

Internet privacy issues are closely related to Internet security mechanisms such as cryptography and network security. In the study of the Internet's security and its risks, a deeper analysis has been carried out by associations such as the *International Telecommunication Union* (ITU), a United Nations agency. Their standardisation sector (the ITU-T), has explored data networks and open system communication and has issued a series of recommendations that cover diverse areas, of which security is one.

In the "Security architecture for Open Systems Interconnection for CCITT applications" recommendation [61] , X800, they suggest eight security dimensions:

a) Access Control
b) Authentication
c) Non-repudiation
d) Data Confidentiality
e) Communication Security
f) Data Integrity
g) Availability
h) Privacy

At the same time, the *Telecommunication Standardization Sector* (ITU-T) in the recommendation X805 [62] identifies the following security threads:

a) Destruction of information or other resources

b) Theft, removal or loss of information or other resources

c) Interruption of services

d) Corruption or modification of information

e) Disclosure of information

Whereas X805 makes specific mention of information disclosure, recommendation X800 makes a special mention of privacy. In their privacy definition they say: *"Because this term relates to the rights of individuals, it cannot be very precise and its use should be avoided except as a motivation for requiring security"*. Therefore, it is not surprising that some of the technological approaches to the preservation of privacy focus on increasing security or preventing privacy violations. For instance, it has been widely publicised that cookies represent a risk to the user's privacy, and the remedy of not using cookies has been suggested. However, cookies can be considered useful tools to aid browsing such as storing session-state and preserving information during Web browsing [74]. However, they can also be used to track the user's behaviour and that can lead to a threat to the user's privacy.

The alteration of browser settings to administer, erase or deactivate cookies has been recommended by popular magazines [139]. However, the disabling of elements such as cookies, JavaScripts or images used to track Web sessions, can limit or, in some cases, prevent a Web site from functioning. Krishnamurthy *et al.* [73], explored the repercussion of disabling some browsing facilities (cookies, third-party cookies, JavaScript, third-party JavaScripts, images, third-party images, etc.) when viewing Web pages from at least 1000 servers. They found that the indiscriminate disabling of these elements affected the Web pages' functionality varying from a small effect, e.g. not showing images, to severe cases when the complete web site did not work properly [73].

Another technological approach directed to mitigate the threat from data mining proposed avoiding the use of centralised data warehouses, due to the high risk that a consolidated repository of the data creates. With a centralised data warehouse, a security violation compromises all the stored information and therefore potentially risks the privacy of the users whose information is stored there. Another approach to protecting the data is by "data perturbation" (modifying the information in such a way that the modified information no longer represents valuable user information), and a series of association rules can be defined to partition data [26].

Another approach for customers who want to get involved in the process of protecting their privacy is the use of software. Anti-virus, anti-spyware, firewall, spam and parental control products, from companies such as McAfee, Symantec and Trend Micro, provide some level of protection against Internet threats.

On the other hand, AT&T has proposed a specialised free plug-in application called "*Privacy Bird*" that allows the user to determine how much their privacy is respected by each web site, according to the privacy policies of that web site. *Privacy Bird* is a downloadable application that alerts users each time they want to visit a Web page. An icon changes colour and there is a message if the web site respects the user's privacy, or not, according to the visited web site's privacy policies [6, 32].

Finally, another technological approach to the protection of privacy uses a third party as mediator. Such is the case of initiatives implemented by companies that specialise in techniques such as automatic deletion of files, anonymity, cryptography, identity theft protection, certificate authority, agents and pseudonymity. These approaches are summarised in Table 4.1.

Third Parties		
Approach	Company	Action
Deletion of files	NIST 800-88 guidelines	Complete clearing of data. Deletion includes: cookies, browser activity, internet history, passwords, credit card information, search history, photos, address bar, cache, history, programs, etc.
Anonymity	Anonymizer, The Cloak	The user goes to a third party that allows them the facility of surfing the web in an anonymous way. Or by encrypting the communication incoming from the user. Therefore, only *'The Cloak'* knows what Internet activities the user is performing.
Cryptography	Credentica (Recently bought by Microsoft)	An *"Issuer"* gives the user a token containing the user's "identity-related assertions". The validity of the token can be verified by a *"Verifier"*. Each time that the user sends information via the id token, the transmission is encrypted, protecting it from being intercepted. These transmissions involve the use of a cryptographic public key and user-generated private key.
Identity theft protection	TrustedID (Valid in the USA)	By setting up an account, a third party deals with the identity theft traditional risks, monitoring credit cards, removing data of pre-approved credit cards, issues three bureau credit reports, etc.
Certificate authority	VeriSign	"Issues public key certificates for a third party. The certificate enables encryption of sensitive information during online transactions, also has information that authenticates its owner and guarantees that it was issued by a confirmable *Certificate Authority* that verifies the identity of the certificate owner. Companies such as VeriSign allow the buyer of the certificate to add an icon named "VeriSign Secured Seal" as a visual backup of the company's presence. A VeriSign Web page also helps the users to verify the authenticity of VeriSign seals."
Negotiations and agents	Joung *et al.*	In the work presented by Joung *et al.* [65], users give certain value to their information, called "credit", the higher the credit, the bigger restrictions in disclosing that information. With the value of the information defined, an agent manages the personal data to fulfil the negotiation.
Pseudonymity	Martinez-Pelaez [88]	"A "digital pseudonym identity card" creates digital identities. The user proportionates his information to a third party. Then with the assistance of an identity card, the user is able to select the information he wants to disclose using a pseudonym."
Third party payments	Pay Pal [112]	"The service allows anyone to pay in any way they prefer, including through credit cards, bank accounts, buyer credit or account balances, without sharing financial information"

Table 4.1: Technological privacy preserving approach, using third parties

The use of biometrics as an identification method and its relation to privacy has generated a classification of systems according to this approach to privacy. This classification is presented next.

4.4.4 Privacy systems

The use of biometrics ("automatic recognition of individuals based on their physiological and/or behavioural characteristics" [64][p4]) has been introduced to assist an automatised identification of users. However, there is a privacy risk factor involved in biometrics. For instance, the use of DNA to identify a person can also be used to determine if that person is susceptible to certain disease, or the retinal patterns can provide medical information about diabetes or high blood pressure [64].

Therefore, due to the close relation between biometrics and privacy, systems have been divided into four different categories, shown in Figure 4.7, according to the way they impact on privacy.

- *Privacy invasive* - "a privacy-invasive system facilitates or enables the usage of personal data in a fashion inconsistent with generally accepted privacy principles" [96][p133].

- *Privacy neutral* - "a privacy-neutral system is one in which privacy is not an issue or in which the potential privacy impact is slight. Privacy-neutral systems are difficult to misuse from a privacy perspective, but do not have the capability to protect personal privacy" [96][p133].

- *Privacy protective* - "a privacy-protective system is one used to protect or limit access to personal information or which provide a means for an individual to establish a trusted identity" [96][p133].

- *Privacy sympathetic* - "a privacy-sympathetic system is one that limits access to and usage of personal data and in which decisions regarding design issues such as storage and transmission of biometric data are informed, if not driven, by privacy concerns" [96][p134].

Therefore, it can be said that a *negative relationship* represents a potential risk to the customer's privacy while the *positive relationship* represents a better controlled privacy environment for the customer's privacy.

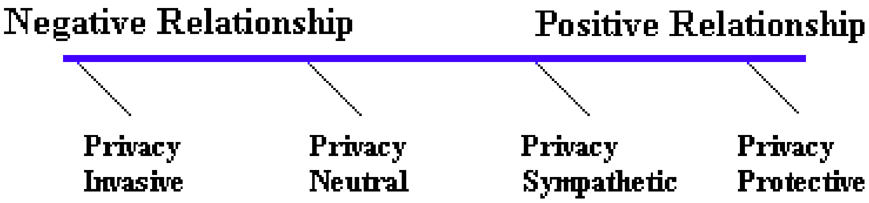


Figure 4.7: Technology impact on privacy [60]

Each technological approach to the preservation of privacy has associated with them a series of advantages and disadvantages. These are presented next.

4.4.5 Approaches to privacy - Advantages and disadvantages

There are a number of advantages and disadvantages in the use of different approaches to preserving privacy. For instance, the use of cookies to preserve sessions has the advantage of giving continuity to e-shopping for future sessions [74]. However their use has so commonly been adopted that preserving privacy by indiscriminately blocking cookies could translate into faulty performance of the Web pages. In general, to use any of these privacy-preserving approaches requires first that the customer is aware of the need to preserve their privacy; second, the knowledge of the existence of the approaches necessary to be able to choose from any of them separately or in combination.

Every presented approach to privacy requires a particular level of expertise and involves associated advantages and disadvantages. These are shown in Table 4.2.

Approach	Advantages	Disadvantages
Privacy awareness		
	Aware customers pay more attention to protecting their privacy, such as reading privacy policies and are more intent on controlling their disclosure of information [106, 32]	The lack of awareness in the customer that does not perceive any privacy threat and does not protect himself. The evolution of technology and the constant creation and modification of regulation requires an enormous effort from the customer who wants to keep up-to-date.
Regulation		
	The current regulations have organisational support. In case of violations to the regulations, a procedure can be followed.	An understanding of terminology and extent of privacy policies, disclaimers and terms and conditions is needed. Regulation is used after privacy violation has occurred. The use of regulation is subject to varied interpretation.
Technology		
	Customers can have a direct participation in the protection of their privacy	Customers need the knowledge of the different protections available to make a conscious selection of the one which most suits their needs. Customer's awareness that the browsing activity is being stored and of the collection methods. Active participation in avoiding spam, phishing attacks, spyware, which lend to fraudulent credit card transactions. Understanding of and use of authentication methods, including keeping a secret and secure password. A periodical checking of their information to verify that their information has not been misused and there is no fraud in credit reports. A constant updating regarding the variety of privacy threats and protection methods.

Table 4.2: Privacy preserved related work - Advantages and disadvantages

A close interrelation between the aforementioned approaches towards preserving privacy (raising *awareness*, *regulation* and use of *technology* (*ART*)) can be found, as shown in Figure 4.8. Rising of awareness (numbers 1,2,3 in Figure 4.8) motivates the user to increase her knowledge of the regulations (number 4 in the figure) and technologies (number 5 in the figure) available to assist her information disclosure and protect her privacy. In an ideal case, an aware customer would value the disclosure of her information, would decide when and under what circumstances to disclose it, and would know, if necessary, to place a trade-off value on the information. An aware person would possess a greater control on her privacy and would be able to use technology to her benefit. Furthermore, being aware of the existence and subject matter of the regulation in place, the customer would know the extent to which the information that can be linked back to her.

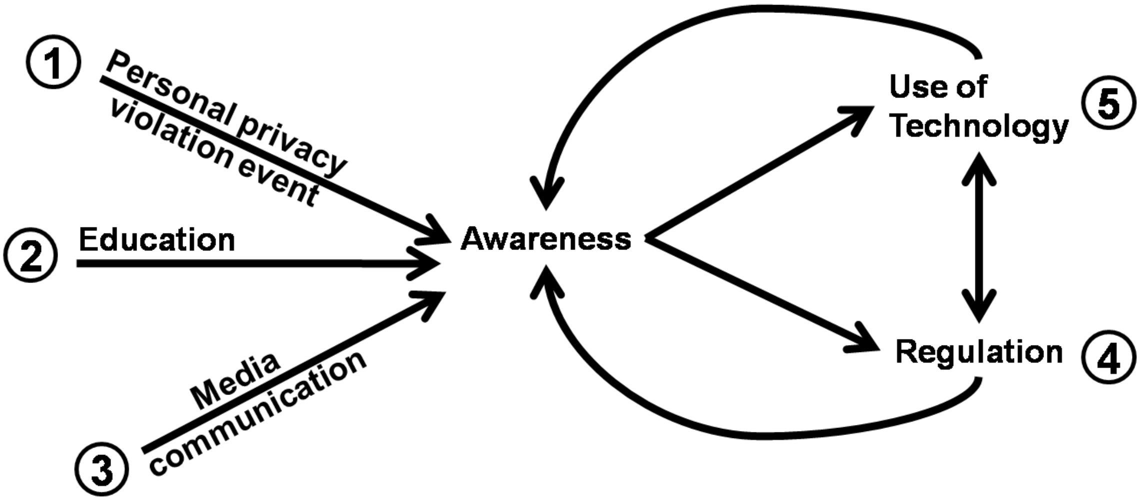


Figure 4.8: *Awareness*, *regulation* and *technology* (ART) cycle

However, since an aware customer would require a permanent update on new technologies and regulations, the second part of the diagram shown in Figure 4.8 applies, whereby new releases of technology and regulations trigger a parallel increase of awareness. This awareness-regulation-technology (ART) cycle requires an effort difficult to muster, even by the most dedicated customers. Furthermore, considering the lack of a universal Internet-regulation, the task of an updated awareness becomes unrealistic. Therefore, a single customer trying to cope with the ART approach individually faces, as yet, a near-impossible challenge. However, technology can provide an environment that incorporates the ART approach and assists customers in preserving their privacy.

A system that seeks to preserve privacy by combining the ART approach would have a better chance of success than systems that use only one technique. Figure 4.9 shows the relationship between the technology systems categories (privacy based) discussed in Section 4.4.4, the control held by the customers discussed in Section 4.4.1, the privacy indices proposed by Westin in relation to the customer's willingness to embrace regulation, Section 4.4.1, and finally regulation in *open privacy regimen* (where the firm has the right to collect and sell customer information including identity and purchasing habits) and *closed privacy regimen* (where customers have the right to remain anonymous) [149] discussed in Section 3.5.4.

Customers using *privacy invasive systems*, or belonging to the *unconcerned* customer group, or in the *no control* state, or using *open privacy regimen* e-commerce stores, face a bigger privacy risk than customers using *privacy protective systems*, in *total control* of their information, willing to use regulations to ensure their privacy (*fundamentalist* group) and using *closed privacy regimen* stores. A proposal to provide a privacy-protective/sympathetic system that aims to protect the *unconcerned* group's privacy and reinforces the *pragmatic* and *fundamentalist* groups' privacy, and that places the customer within an environment with elements to facilitate a more controlled and regulated information disclosure, is presented in the next chapter.

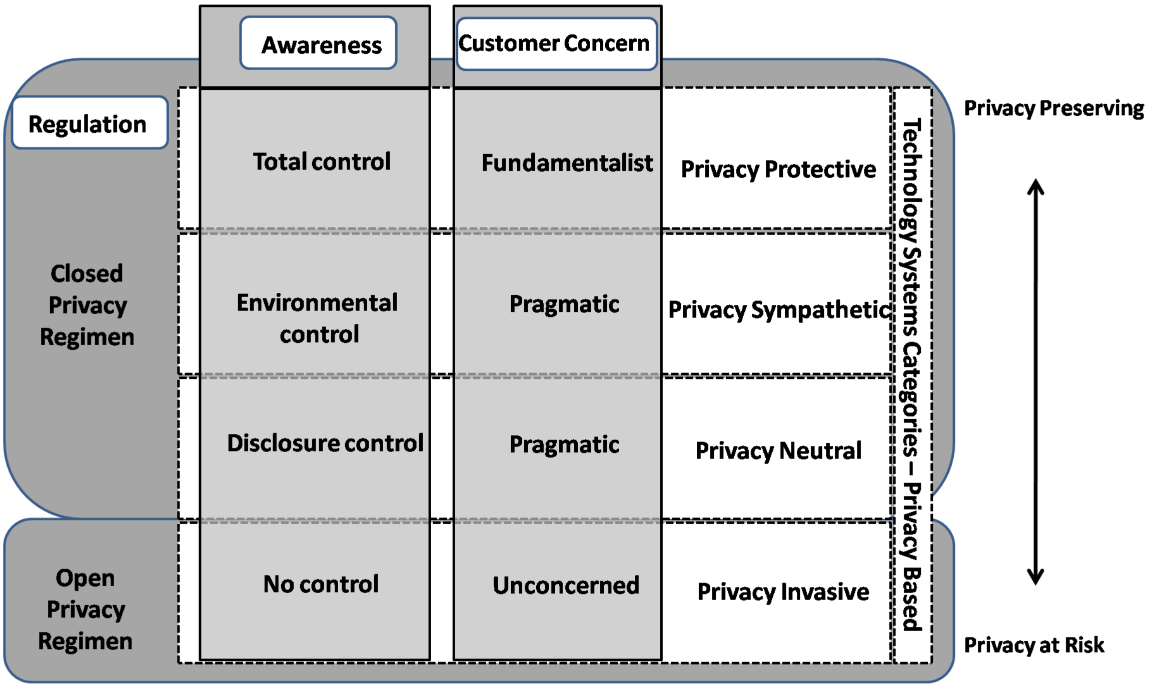


Figure 4.9: Privacy perspective combining *awareness*, *regulation* and *technology*

4.5 Conclusions

Privacy is an elusive concept; its definition has been related to multiple areas and concepts. However, for the purpose of this work three main aspects are singled out: the control over disclosure, the control over body or personal information and the right to be left alone. Since information is the main element of the Web, privacy is at risk, mainly due to the facility of losing control over the disclosed information and the difficulty of setting boundaries (in regards to the right to be left alone). Whether information can be voluntarily disclosed, such as in the case of social networks (Facebook, Bebo, etc), Web users face a different privacy risk with the information that is inadvertently disclosed through covert observation of the user.

As mentioned in Chapters 2 and 3, a motivation for the collection of information is the retailers' need of information to perform their marketing analysis. However, as noted, customers are not always aware that the collection of information occurs. They do not know the amount and detail of collected information and there is no possibility to amend it. When customers realise that their information was collected without their knowledge, trust is lost [10]. However, when customers are aware of the value that their information can have, they can seek to retain control over their information and perhaps trade it for benefits [53, 106].

With the existing privacy preserving approaches that use one or in the best cases two of the ART approach concept (*awareness*, *regulation* and use of *technology*), customers are left with inflexible means of protecting their privacy, that requires their constant update in the use of emergent technology (such as cryptographical keys, or non-flexible negotiation such as Privacy Bird), and the need of a constant update in the existence, content and extent of current legislation, making their efforts of protecting their privacy a difficult task. Therefore, a privacy preserving approach to e-commerce that empowers the customer regarding their information disclosure, provides a regulation element and encourages the raising of awareness is needed. This proposal is presented in the next chapter.

Chapter 5

Problem statement

5.1 Introduction

E-commerce, in comparison with the rest of commercial endeavour, has a short history, with 1998 marking the beginning of the e-commerce era according to the OECD [108]. During these years, the success of e-commerce has been noticed by e-tailers who, in order to take full advantage of this popularity, have adopted strategies, such as the case of personalisation. Personalisation, as described in Chapter 3, tailors the Web experience to a particular user or set of users, and has been used by e-tailers to assist their business analysis, define their business strategy and to encourage purchases. However, personalisation can present undesired side effects such as violation of privacy.

This chapter explores the problem of preserving privacy while e-shopping from the point of view of both customers and e-tailers. It also analyses the current situation and proposes a preserving privacy shopping environment (PPSE) as a possible solution.

5.2 Preserving privacy while e-shopping

5.2.1 Business perspective

Customers are at the centre of e-commerce; whether it is B2C, C2C or even B2B. From the business perspective, by having a better categorisation of the customers and the market, a closer match can be made between the customers' needs and the information about the products or services on offer by the company [89]. The categorisation of both customers and potential customers and the determination of the preferences of these groups is the objective of *market segmentation.*

Market segmentation, discussed in Section 2.3, is a valued technique that aims to facilitate better directed marketing by categorising customers or potential customers. Hence, it is not surprising that to direct their marketing and assist their business strategy, e-tailers exploit the advances of e-commerce technology and use techniques such as personalisation.

Personalisation techniques gather a certain amount of information from the customer's while interacting with systems. These results can be used to gauge the customer's preferences, develop a better customer profile and provide a better offer of products that match their needs. Customers, then, can be presented with suggestions related to their perceived preferences, offering additional items that the customer might purchase, while maintaining the original focus of their shopping.

5.2.2 Customer perspective

However, personalisation can be perceived as a two-edged sword; on the one hand, personalisation represents a benefit to the customer, assisting them with their shopping. On the other hand, the data gathered to facilitate personalisation can equally be used against the customer. For instance, customers could be given '*personalised*' prices based on previous purchases or perceived income status [123], not always to the customers advantage.

Another problem, from the customer's point of view, is the situation where they do not know that their personal information is being collected, not knowing also the risks that this represents [94, 140]. People seem to react in different ways to privacy violations. Some people, as soon as they realise that their data is being collected or their activities tracked, lose the trust in the e-tailer [94, 124] and abandon the store. Some other people are willing to give their private information data in exchange for certain benefits [145, 67].

A possible explanation for the careless disclosure of their data, apart from a lack of understanding of the potential risks involved, could be the lack of flexibility that customers face when they do their shopping with stores that could present a threat. If the customer wants to acquire the specific goods provided by a store that does not protect their privacy, disclosing data in exchange for goods could well appear to be the lesser evil [100]. This is not the only concern with respect to customers' disregard for their own privacy. Studies show that some e-commerce customers assume that they will be presented with the best options just because of the appearance of the Web site [140].

Therefore, the e-tailer's intention of making the widest possible use of the customers' personal information, and the customers wish to protect their privacy are in conflict.

5.2.3 Towards a fair compromise

To address the problem of protecting the customer's privacy while doing e-shopping, this research proposes to create an environment where e-tailers are able to collect the necessary information that may allow them to carry out their business planning (including the use of personalisation), while customers' privacy may remain protected. In this environment, a relationship between business needs and customer information can be achieved when customers consent to disclose, in a controlled environment, information needed by e-tailers. In such an environment, the need to submit false information as a protective measure, is substantially reduced. Customers would be aware of the information that is being stored about them, know who can access it and its potential usage.

This research proposes that, in addition to customers and e-commerce stores gaining equal benefit and preserving privacy, parties share responsibilities. Therefore, the following thesis statement is proposed.

5.3 Thesis Statement

> ***It is possible to develop a privacy preserving shopping environment (PPSE), which respects the customer's privacy needs while allowing the company to gather and use sufficient reliable customer-specified data to achieve a level of personalisation which can be used to encourage customer loyalty.***

To support the thesis statement, the following components, shown in Figure 5.1, were designed and implemented as part of a privacy preserving shopping environment (PPSE):

- a third party, named Alter-Ego, whose objectives are to facilitate and mediate the customer's disclosure of information to the e-tailer; and
- an agreement between the e-store and the Alter-Ego, called *personal level agreement* (PLA), which has the objective of formalising the exchange of sensitive information and preferences between customer and e-tailer.

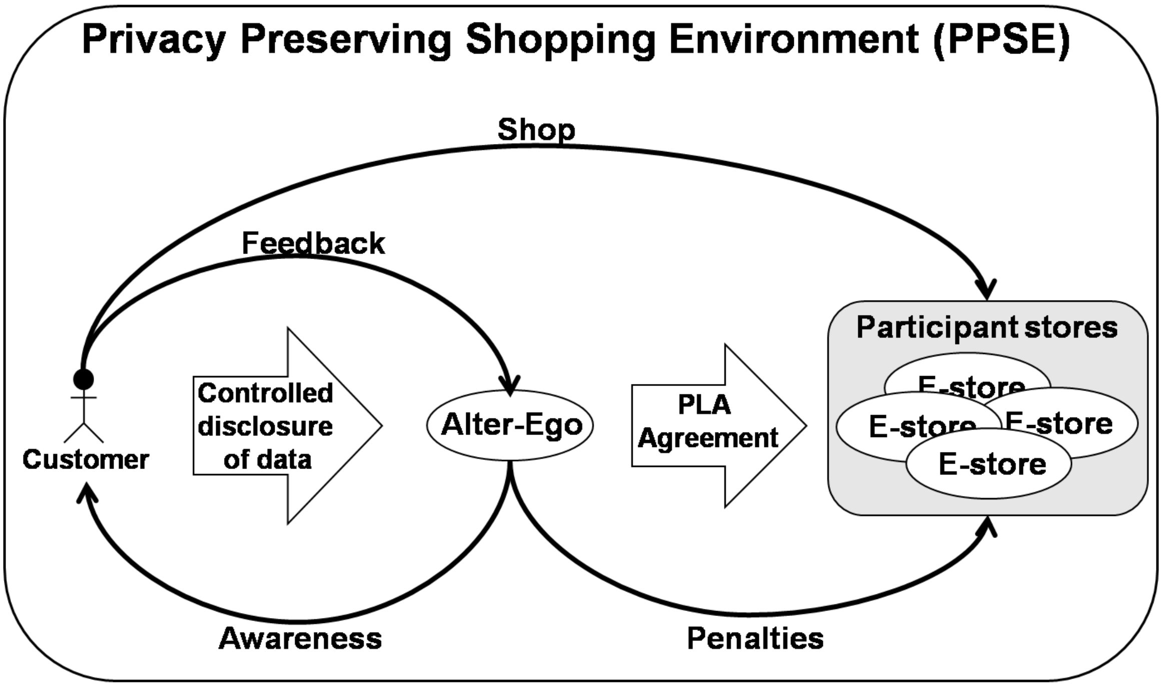


Figure 5.1: Privacy protective/sympathetic system proposed.

As Figure 5.1 shows, customers store, in a controlled way, their information in the Alter-Ego. This information excludes data that could be used to identify the client, i.e. name, address. Customers can disclose their information to the participant stores in a regulated way via the Alter-Ego using the PLA agreement. Having disclosed the desired information, customers can do their shopping directly with the participant store.

Figure 5.1 also shows the Alter-Ego raising awareness, customers giving feedback to assist the regulation process, and finally, penalties to be applied to participant stores that do not comply with the agreement.

5.4 The Privacy Preserving Shopping Environment (PPSE)

The existing efforts used to preserve privacy, analysed in Chapter 4, are categorised as: raising *awareness*, *regulation*, and use of *technology* (*ART*). Currently in industry or research, one (or at most two) of the aforementioned approaches are used. The PPSE aims to combine all three.

5.4.1 PPSE - Awareness

One disadvantage of initiatives that aim to preserve privacy by raising awareness, such as "*Privacy International*" introduced in Chapter 4, is that only customers who are already concerned about their privacy deliberately look for and find the information provided by organisations. Another disadvantage of these methods is the frequency with which the information is updated and the lack of means to verify that the information was properly received and understood.

When the availability of the information to create awareness is not easy to find, customers do not perceive privacy risks and do not undertake any protective measures. In a paper published by Conti and Sobiesk [30], it was shown that out of 352 undergraduates and a comparison group of 25 middle aged adults, 80% of the participants were comfortable with the privacy that they had when they used search engines, even if 99% of the participants believed that at least some of the search keywords they used were retained.

Raising privacy awareness has a positive effect. Olivero *et al.* [106], carried out a series of interviews and observed that increased awareness of information gathering activities resulted in participants increasing their demands for more control over their personal information.

Customers are moving towards a more pro-active role in the protection of their privacy. The results of 13 years of surveys carried out by Westin [75], used to determine the trend of privacy perceptions, show that the number of *unconcerned* participants (those that trust the collection of information by organisations, are not in favour of new privacy regulations and do not use controls to protect their privacy) had decreased. The number of *Fundamentalists* (participants who distrust organisations asking for their personal information, are worried about computerised-gathered of information and its uses, and favour the update and the creation of new regulation to protect their privacy) was constant. It was the number of *pragmatics* that increased (i.e. those participants that weigh up the benefits of protection and regulation against the amount of information they are prepared to disclose, believing that trust should not be freely given but earned and seek to have opt-out options against the indiscriminate collection of information) [75].

The PPSE approach aims to raise customers' awareness by continual and updated presentation of information about the risks and methods of privacy protection. By making privacy awareness literature available to the customer, the PPSE aims to increase customer knowledge and give the customer greater control over their personal information. With this knowledge, customers have the means to perceive privacy risks and obtain the best prevention from the provided shopping environment. Raising customers' awareness enables them to make a conscious decision to protect their privacy and balance their choice of Web features, i.e. personalisation, against their need for privacy.

5.4.2 PPSE - Regulation

Laws, conventions, privacy policies and even the *World Wide Web consortium*'s "*Platform for Privacy Preferences*" (P3P), have been designed/implemented to ensure that privacy is respected. However, they present a major drawback: the requirement of reading and understanding the legal, and sometimes technological, terminology associated with privacy policies, terms conditions and disclaimers. For many, this represents an impossible challenge. Furthermore, legislation is not necessarily enforceable in all countries.

Although the PPSE is not presented as a system to be enforced by governments, it is a well regulated space where customer can resort to the protection of their privacy. As part of the PPSE environment, the third Web party, Alter-Ego, mediates between the customer and the company, facilitating the customers' disclosure of data controlled by the *personal level agreement* (PLA) [46]. At the same time, a basic privacy policy that all participant stores must abide by is set up by the PPSE. This basic privacy policy, together with a close monitoring of the participant stores' compliance, gives the PPSE the ability to maintain transparency in the handling of data and in privacy preservation. Hence, the customers can be assured that the participating Web stores comply with fair privacy policies, laws and conventions. With such close monitoring and the basic privacy policy established, the customer would not need to read privacy policies each time he or she enters a participant store. The customer could be certain that the participant stores comply with the established privacy policy.

On the other hand, customers are encouraged to participate in the regulative process by means of giving feedback and ranking their privacy experience while shopping with the participant stores. This is explored in Section 5.5.3.

5.4.3 PPSE - Technology

Regarding the efforts towards preserving privacy, the majority of the proposals explored in Section 4.4.3 involved a significant technological component, such as the issuing of cryptographic keys to verify identity. Unfortunately, customers still need to have an understanding of the privacy risks and certain technology expertise in order to want to use any privacy-protective technology.

In addition to the need of a certain level of expertise, customers have a major disadvantage: in cases where software is used to identify a Web site that does not conform to uniform privacy policy practices, the only option that a customer has is to avoid that Web site. However, if customers are in need of products that a particular site offers, they may well decide to ignore the warning and purchase there anyway, not knowing if the privacy policies are fair or not. With the PPSE, the customer has the facility of disclosing as much or as little information as desired and still continue shopping with the participant stores that conform to the PPSE fair privacy policy.

One way to assist customers is by means of a third party mediator. Third party mediators have successfully been used to assist customers and companies. Examples are:

Credentica, issues the user with cryptographically protected ID tokens to protect their transmissions from being intercepted.

PayPal, "allows any business or consumer with an email address to securely, conveniently and cost-effectively send and receive payments online" [113].

Therefore, it can be expected that customers being targeted for the approach proposed by the PPSE would probably be familiar with some other third party Web site that assists them in the mediation, management and facilitation of their shopping.

At the same time, as discussed in Section 4.4.4, systems can be divided into:

- *privacy invasive* systems that facilitate or enable the use of personal data in an inconsistent way with generally accepted privacy principles;
- *privacy neutral* systems where privacy is not an issue, are difficult to misuse from a privacy perspective, but have no capability of protecting personal privacy;
- *privacy protective* systems used to protect or limit access to personal information, or which provide means for an individual to establish trusted identities; and

- *privacy sympathetic* systems that limit access to and usage of personal data and in which decisions regarding design issues such as storage and transmission of information are informed, if not driven, by privacy concerns.

Based on these categories, privacy invasive and privacy neutral systems pose a potential risk to the customer's privacy, as Figure 4.7 in Section 4.4.4 shows, while privacy sympathetic systems represent a more controlled privacy environment for preserving the customer's privacy.

The PPSE proposes a *positive relationship* between customers and e-tailers, by means of a privacy protective / sympathetic system, and an easy-to-use third party, where the customer is given the flexibility to decide what information to disclose to which participant store. This flexibility, together with the confidence that the participant stores are compliant with the basic privacy policy defined in the PPSE, gives the customer the advantage of shopping while being reassured that the confidentiality of their data is being respected.

5.4.4 PPSE - Advantages and disadvantages

The advantages of the PPSE are:

- **To the customer:**
  - Easy access to updated, privacy-relevant information to encourage their privacy-awareness, including awareness of the risks of indiscriminately disclosing personal information, details of how to protect themselves and suggestions of recovery procedures in cases of privacy-loss.
  - Their privacy would be protected by means of the PPSE, and the disclosure of data will be supported by the PLA agreements with the participant e-stores.
  - By means of reliance on the close surveillance of the participant stores' adherence to the PPSE, customers would have no need to read each clause of each participant stores' privacy policy. Only compliant stores would be listed as participant stores.
  - An easy-to-use environment, Alter-Ego, is provided which assists customers in doing shopping in a privacy protective way.
  - The customers shopping using the PPSE approach and the Alter-Ego Web site would have:

    - their non-identifiable preferences and sensitive information stored in a repository;
    - their information made available without the need to provide it every time they do their shopping; and
    - the flexibility of deciding the level of disclosure they want to use to enter the store, knowing that the participant stores comply to the PPSE fair privacy policy.

- **To the e-commerce store:**
  - Since customers would disclose data voluntarily and within a trusted relationship, there would be no need of masking themselves or providing false information. Therefore, the information received via the Alter-Ego is expected to be more reliable than the information inferred from simple analysis of raw browsing data.
  - Detailed preferences and sensitive information sent by the Alter-Ego would be ready to be used in the e-stores' market segmentation, therefore the stores would have the opportunity to suggest what information they would like and need.
  - With the data provided by Alter-Ego, even stores with no personalisation or customisation functionality, would have the opportunity of using the parameters, i.e. list of ingredients linked to allergies, to implement (or integrate) their personalisation.
  - The stores in the PPSE environment that conform with the PPSE precepts would benefit from positive customer feedback, increasing their reputation.

- **To preserve privacy:**
  - The PPSE provides a way of evaluating customer privacy awareness and customer reaction to the presentation of information, enhancing customer awareness of privacy.
  - The customer feedback will be used by Alter-Ego to assist the close monitoring of the behaviour of participant stores, and achieve community regulation.
  - The feedback given by the customers would affect e-stores' reputation and warn other customers about risks.

It would be unrealistic to assume that the PPSE environment approach does not have disadvantages. The following disadvantages have been detected.

- **PPSE disadvantages:**
    - A multidisciplinary team would be required to carry out the following actions:
        - content management administration of the awareness zone;
        - constant revision of new categories added to gold disclosure level (disclosure levels are explained in Section 5.5.1);
        - in customers' participation:
            - moderators for the forums,
            - monitors of customer feedback to avoid manipulation by e-commerce stores pretending to be customers, influencing or modifying other participant's opinion,
        - monitoring of any variation or change in the participant store's privacy policies and personalisation or customisation; and
        - user support e.g. in case of recovery of lost information or incompatibilities with their systems.
    - Companies that already have personalisation techniques implemented would have to invest time and resources to make adjustments to be compliant if they want to enjoy the advantages that the PPSE environment proposes.

To solve the disadvantages, the PPSE would have to be implemented as more than a single person initiative. The PPSE could be implemented either as a foundation to preserve privacy with charity funding or donors or as a private company.

In the case of implementing the PPSE approach as a private company, a business case would need to be in place to evaluate the best ways of making it profitable. However, one or more of the following proposals could be used to obtain profit from the PPSE:

- Customers could be charged according to the time that they use the PPSE facilities, i.e. free use for a certain time usage and membership subscription, or

- Customers could be charged certain amount of money according to the disclosure level, or

- Participant stores could be charged to be listed on the Alter-Ego Web site, or
- Participant stores could be charged according to how much the PPSE is used, i.e. a charge per sale, or
- Participant stores could be offer consultancy or training on customising and adapting the existing personalisation of the newcomer stores.

5.5 The Alter-Ego

Regardless of the variety of current privacy-preserving tools available, as discussed in Section 4.4, the use of these initiatives remains in the expert customer domain, leaving less knowledgable customers unprotected. Customers, both novices and experts, need to be provided with an easy-to-use, protective mediator which use does not require a major investment of time or expertise.

Alter-Ego is an easy-to-use third party that mediates between customer and participant stores, storing the customers' preferences and sensitive information and facilitating the disclosure of information to participant stores. In Alter-Ego, the customer is given the flexibility to:

- decide what information will be sent to a participant store;
- have access to a space where the raising of customers' awareness of privacy issues will be addressed; and
- have a space to give feedback about their privacy preserving experience with the participant stores, and rate participant stores in order to promote and facilitate regulation.

The Alter-Ego elements are:

Awareness A set of Web pages which inform customers about news related to privacy issues and provides help in case of privacy loss;

Mediation A set of Web pages linked to a repository where customers can store their preferences and sensitive information, and have total control of its disclosure;

Redirection Web pages which give the customer the option of selecting the kind of data to be disclosed to each participant stores, using 3 different disclosure levels (*bronze*, *silver* and *gold*, explained in Section 5.5.1); and finally

Regulation by participation A feedback zone that allows customers to rate how well preserved their privacy was by the participant stores.

5.5.1 Alter-Ego - Disclosure levels

The Alter-Ego allows the customer to disclose as much or as little information as they perceive necessary. The information provided can be used by the stores to offer customers customisation and personalisation. As discussed in Section 4.4.1, a categorisation of customer's privacy perception based on a series of surveys with respect to different periods of time and different areas, such as "consumer privacy concern index" and "medical sensitivity index" has been defined by Westin's in his "Private Index" [75]. Westin's private index [75] categorises customers' perception of privacy into: "*high, medium and low*", and the participants of the groups fitting in those categories are called "*fundamentalist, pragmatic and unconcerned*".

The Alter-Ego, on the other hand, proposes using three levels of information disclosure according to the customer's privacy needs. The levels are *low* disclosure *(**bronze**)*, *medium* disclosure *(**silver**)* and *high* disclosure *(**gold**)*, and are linked to the amount of data that customers are willing to disclose to the e-commerce store. By providing customers with three different options to preserve their privacy, all three categories of customers in Westin's index could match their privacy perceptions and expectations and freely decide which information will be disclosed. Therefore, the more data the customer discloses, the more customer data gathered by the store, and the more detailed the personalisation, or customised pointed-personalisation in the particular case of this research, that can be provided by the store.

With detailed user-specified data, the store will have data to formulate a better market segmentation and at the same time, the customer's privacy and confidentiality will be respected.

Alter-Ego disclosure level	**Information disclosed**	**Westin's privacy index**
Bronze (Low)	***Anonymous access*** No sensitive data is collected. Customer has anonymous access.	Fundamentalist
Silver (Medium)	***Preferences data only*** Basic preferences are disclosed. No identification of the customer. No link to previous purchases. Basic personalisation is provided based on customer's preferences.	Unconcerned
Gold (High)	***Sensitive information*** Preferences and sensitive information are disclosed. The veracity of the data that will be stored is previously confirmed with customer. Full personalisation. Recommendation lists based on previous purchases or similar customer's purchases are provided to encourage customers and reward their sharing of private information.	Pragmatist

Table 5.1: Alter-Ego disclosure levels matching Westin's privacy concerns categories.

5.5.2 Alter-Ego - Division of information

The first division of information held in the Alter-Ego is based on the UK *Data Protection Act* [105], that divides information into two main groups; *personal data*, the information that can identify a living individual, and *sensitive personal data*, the information about the individual in areas such as religious beliefs, physical or mental health or condition, sexual orientation. To preserve the customer's privacy, the Alter-Ego avoids the collection, use or disclose of personal data (information that could lead to the participant's identification, such as name or address), limiting the collection of information into three categories, also shown in Table 5.1:

- Bronze - low disclosure level, corresponds to *anonymous access.*
- Silver - medium disclosure level, corresponds to *preference data only.*
- Gold - high disclosure level, corresponds to *sensitive information.*

At the *bronze* disclosure level, anonymity is offered to the customers who decide not to disclose any data. Customers can browse the store without revealing who they are. No information is collected that might link the user identity to their browsing activity. However, since the store is collecting no data from the customers, no customisation, personalisation or recommendations can be offered. For its characteristics, this level is directed to Westin's "fundamentalist" group.

At the *silver* disclosure level, disclosure is achieved by presenting customers with a list of preferences. From these preferences, the customer can decide if they want their choice of preferences to be used for personalisation. The disclosed information can be used by the store to support their marketing strategies.

The preferences presented in the Alter-Ego for the *silver* disclosure level, *preference data only*, include specific choices about food preferences such as vegetables, fish, pork, which although apparently have no reference to the customer's privacy, have been found to have a link to certain attitudes and beliefs that customers might find embarrassing to share [93]. Besides, the inferences about the customer's reasons for the selection and consumption of these elements may represent a potential risk to privacy. For example, Molina *et al.* [93] presented a study where meat was perceived to be an upper class food selection by Brazilians and its consumption reflected a higher social status, whereas the consumption of fruits and vegetables were related to lower social status. Their preferences were linked to the belief that the consumption of meat would give them a higher social presence [93]. Another case can be found in the customers' avoidance of meats, especially pork, since it could be used to infer a link to religious beliefs.

In addition, at *silver* disclosure level, customers are presented with five different categories (*intensity of preference*) for each of the preferences. These non-ordinal categories indicate the intensity of the preference, and provide a finer granularity in the disclosure of the customers' options. The *intensity of preference* categories are: *always, sometimes, maybe, never, don't care.* From this information, the store can use the intensity preference for a particular preference and feed it to their personalisation engine. At the same time, this information can be used to support the store's data analysis. The silver disclosure level is directed towards the customers who have an "unconcerned" perception of privacy.

Finally, at *gold* disclosure level, the options provided to the customer are those that can be considered sensitive information such as health issues or religious preferences and give a more detailed profile of the customer. For instance, there is a privacy risk in cases of customers selecting halal meat, since its delivery address can be used to trace Muslim communities.

Customers choosing the gold level of disclosure can indicate the intensity of their preferences (using the granularity provided by the five *intensity of preference* categories for each of the options presented by the gold level) or introduce new elements to assist their shopping. The introduction of elements allows customers to have a participatory role and a better personalisation of their shopping, which would add dynamism and flexibility to their shopping experience. For example, customers can add the ingredient "coffee" and the intensity of preference "never" to avoid items containing coffee.

At the same time, the gold level also includes the disclosure of valuable data for the store's marketing purposes such as gender or age. Therefore, due to the amount and detail of the disclosed information, customers using gold level would be presented with full personalisation and recommendations. At the same time, since the gold level makes use of previous purchases to offer recommendations, customers using gold level can be presented with recommendation lists based on other customers' choices or their own previous purchases, facilitating dual usage of the information (privacy and search requirements).

Finally, customers using gold level would also have the opportunity to access the information that the store holds about them and amend the information associated to their preferences or sensitive information.

The gold level is directed towards the "pragmatist" group that wants to be convinced of the benefits of the applications before committing themselves to its use.

With the proposed division of information, the Alter-Ego allows customers to select their desired disclosure level and to disclose their information to a participant store. From this disclosure system perspective, the change of level by the customer represents a change of commitment between the store and the customer.

5.5.3 Alter-Ego - Regulation

In order to fully implement the PPSE, Alter-Ego allows customers to participate in the regulation process by encouraging customers to rank participating stores and by following up cases of misbehaviour. Ranking has been successfully used by companies such as eBay, that implements a feedback system to assist buyers and sellers to build their own reputation. A reputation system, such as the one used by eBay, faces three challenges; "provide information that allows buyers to distinguish between trustworthy and non-trustworthy sellers", "encourage sellers to be trustworthy", and "discourage participation from those who aren't"

[119][p3]. Resnick *et al.* [120] in his analysis of data from eBay concluded that, under certain circumstances, the feedback net "makes up for the lack of traditional feedback mechanisms" [120][p23]. A positive ranking in a reputation system, such as the one provided by eBay, has a beneficial effect on the sellers. Resnick *et al.* [121] show that buyers were willing to pay, on average, 8% more to sellers with high positive feedback than to a new sellers.

5.5.4 Participant stores

To qualify as a participant store in the PPSE environment, the store needs to agree to comply with the PLA agreement and the privacy policies required by the PPSE. The participant stores would have to provide services to match the three level Alter-Ego information disclosure levels and respect the associated confidentiality levels.

5.6 Personal Level Agreement (PLA)

The *personal level agreement* (PLA) is an agreement between the customer and the participant e-commerce store and regulates the customer information provided by Alter-Ego to the participant store. Its main objective is to formalise the transfer of sensitive information and preferences from customers via Alter-Ego to the participant store.

- The e-commerce store agrees to the following:
  - The confidentiality of the customer's private data will be respected and the data provided will be used exclusively for their own marketing and business purposes.
  - The information collected using this agreement will not be disclosed to other signatories or third parties.
  - The information disclosed by the customer using the Alter-Ego, will be used to provide extra services, such as personalisation;
  - Customers using *gold* disclosure level will be allowed to view and amend the information held about them in relation to the preferences and sensitive information associated with them.
  - Any contravention of the rules by the participant stores, found by the PPSE or reported by customers, will be investigated and penalised accordingly.
- The customer commits to the following:

- To use the Alter-Ego third party mediator Web site for their shopping;
- When ranking their privacy experience with the participant store, to provide objective and truthful feedback;

5.7 Conclusions

The PPSE is a novel integral proposal that preserves privacy of customers while they do their e-shopping. The PPSE incorporates the three different approaches found in previous privacy preserving methods: raising *awareness*, *regulation*, and the use of *technology*. At the same time, it offers a space where customers' information can be protected and their privacy respected.

The main advantages of this proposal are: the flexibility with which customers can store and disclose their information, the existence of a series of participant stores that respect customer's privacy and abide by the same privacy policies, and a repository of information where customers can manage their preferences.

Chapter 6

Design and implementation

6.1 Introduction

Support the thesis statement is divided into two parts. The first part states: "*It is possible to develop a privacy preserving shopping environment (PPSE), which respects the customer's privacy needs while allowing the company to gather and use sufficient reliable customer-specified data to achieve a level of personalisation*". The second part states: *"which can be used to encourage customer loyalty"*. To support the first part of the thesis statement, a prototype of the PPSE environment was created, and to support the second part a user test to investigate the customers' satisfaction and potential loyalty was developed and performed.

This chapter presents a substantiation of the first part of the thesis statement: the creation of the prototype PPSE. This includes the Alter-Ego third party site, and a structure for the PLA requirements. For evaluation purposes, a simulated participant store called *bshop*, was also developed.

The elements of the PPSE are presented next. Section 6.3 presents three potential implementation approaches and the reasons for selecting the chosen one. Section 6.4 describes the design and implementation of the Alter-Ego. Finally, Section 6.6 details the design and implementation of the bshop.

6.2 Elements of the PPSE

The PPSE, as introduced in Chapter 5 and shown in Figure 5.1, contains the following elements: the *Alter-Ego*, a trusted third party that facilitates and mediates the customer's disclosure of information and the e-tailer's user-specified data requirements, and the *personal level agreement* (PLA), that formalises the exchange of non-identifiable sensitive information and preferences between customer and e-tailer.

While the PLA was described in Section 5.6, this section focuses on the Alter-Ego, a trusted third Web party, which has the following functionalities:

- Raise privacy awareness in the customer, can be obtained by: previous personal privacy events, press communication or education advising customers of the risks of indiscriminate information disclosure. The rise of awareness, located in the Alter-Ego, has an informative approach. Therefore, customers using the Alter-Ego third party would be presented with information about the importance of protecting their privacy, the reasons for using the PPSE and the suggestion of a series of general steps to protect themselves even if customers opt not to use the PPSE.
- Facilitate customers' controlled disclosure of data, by providing customers with three different disclosure levels: *bronze*, *silver* and *gold*, according to the amount of information that will be disclosed. At the same time, the store offers personalisation features according to the disclosed information. To recap:
    - At the *bronze level*, the customer discloses no information, the store has no elements to offer personalisation.
    - At the *silver level*, the customer discloses preference data only, allowing the store to provide some preferences-based personalisation.
    - At the *gold level*, the customer discloses preference and sensitive information, allowing the store to provide full access to personalisation features.
- Provide a feedback mechanism to encourage customer participation in the regulation process, by means of a customer ranking, as shown in Figure 6.1. Customers could be able to rank participant stores, and cases of misbehaviour could be followed up. This ranking mechanism has been successfully used by companies such as eBay to assist buyers and sellers in building a reputation [120, 121]. Since reputation systems using ranking mechanisms have already been successfully used, it is reasonable to conclude that a reputation system would have the same positive effect within the PPSE. Therefore, due to the characteristics of user test including time limitations (45 minutes), not-repeatable nature and the use of a student e-shop (without providing the store

service), the feedback via ranking mechanism of the PPSE consequent penalisation of non-compliant e-stores, was not implemented in the PPSE prototype.

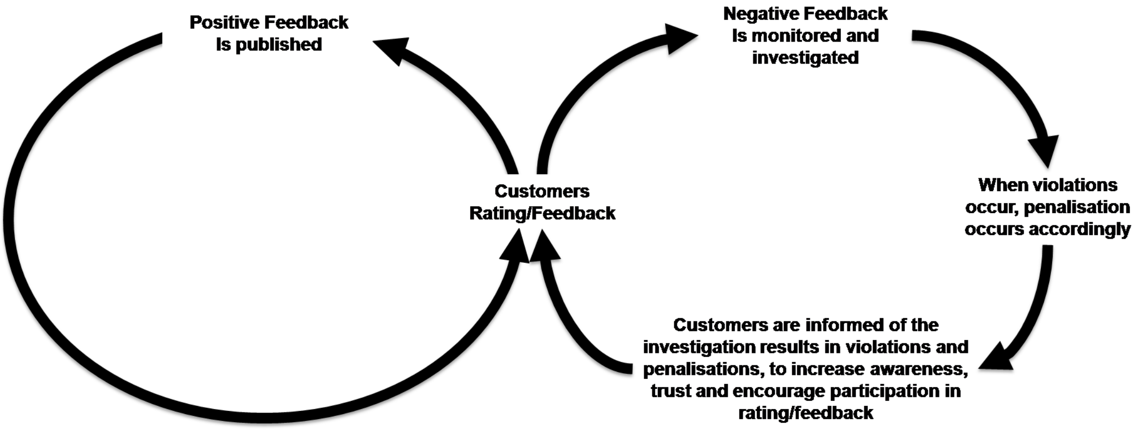

Positive Feedback Is published
Negative Feedback Is monitored and investigated
Customers Rating/Feedback
When violations occur, penalisation occurs accordingly
Customers are informed of the investigation results in violations and penalisations, to increase awareness, trust and encourage participation in rating/feedback

Figure 6.1: Regulation by feedback and customer ranking.

- Impose penalties on participant stores in relevant cases,
- Implement the PLA agreement, and
- Direct customers to do their shopping with their selected participant stores according to their desired disclosure level.

Three different designs of the Alter-Ego were explored. The details of each of these options, their advantages and disadvantages, and the criteria for the selection of the chosen prototype design option are presented next.

6.3 Alter-Ego - Design options

6.3.1 Design option 1

> ***Retain the customer's personal information within the customer's own machine using cookies.***

This proposed design would give the customer the facility of manually creating their own cookies containing their profile. This information would, in future, be shared with the participant e-commerce sites using the PPSE approach.

Operation: In this option, shown in Figure 6.2, the customer goes to the Alter-Ego that contains forms to be completed with the customer's information. Alter-Ego creates a cookie containing the disclosure level and the customer's information and returns it for storage on the customer's own machine. In this proposal, Alter-Ego is used as an interface that allows the customer to control the creation, update and deletion of the cookies and the information contained in the cookies.

In this option, an interaction with the participant e-store could be adapted from the proposal presented by Shankar [132]. In this proposal, a cookie policy is defined by the user, and each time the user enters a Web page, the system compares the user's cookie policy with the Web site's cookie policy and either accepts or denies the user's access/ disclose of information into the Web site, and has the possibility of automating both the acceptance or denial, so that the customer does not have to repeat the decision for each different Web site.

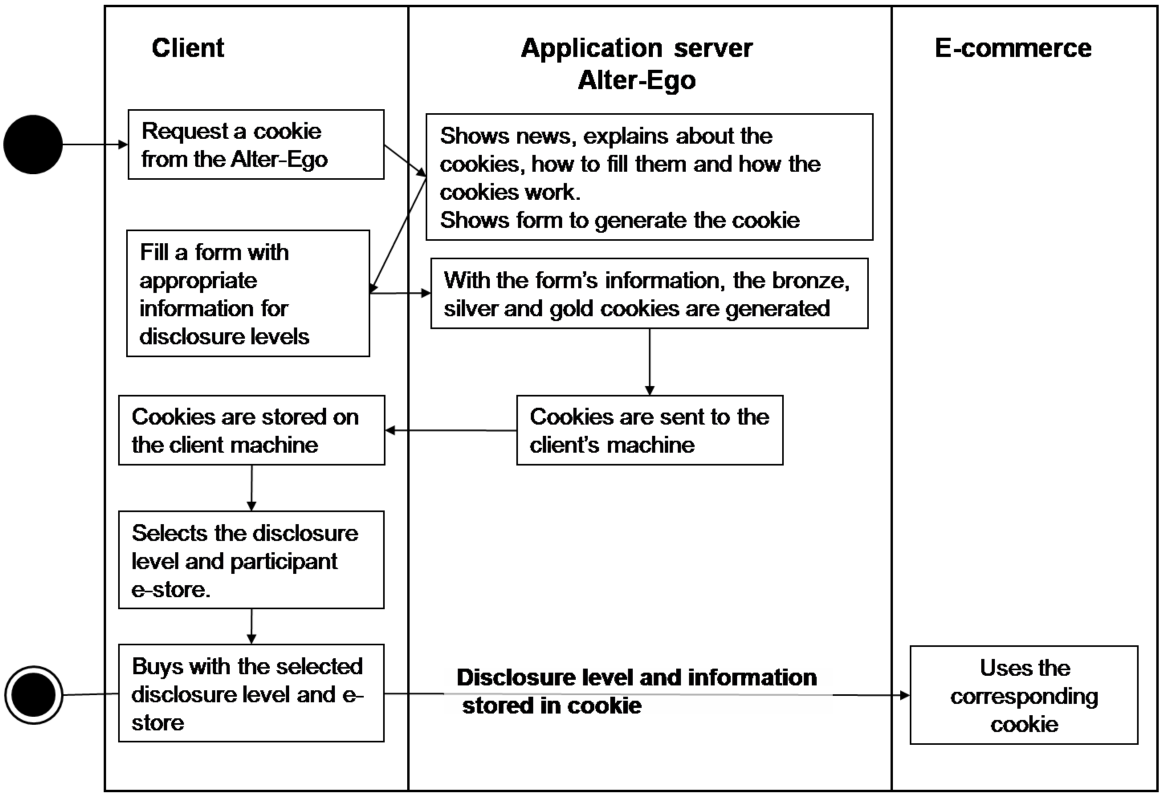

Client
Application server
Alter-Ego
E-commerce
Request a cookie from the Alter-Ego
Shows news, explains about the cookies, how to fill them and how the cookies work.
Shows form to generate the cookie
Fill a form with appropriate information for disclosure levels
With the form's information, the bronze, silver and gold cookies are generated
Cookies are stored on the client machine
Cookies are sent to the client's machine
Selects the disclosure level and participant e-store.
Buys with the selected disclosure level and e-store
Disclosure level and information stored in cookie
Uses the corresponding cookie

Figure 6.2: WebML activity diagram. Design option 1 - *Cookies*.

Implementation: To probe this approach, a proof of concept script was created using

JavaScript. This implementation contained a form with which the customer would enter the information so that the system could generate a cookie. Two possibilities were explored:

- The first one allowed the customer to record all their data and tag each item as *gold*, *silver* or *bronze*. This case created one cookie holding all the information pertaining to the three different disclosure levels.
- The second case required the customer to step through three different pages to create three different cookies.

In the first case, using a cookie with three different disclosure levels in the cache, the retailer's use of only the part of the cookie selected by the customer cannot be guaranteed. The temptation of using the *gold* information when the user selects *bronze* would probably be difficult to resist or even to detect and control. In case number two, when the customer is presented with a Web page per disclosure level, three different steps were required to obtain bronze, silver and gold, requiring an extra effort from the customer.

Advantages: This approach ensures that the main functionality is supported. Customers can store, update or delete the information and disclose only the information that they choose to. This proposal is backed up by the P3P initiative, that dictates the controlled interchange of information from the user, by the exercise of certain guidelines and the observance of rules for the use of cookies [33]. At the same time, this proposal allowed the customer to store, update and delete their own cookie containing whatever information they desired to share. However, this choice would be present only during the creation of the cookie and the customer would not have the same control after the cookie has been sent to stores.

Disadvantages: A disadvantage present in this design option was the cookie's name. To be able to use the cookie in a participant e-store, the cookie's name needs to match the e-store's site address. To solve this, the Alter-Ego would need to create the cookie and direct the customer to the e-store to match the e-store's address. Therefore, a dynamic access to the participant e-commerce sites would not be fulfilled using this particular option.

Finally, the usage of cookies has a clear disadvantage when the user stops refreshing the cookies, or when cookies expire, or when cookies have been perceived, and publicised, as a privacy threat, or when, in some cases, a customer does not permit cookies on their machines, or when using computers from public places (e.g. libraries) [139].

Conclusion: This option was not selected due to the lack of control and the weaknesses presented during the prototype implementation.

6.3.2 Design option 2

> ***Retain the information within the customer's own machine stored as a text file.***

This proposed design could give customers the ability to create their own text files containing personal profiles. This information could be shared with the rest of the participant e-commerce sites using the PPSE approach.

Operation: In this option, illustrated in Figure 6.3, the customer completes a form presented by Alter-Ego. The user is able to store, update and delete their information, storing the resulting file on their own machine (the client machine). The Alter-Ego Web site would retrieve the contents of these files according to the user's privacy preferences and use this to mediate the e-shopping experience.

In this option, an adaptation from the semantic mediation process framework presented by Park *et al.*[110] could be implemented to facilitate the inter-operability among the heterogeneous and distributed information sources.

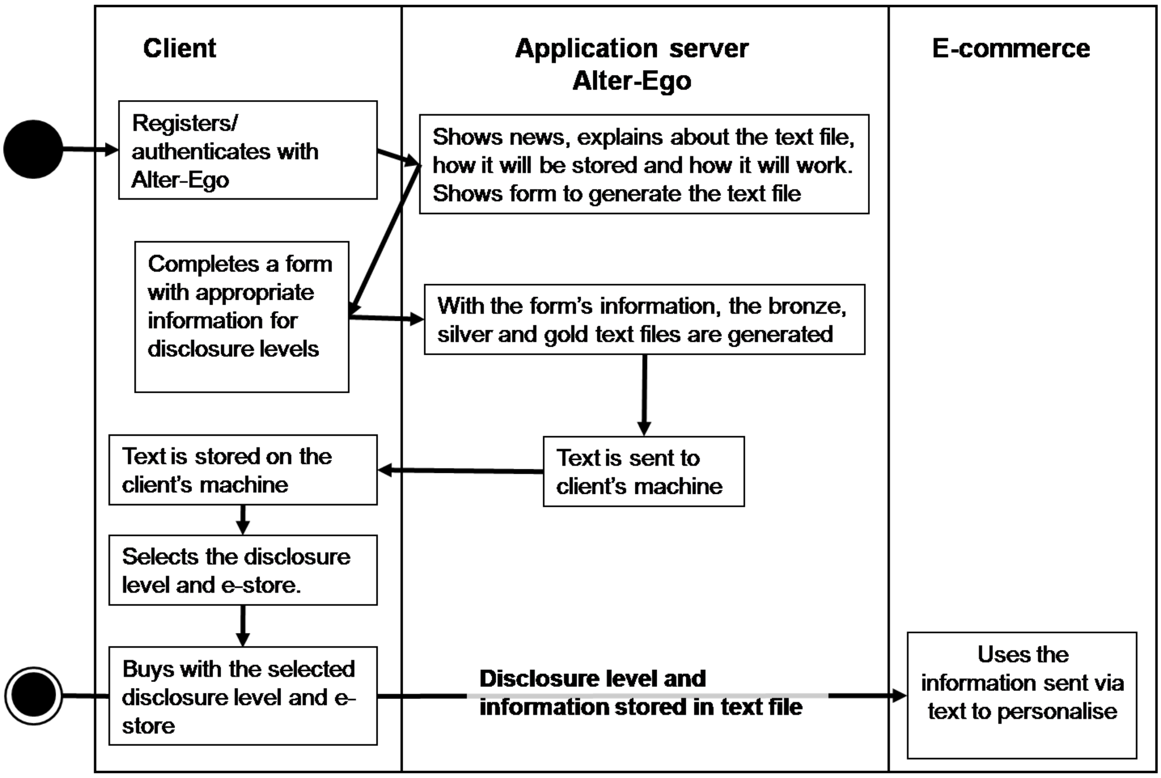

Client
Application server Alter-Ego
E-commerce
Registers/ authenticates with Alter-Ego
Shows news, explains about the text file, how it will be stored and how it will work. Shows form to generate the text file
Completes a form with appropriate information for disclosure levels
With the form's information, the bronze, silver and gold text files are generated
Text is stored on the client's machine
Text is sent to client's machine
Selects the disclosure level and e-store.
Buys with the selected disclosure level and e-store
Disclosure level and information stored in text file
Uses the information sent via text to personalise

Figure 6.3: WebML activity diagram. Design option 2 - Text file.

Implementation: Two ways of implementing this option were tried. In the first, the customers entered a Web site that presented them with a form to be completed online. The customers then saved the generated text file on their own machines. The second option gave rights to the server-side program to store the generated text file on the customer's machine. In both options, JavaScript code contained a form to be completed by the customer. The completed form stored data in a text file.

The process of writing a file on the user's computer was approached in two different ways. In the first static version, the customer was asked to download the generated file and save it in their desired location. In the second, to add dynamism to the process, the text file was stored directly on the customer's computer. To store text directly in the customer's machine and have secure transactions, Park *et al.* [110] proposes the use of Java Web Start ("an easy, robust, and secure way to deploy applications directly from the Web"[144]). Java Web Start is recommended because "it automatically saves the downloaded JAR (Java Archive) files in the client machine at initial activation, and thus eliminates the subsequent download of the same files again when the user executes the application the next time" [110][p618].

However, this process would require the customers to lower the browser security setting to its minimum level to allow direct access to the customer's machine.

Advantages: In this option the user is always in control of the information that is stored on the client computer, with the option of altering it at any time without the need to be connected online. The customer's control is more direct than the cookie option and disclosure can be as limited as the user chooses. This proposal ensures the quality of collected data and solves one storage issue.

Disadvantages: Writing to the customer's machine was the main problem faced during the exploration of this design option. In the case where the customer was asked to download the files, they were faced with an extra work and memory load. At the same time, the process had a lack of dynamism that might translate privacy protecting shopping into a tiresome experience. On the other hand, the case when the information was directly written on the customer's machine required not only a conscious reduction of the browser's security settings that would leave the customer weakened against malicious attacks, but would also require a greater expertise in the customer to know how to modify those settings. Hence, the (more experienced) customer would be in control of the information of three different text files (gold, silver and bronze) on a Web site, but the computer would be in a low level of protection such that any hacker attack would succeed in accessing the computer's records and would jeopardise the rest of the files.

Conclusion: This option was discarded due to the dangers involved, the inconveniences to the users and the lack of dynamism.

6.3.3 Design option 3

Store the customer's personal information online, using a Web portal.

This option gives customers the opportunity to store their information online and use a Web portal constructed based on the PPSE approach (and fulfilling the PLA agreement). In this option, the customer would have a mediating portal to assist their privacy preserved shopping.

Operation: In this option, illustrated in Figure 6.4, customers go to a Web portal that provides them with the elements to assist their shopping. Portals are Web-based applications that provide multiple functionality and information on the same space [1], and have been used in areas as diverse as medicine, commerce, comparative services, etc. The Alter-Ego Web portal would have characteristics of an *information presenting Web portal (IP Web portal)*, providing users with online information and information-related services, together with a channel of communication [39]. Hence, the Alter-Ego Web portal would manage the storage and disclosure of preferences and sensitive information. At the same time, the portal would mediate the customer access to a number of e-commerce sites, e-grocery for this case, with the following main functionalities:

First, the portal would provide to its registered customers an interface to guide and assist them in the storage of their information.

Second, after storing their preferences, customers could be able to, within the Alter-Ego Web portal, easily specify the disclosure level of information and select the participant stores where their information would be disclosed.

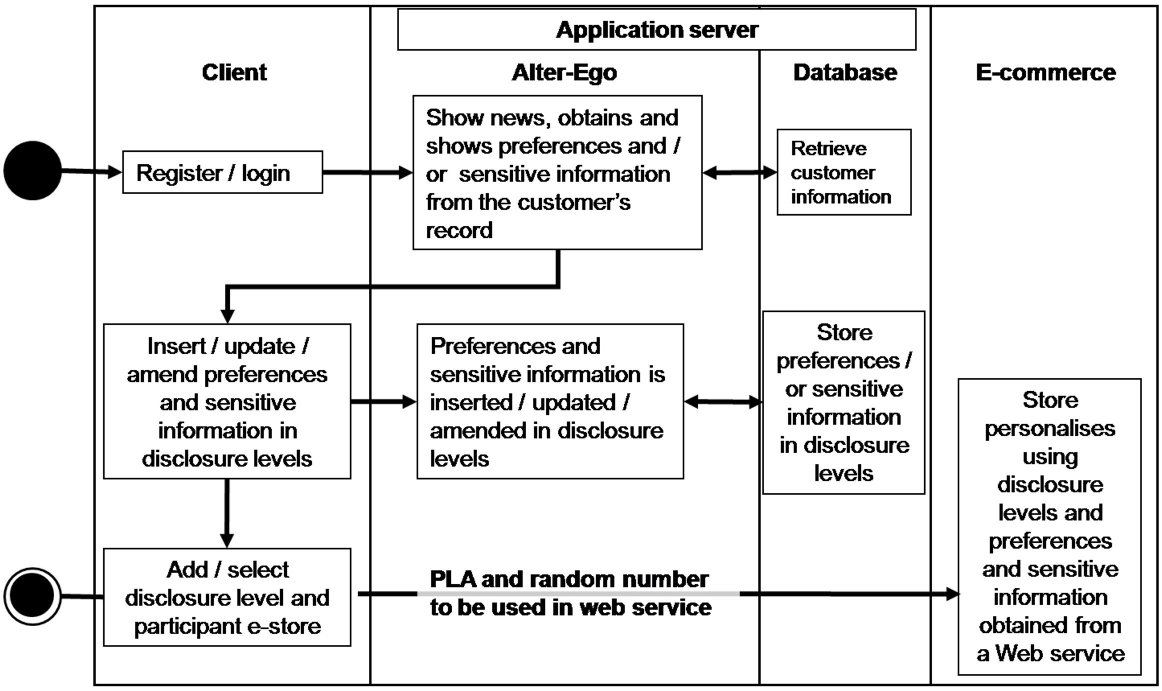

Application server
Client
Alter-Ego
Database
E-commerce
Register / login
Show news, obtains and shows preferences and / or sensitive information from the customer's record
Retrieve customer information
Insert / update / amend preferences and sensitive information in disclosure levels
Preferences and sensitive information is inserted / updated / amended in disclosure levels
Store preferences / or sensitive information in disclosure levels
Store personalises using disclosure levels and preferences and sensitive information obtained from a Web service
Add / select disclosure level and participant e-store
PLA and random number to be used in web service

Figure 6.4: WebML activity diagram. Design option 3 - Web portal.

Therefore, customers visiting the portal would be presented with a means of keeping their information private and, since an authentication or registration process would be required (using an e-mail and password) they would have the advantages of controlling it. For authenticated customers, a different zone would facilitate the storage of their preferences and sensitive information.

To facilitate the customers' disclosure of information, the Web portal would allow customers to select what information, from their stored records, would be disclosed to which participant e-grocery store. A list of participant stores would be listed in the Web portal. A Web service could be used to facilitate a flexible and dynamic exchange of information. A Web service, as shown in Figure 6.5, is an "interface positioned between the application code and the user of that code"[137][p2]. Since a Web service allows any language supporting the Web service to have access to the application's functionality, its use provides the flexibility required by the portal to disclose information to participant stores.

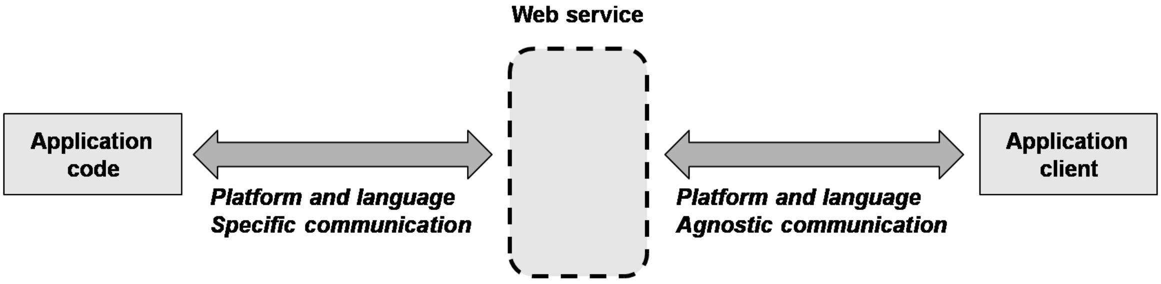

Web service
Application code
Platform and language Specific communication
Platform and language Agnostic communication
Application client

Figure 6.5: Web service. *Adapted from [137]*

Advantages: The advantages of this proposal are:

- In order to be listed on the portal, the participant stores must comply with the PPSE privacy policy, providing the customers with a measure of trust in the integrity of the stores.
- Customers' information would be stored in a controlled and secured environment, giving them the chance of managing the disclosure of their information in a controlled way, as introduced in Section 5.5.2.
- To prevent participant stores from obtaining unlimited customer-related information, when the customer discloses information to a participant store, the Alter-Ego Web portal would send an anonymous identifier and the disclosure level to the selected

store. The participant store would require both data to use the Web service that would deliver the customer's profile upon request.

- The Alter-Ego Web portal would only store preferences and sensitive information. It would not store personal information. Therefore someone hacking the Alter-Ego Web portal would have no chance of identifying the customer, reducing the potential benefit of hacking attacks.
- Customers would not need to enter their preferences each time they visit a different store.
- If customers decide to use the PPSE to assist their shopping for another person with a different set of preferences than themselves, they could change their preferences dynamically when doing their shopping directly in the store, without affecting their original preferences in the Alter-Ego Web portal.
- Since the customers' privacy would not be threatened, the need to protect themselves by giving false information would decrease. Accurate information would more likely be provided, and the e-grocery store could rely on and trust this customer information to perform market segmentation and other marketing studies, as stated in Section 5.2.1.

Disadvantages: As stated in Chapter 5, the main disadvantage posed by the Web portal would be the continuous level of human involvement required to have the Web portal working under optimal conditions. For instance, the information presented to create awareness (also called *news*) would need to be kept updated.

Therefore, since the advantages outweigh the disadvantages, and assuming that the maintenance problems were already solved, the PPSE uses a trusted third Web party portal to implement the Alter-Ego. The following Section will discuss the implementation.

6.4 Alter-Ego Web portal

The conceptual framework of the trusted third party, Alter-Ego portal ensures that while the e-grocery store collects personal information from the customer, the sensitive information and preferences are held separately. By having the Alter-Ego implemented as a portal, as shown in Figure 6.6, with the customer's non-identifiable preferences and sensitive information stored in Alter-Ego's database, the customer can choose any of the participant stores to disclose as much or as little information as desired [47]. When the customer chooses the disclosure level that will be used to do their shopping with the e-grocery of their choice, the Alter-Ego portal sends the information (an anonymous identifier and the disclosure level) to the selected e-grocery. The store uses a Web service to retrieve the customer's information. Using the retrieved information, the e-grocery store can then offer the agreed level of "*personalised personalisation*" [48]. Customers retain control over their disclosed preferences and sensitive details.

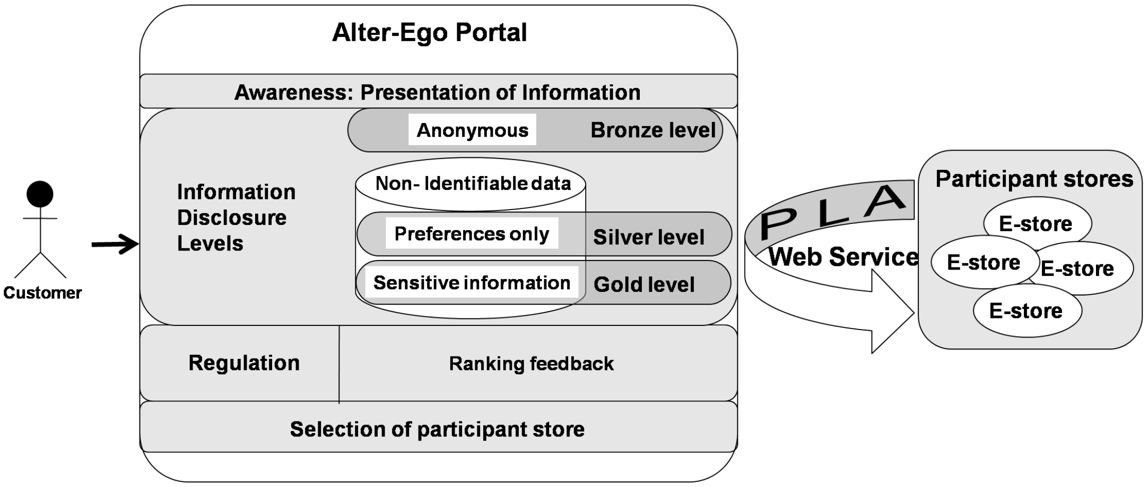


Figure 6.6: Alter-Ego Web portal.

To measure the user perceived service quality of information on Web portals, Yang *et al.* determines five service quality dimensions for IP Web portals: usability, usefulness of content, adequacy of information, accessibility and interaction [161]. To these dimensions, Lin *et al.* adds the importance of playfulness in computer mediated environment as an important element in the customers' satisfaction and an encouraging factor in their continued use of a Web site [82]. These dimensions were taken into consideration during the design and implementation process. The design and implementation of the Alter-Ego Web portal is presented next.

6.4.1 Alter-Ego Web portal - Overview

After authenticating, customers enter the Alter-Ego Web portal home page. From the home page, they have different navigation options, as shown in Figure 6.7. They can decide to enter, edit or delete their information using the links to silver or gold levels. Customers have also the choice of going to "*select & shop*" to select their disclosure level and the store to which their selected information will be disclosed.

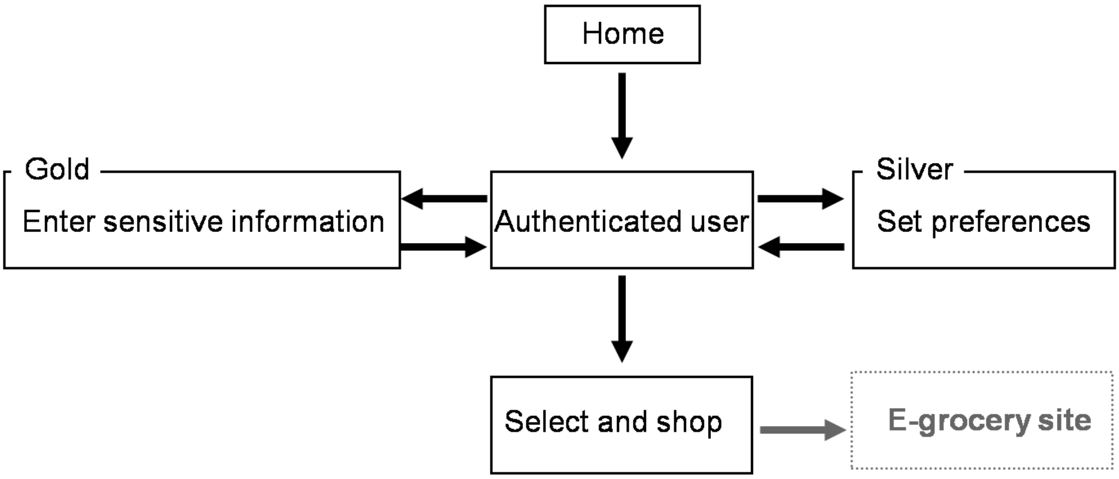


Figure 6.7: Alter-Ego Web portal - User state transition diagram - Overview.

The WebML activity diagram of the Alter-Ego Web portal overview, shown in Figure 6.8, shows the customer registration/login process. The entering/updating/deleting preferences, for silver level, the entering, updating and deleting sensitive information, for gold level. Finally, it shows the selection of the disclosure level and participant store.

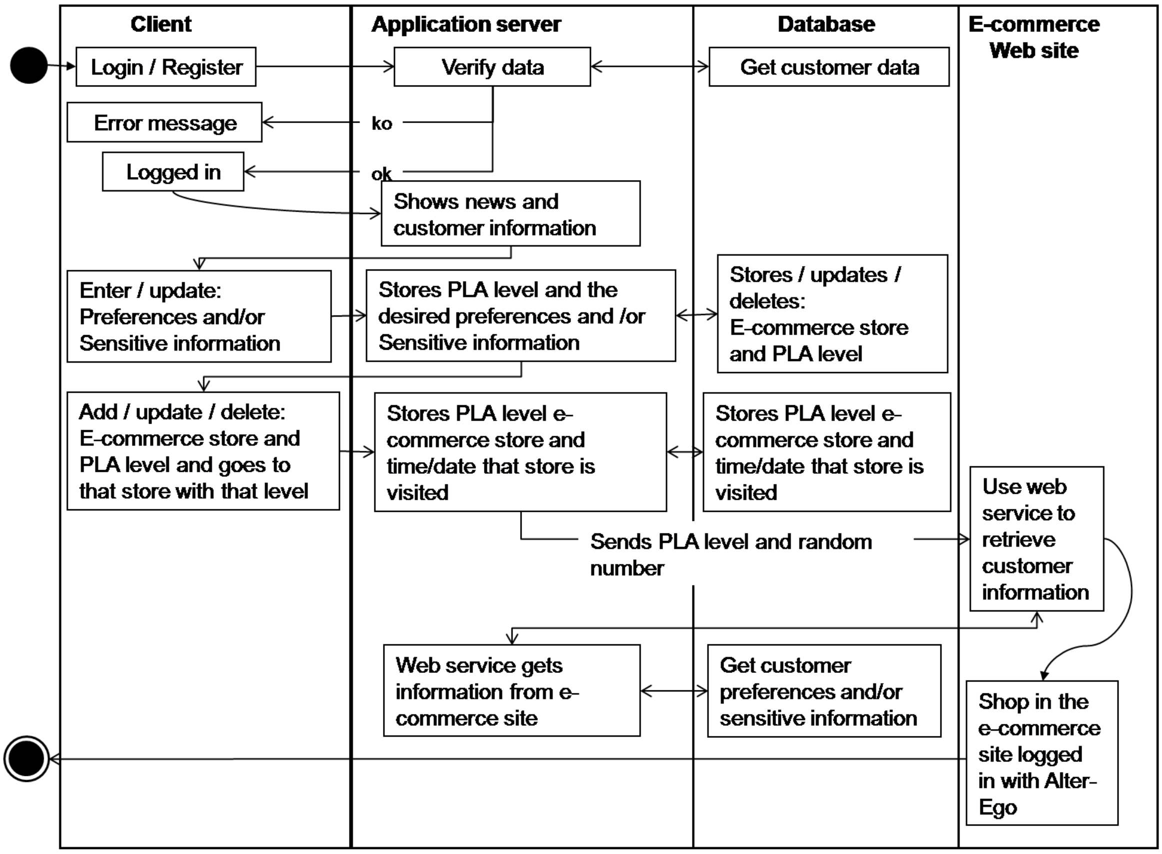

Client
Application server
Database
E-commerce Web site
Login / Register
Verify data
Get customer data
Error message
ko
Logged in
ok
Shows news and customer information
Enter / update: Preferences and/or Sensitive information
Stores PLA level and the desired preferences and /or Sensitive information
Stores / updates / deletes: E-commerce store and PLA level
Add / update / delete: E-commerce store and PLA level and goes to that store with that level
Stores PLA level e-commerce store and time/date that store is visited
Stores PLA level e-commerce store and time/date that store is visited
Use web service to retrieve customer information
Sends PLA level and random number
Web service gets information from e-commerce site
Get customer preferences and/or sensitive information
Shop in the e-commerce site logged in with Alter-Ego

Figure 6.8: Alter-Ego Web portal - WebML activity diagram - Overview.

The Alter-Ego Web portal home page layout is shown in Figure 6.9, showing the content area that contains the privacy news and information to create awareness in the interested customers. The login area is located on the right hand side. During the layout construction of the entire Alter-Ego Web portal, the opinion of five colleagues from the Department of Computing Science of the University of Glasgow was requested. The five participants volunteered to do an exploration of Alter-Ego Web portal, *speaking out loud* about any difficulties that they found and any suggestions to improve the navigation and the usability. Their opinions were recorded and notes about their navigation, their reactions to the presentation and content of the diverse elements of the portal were taken. Changes to the layout were made when the opinion of at least three participants concurred. As a result of this feedback, the Alter-Ego Web portal situates the login area in the right hand side (shown in Figure 6.10), while the registration is on the left hand side (shown in Figure 6.13 and Figure 6.15).

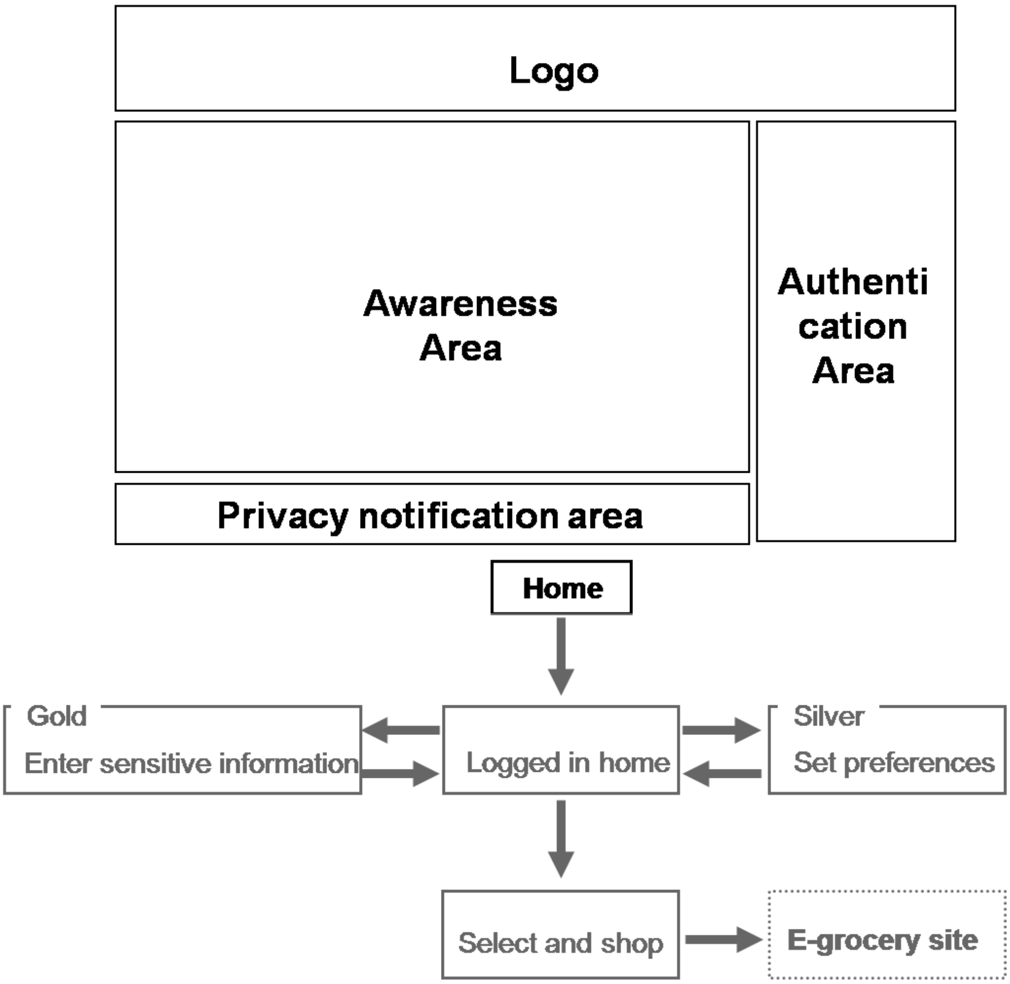
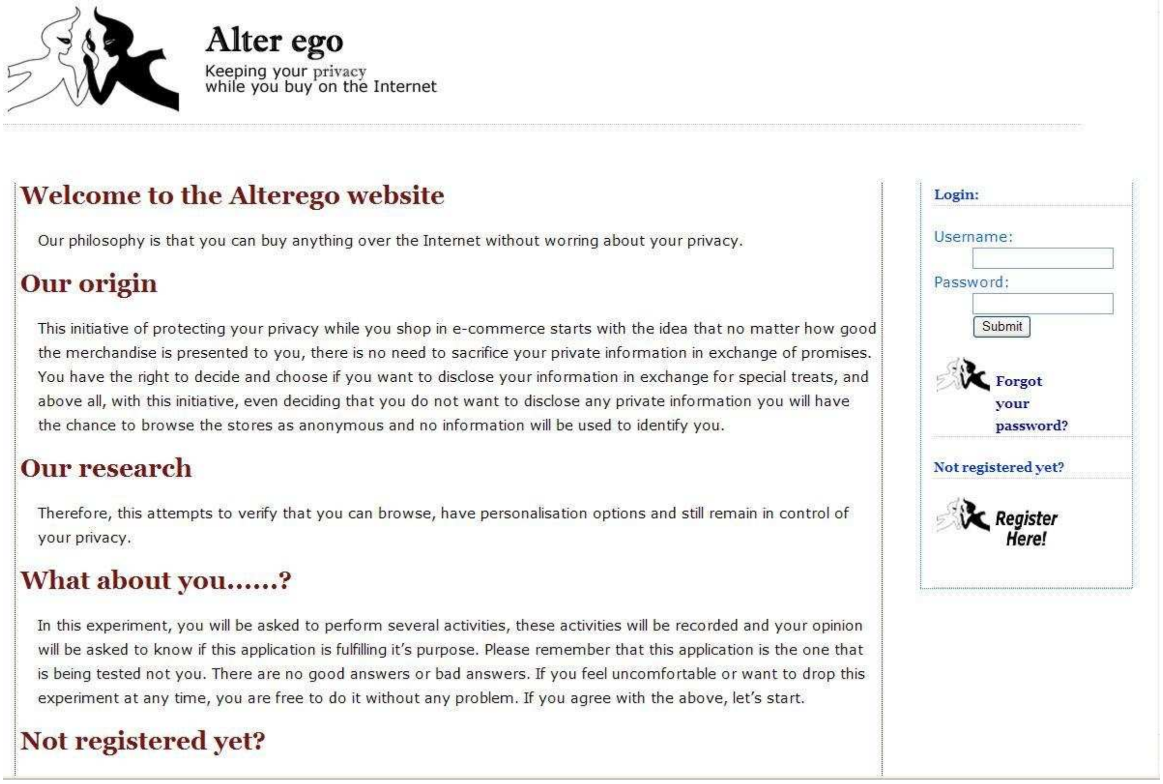


Figure 6.9: Alter-Ego Web portal - Home page - Layout.

The implementation of the Alter-Ego Web portal home page, shown in Figure 6.10, includes a logo composed of an image and a phrase. The image has the purpose of visually reinforcing the idea of duality and mediating environment and the phrase “Keeping your privacy while you buy on the Internet” is used as a message to the customers to clarify the purpose of the portal. A strong contrast between background and text is used for the presented content, using black text over white background, due to the acceptance and display quality that this combination presents [83] .



Welcome to the Alterego website

Our philosophy is that you can buy anything over the Internet without worring about your privacy.

Our origin

This initiative of protecting your privacy while you shop in e-commerce starts with the idea that no matter how good the merchandise is presented to you, there is no need to sacrifice your private information in exchange of promises. You have the right to decide and choose if you want to disclose your information in exchange for special treats, and above all, with this initiative, even deciding that you do not want to disclose any private information you will have the chance to browse the stores as anonymous and no information will be used to identify you.

Our research

Therefore, this attempts to verify that you can browse, have personalisation options and still remain in control of your privacy.

What about you......?

In this experiment, you will be asked to perform several activities, these activities will be recorded and your opinion will be asked to know if this application is fulfilling it's purpose. Please remember that this application is the one that is being tested not you. There are no good answers or bad answers. If you feel uncomfortable or want to drop this experiment at any time, you are free to do it without any problem. If you agree with the above, let's start.

Not registered yet?

Figure 6.10: Alter-Ego Web portal - Home page (login) - Implementation.

6.4.2 Alter-Ego Web portal - Architecture

The implementation of the Alter-Ego Web portal uses a multi-layered architecture [148, 51], containing a Web server (running PHP), a database (MySQL) and a Web service. This is shown in Figure 6.11. Although this architecture is primarily designed for distributed applications, it has proved possible to situate the entire PPSE environment with a single host, so that it could easily be used for the proof of concept.

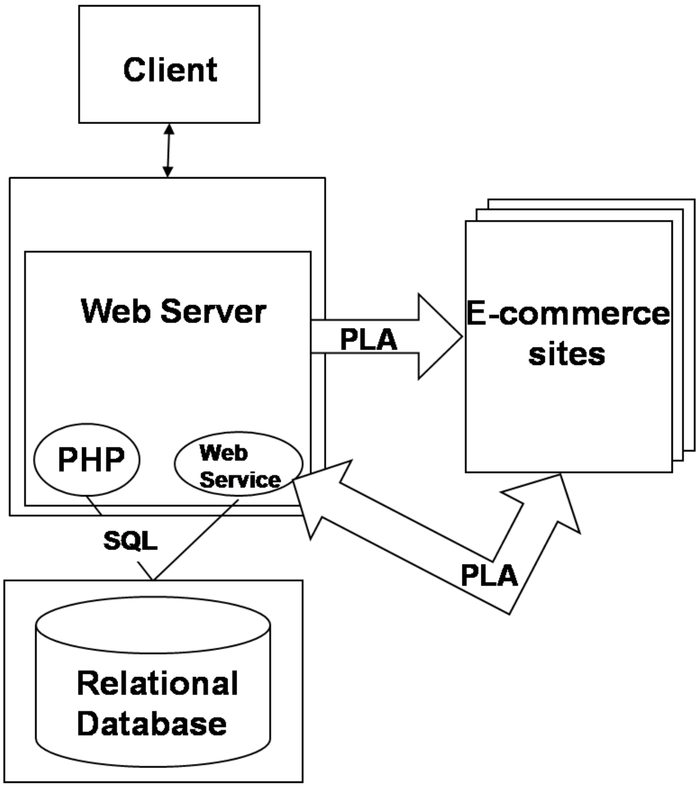


Figure 6.11: Alter-Ego Web portal - Architecture.

However, in the case of implementing this proposal in a more formal and operational way, the Alter-Ego Web portal architecture should contemplate a more robust design such as that shown in Figure 6.12.

This more robust architecture considers the activation of a backup machine via switches. The switch activates in case of failure, and a backup for the switch is present should the first switch fail. Session data is stored to maintain the customer's session state during their visit. Therefore, in case of failure, and switching from the original Web server to its backup, this change would not be perceived by the customer, allowing transparent recovery. Finally, a replication of the database would ensure a preservation of the data in case of any failure.

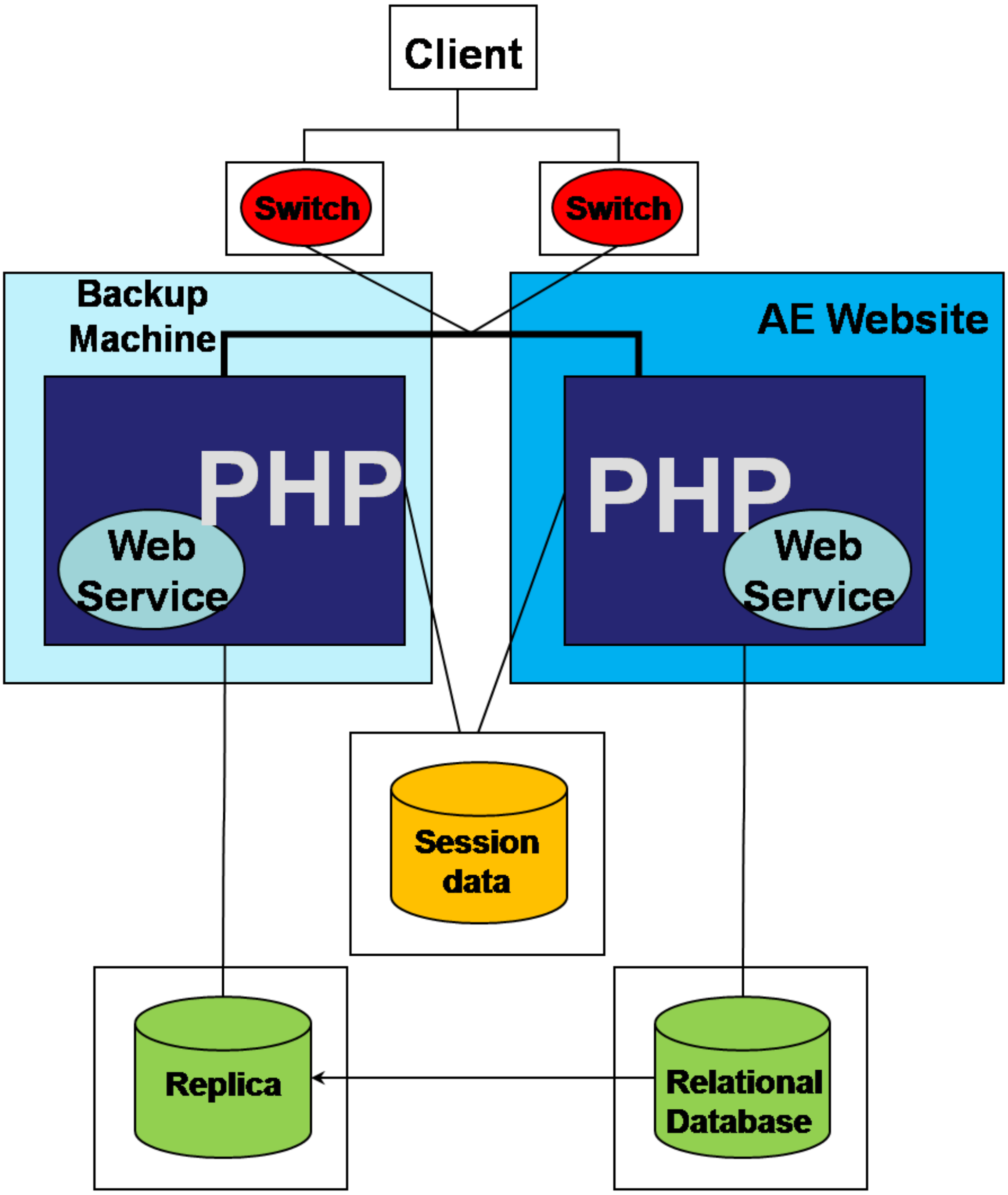


Figure 6.12: Architecture proposed for the Alter-Ego Web portal in case of operational implementation.

6.4.3 Alter-Ego Web portal - Register/login

The registration area is placed on the left hand side of the screen, as shown in Figure 6.13. As well as the preferences expressed by the five volunteers during the usability test, two other factors contributed to the decision of locating the registration area in a different position from the login area. The first relates to the fact that the rest of the Alter-Ego's navigation menu is located on the left hand side of the layout. Having the registration and the menu on the left hand side would give the customers a visual continuity while navigating the rest of the portal. The second factor is the F-shape pattern that users follow when they read Web content. Firstly, users follow a horizontal movement followed by a second horizontal movement in a shorter area than the first. Finally, users scan the content situated on the left in a top-down vertical movement [99].

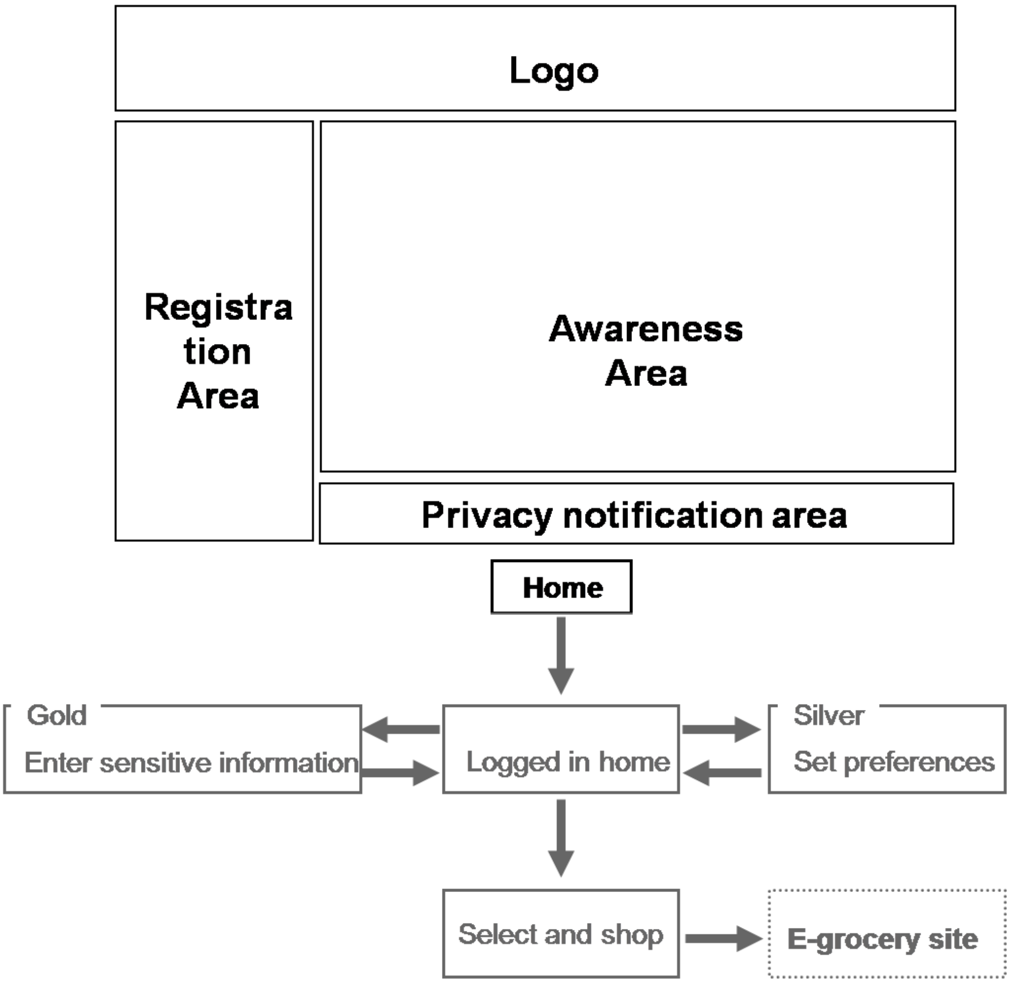


Figure 6.13: Alter-Ego Web portal - Register page - Layout.

The detail of the registration and logging in activities are shown in the WebML activity diagram presented in Figure 6.14. Newcomers, are required to register by providing a valid e-mail address and password. This information is verified against the records of registered customers to verify that there is no previous registration under those specific details. If this search is unsuccessful, the customer is registered, otherwise, an error message appears. The login process follows a similar process. In this case, customers provide the e-mail address and password which they provided at registration, and the system verifies the information against the database records and allows or denies access. In case of forgotten passwords, a reminder can be sent to the registered e-mail.

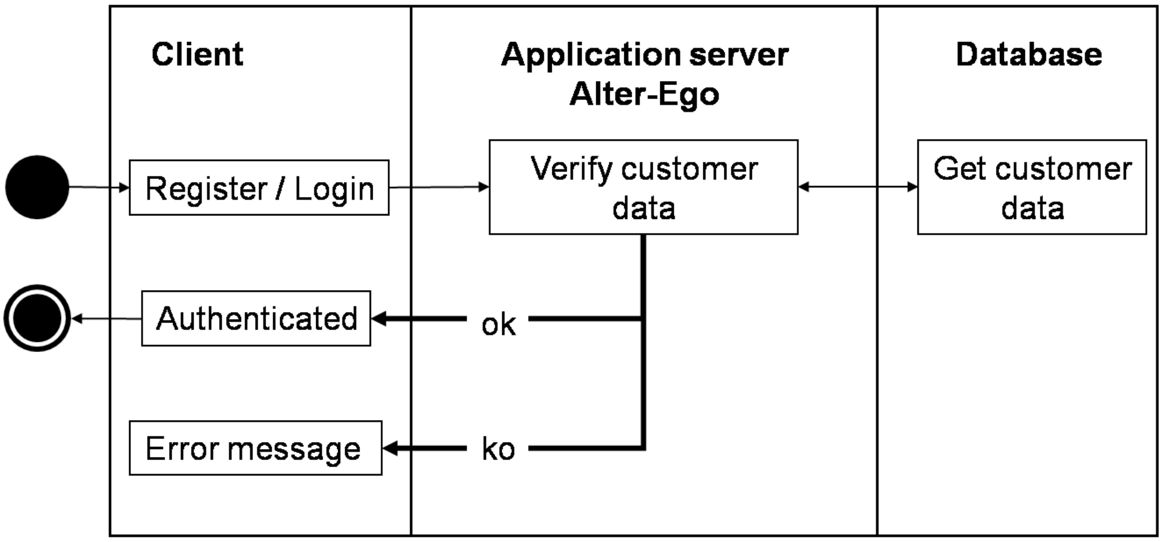
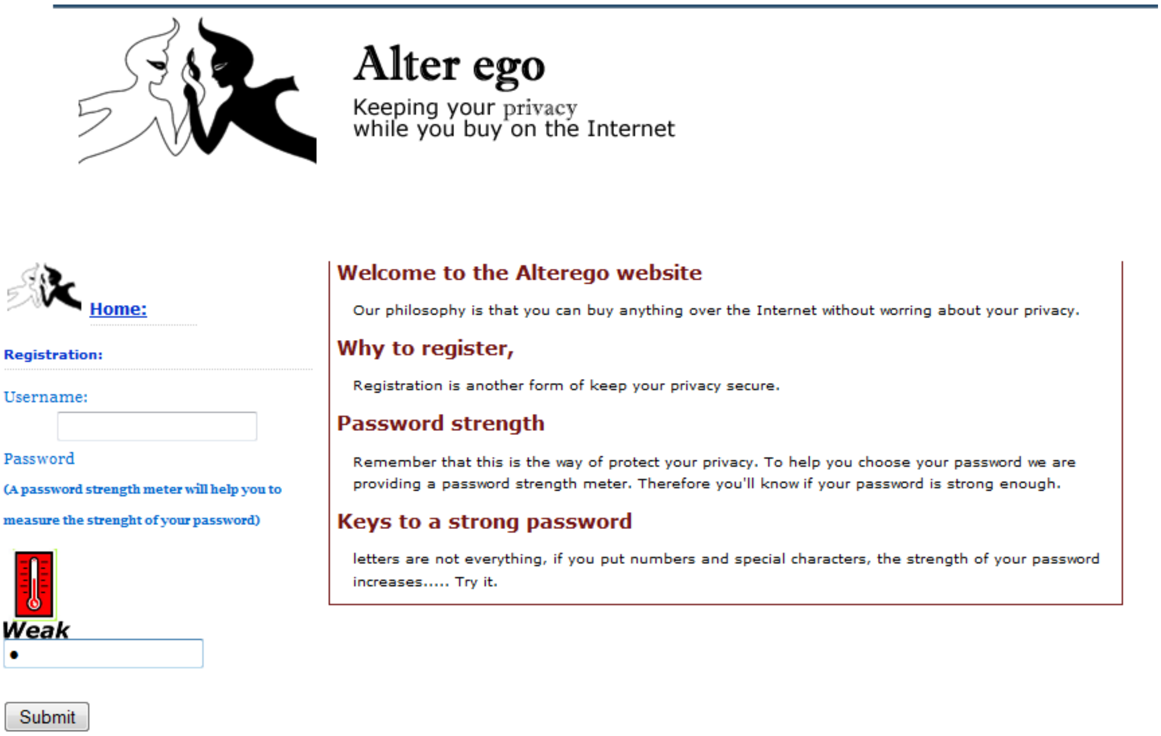


Figure 6.14: Alter-Ego Web portal - Register/login - WebML activity diagram.

During the implementation, a visual aid is included to assist the customers in their password selection. As the customer enters the password, the icon changes as the password increases or decreases its strength. Passwords are catalogued as *weak, medium* or *strong.* Consequently, passwords that contain only letters or are short (less than 6 characters) are considered weak, while a combination of length, letters, numbers and special characters increases the strength of the password [28, 86]. The implementation of the Alter-Ego Web portal registration process is shown in Figure 6.15.



Alter ego

Keeping your privacy while you buy on the Internet

Home:

Registration:

Username:

Password

(A password strength meter will help you to measure the strenght of your password)

Weak

●

Welcome to the Alterego website

Our philosophy is that you can buy anything over the Internet without worring about your privacy.

Why to register,

Registration is another form of keep your privacy secure.

Password strength

Remember that this is the way of protect your privacy. To help you choose your password we are providing a password strength meter. Therefore you'll know if your password is strong enough.

Keys to a strong password

letters are not everything, if you put numbers and special characters, the strength of your password increases..... Try it.

Figure 6.15: Alter-Ego Web portal - Registration - Implementation.

After customers successfully log in, they are presented with the Alter-Ego Web portal home page, as shown in Figure 6.16. From this home page, customers have the following options: reading the presented information (to raise their privacy awareness), interact with the disclosure levels entering or modifying their information (as shown in Figures 6.19, 6.22 and 6.24), or select the disclosure level and e-grocery store to shop (as shown in Figure 6.28).

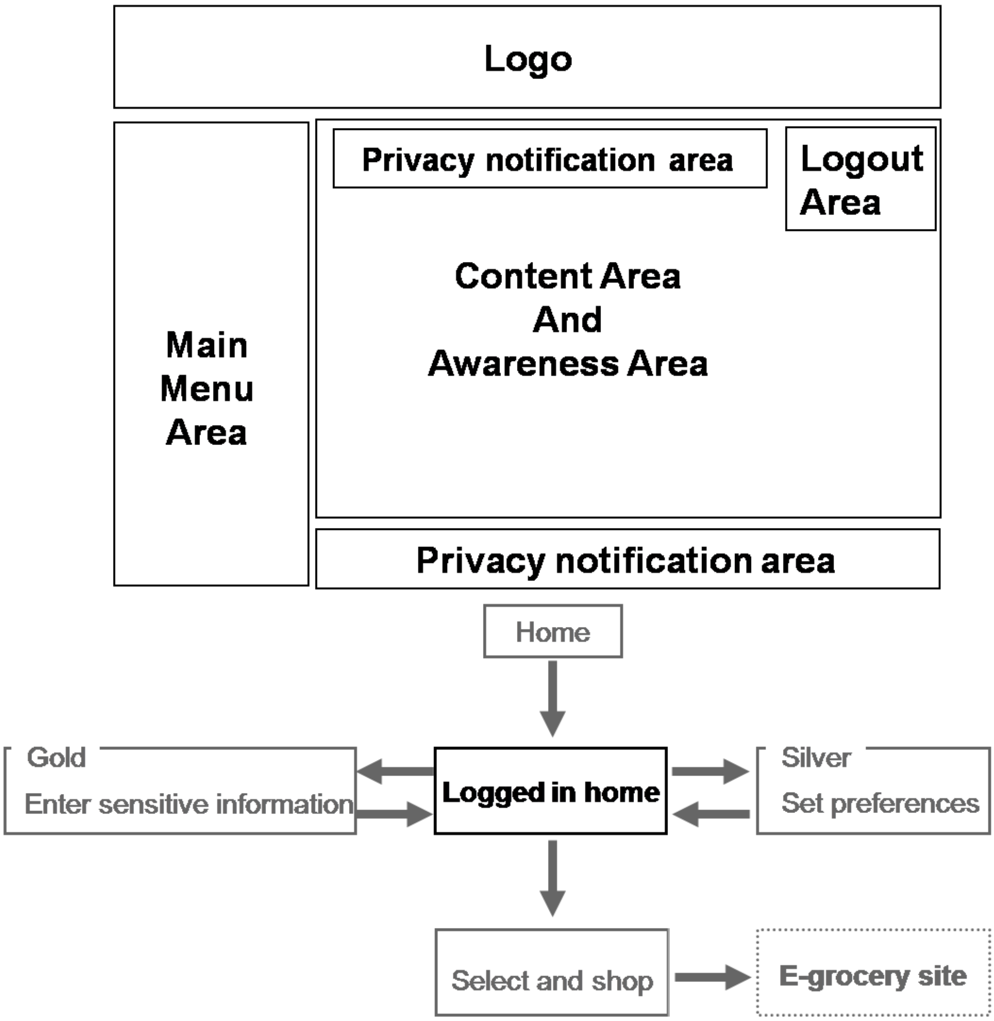
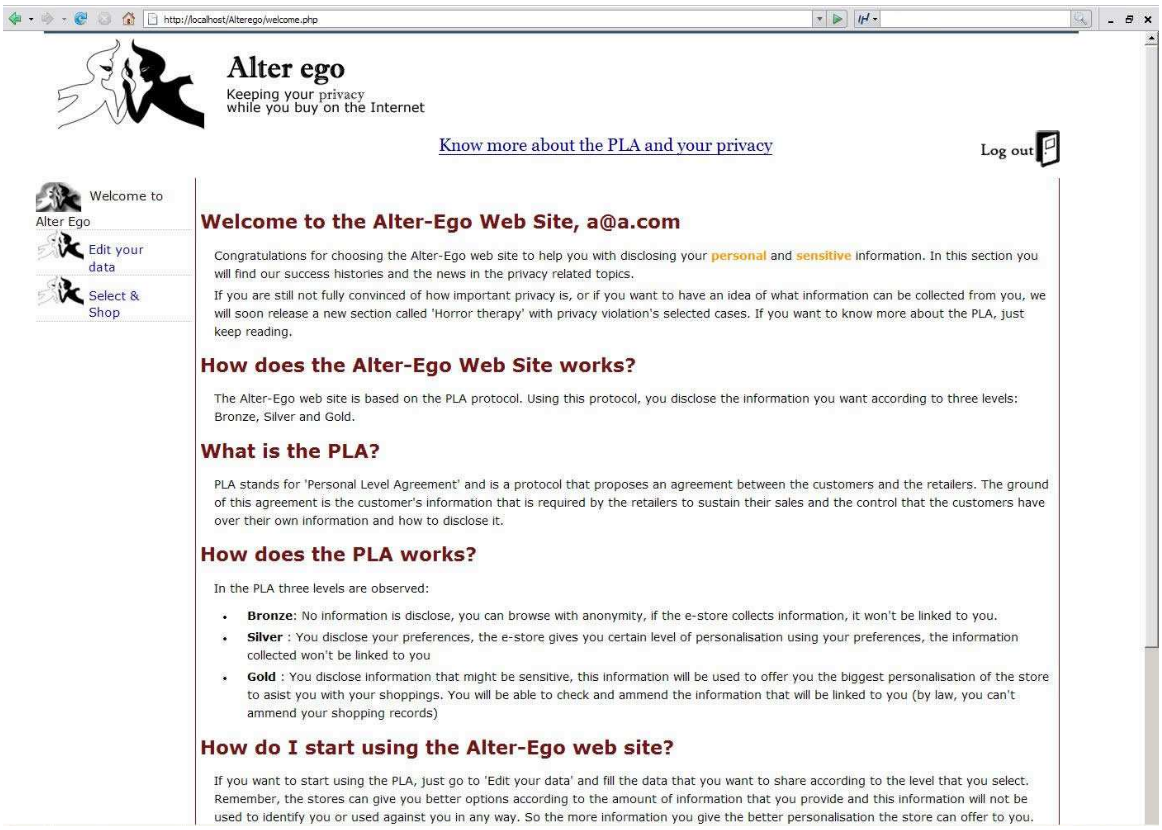

Logo
Privacy notification area
Logout Area
Main Menu Area
Content Area And Awareness Area
Privacy notification area
Home
Gold
Enter sensitive information
Logged in home
Silver
Set preferences
Select and shop
E-grocery site

Figure 6.16: Alter-Ego Web portal - Logged in home page - Layout and transition diagram.

In the implementation of the Alter-Ego Web portal home page, shown in Figure 6.17, the F-shape reading pattern is used to guide the customers' attention to particular points. In the first horizontal line the logo is presented, in the second line of the F-shape, links to allow the customers to learn more about privacy and logout are presented. The vertical of the F-shape reading pattern contains the navigation menu to the rest of the portal.


http://localhost/Alterego/welcome.php
Alter ego
Keeping your privacy
while you buy on the Internet
Know more about the PLA and your privacy
Log out
Welcome to Alter Ego
Edit your data
Select & Shop

Welcome to the Alter-Ego Web Site, a@a.com

Congratulations for choosing the Alter-Ego web site to help you with disclosing your personal and sensitive information. In this section you will find our success histories and the news in the privacy related topics.

If you are still not fully convinced of how important privacy is, or if you want to have an idea of what information can be collected from you, we will soon release a new section called 'Horror therapy' with privacy violation's selected cases. If you want to know more about the PLA, just keep reading.

How does the Alter-Ego Web Site works?

The Alter-Ego web site is based on the PLA protocol. Using this protocol, you disclose the information you want according to three levels: Bronze, Silver and Gold.

What is the PLA?

PLA stands for 'Personal Level Agreement' and is a protocol that proposes an agreement between the customers and the retailers. The ground of this agreement is the customer's information that is required by the retailers to sustain their sales and the control that the customers have over their own information and how to disclose it.

How does the PLA works?

In the PLA three levels are observed:

- **Bronze**: No information is disclose, you can browse with anonymity, if the e-store collects information, it won't be linked to you.
- **Silver** : You disclose your preferences, the e-store gives you certain level of personalisation using your preferences, the information collected won't be linked to you
- **Gold** : You disclose information that might be sensitive, this information will be used to offer you the biggest personalisation of the store to asist you with your shoppings. You will be able to check and ammend the information that will be linked to you (by law, you can't ammend your shopping records)

How do I start using the Alter-Ego web site?

If you want to start using the PLA, just go to 'Edit your data' and fill the data that you want to share according to the level that you select. Remember, the stores can give you better options according to the amount of information that you provide and this information will not be used to identify you or used against you in any way. So the more information you give the better personalisation the store can offer to you.

Figure 6.17: Alter-Ego Web portal - Logged in home page - Implementation.

6.4.4 PPSE - Disclosure levels

As discussed in Chapter 5, there are three disclosure levels which Alter-Ego Web portal must manage: *bronze, silver* and *gold.* From the Alter-Ego Web portal home page, customers who want to edit their information in the disclosure levels are presented with a horizontal menu which facilitates navigation between levels. From the home page, and before selecting any disclosure level, the content area presents their characteristics with information describing them, and instructions on how to use them. At the same time, a logout icon is presented in case the customer decides to end the session. The layout and transition diagrams are shown in Figure 6.18.

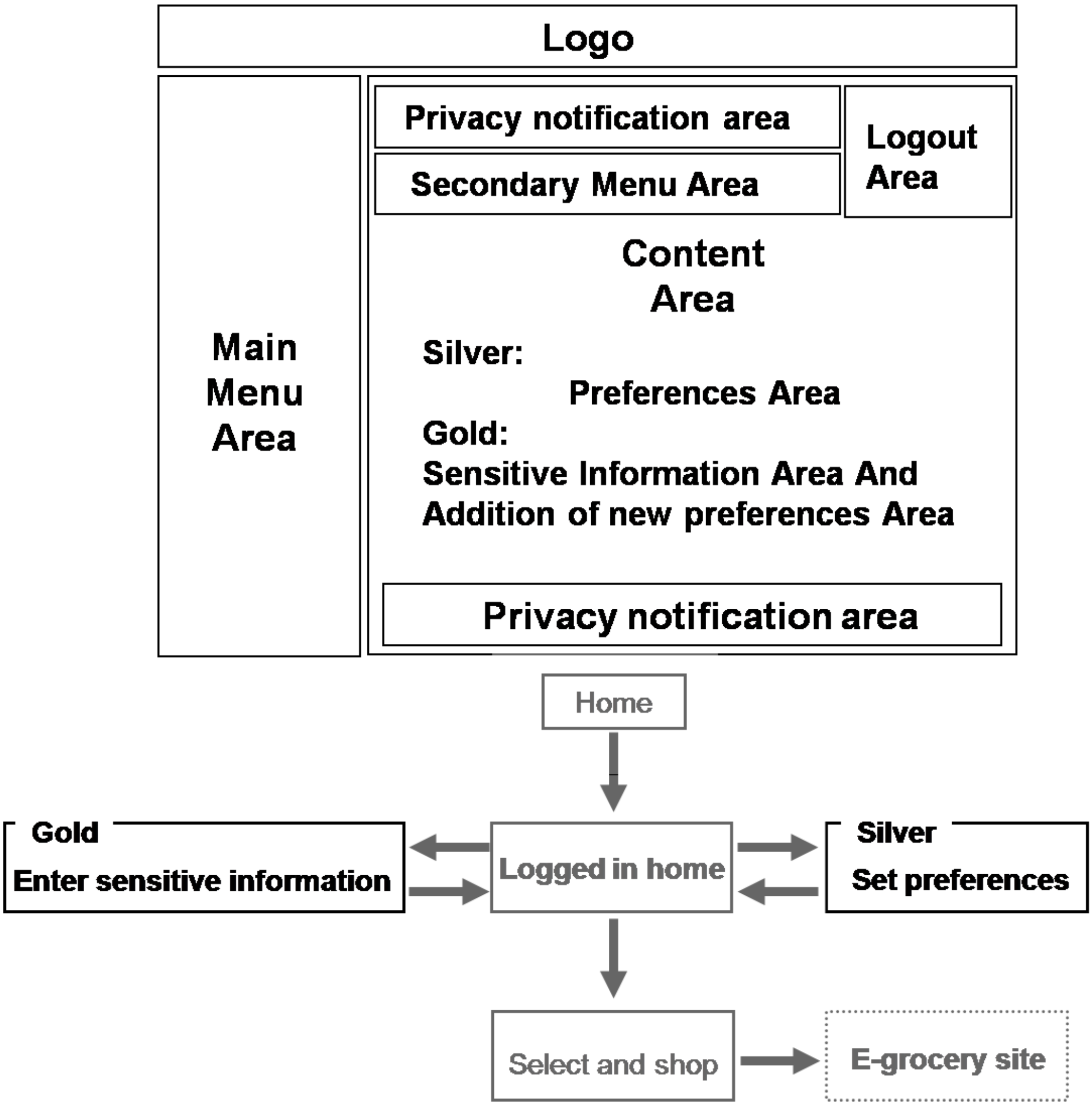
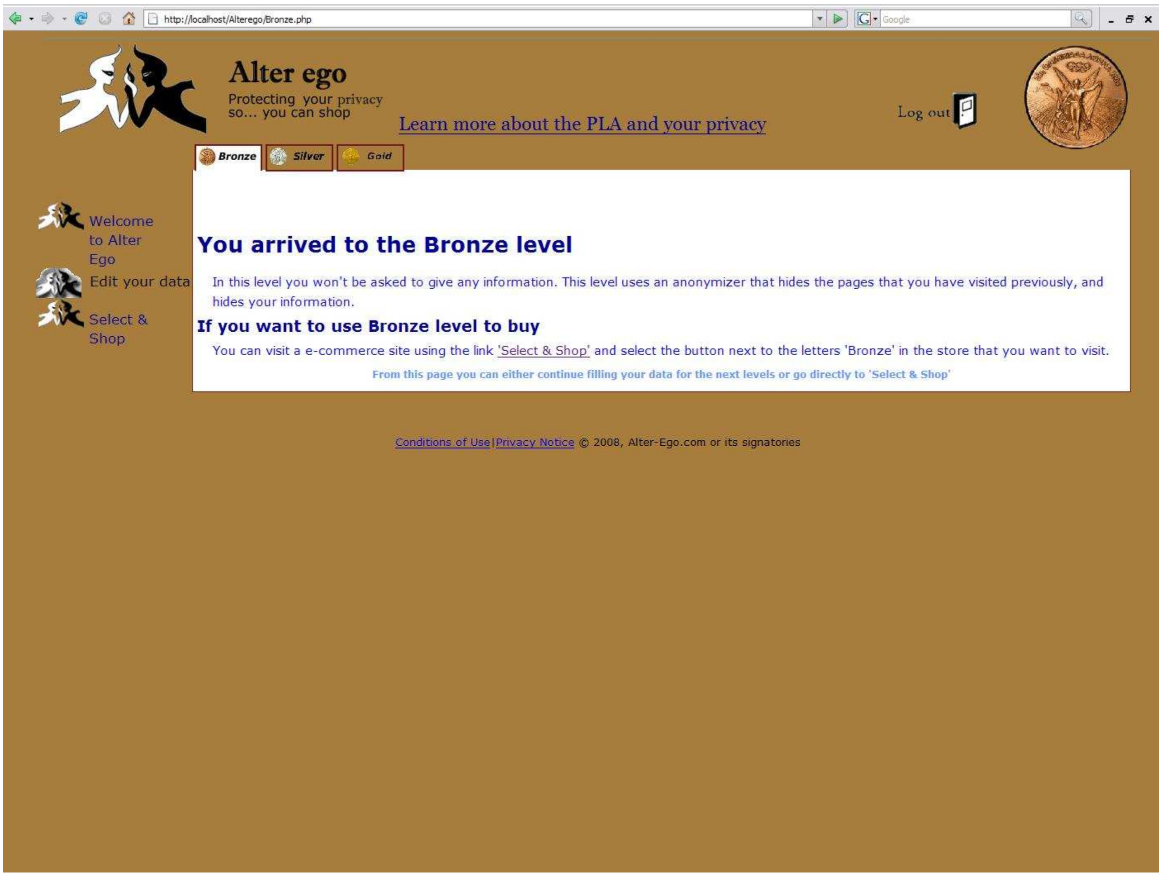


Figure 6.18: Alter-Ego Web portal - Disclosure levels - Layout and transition diagram.

To edit the information in the disclosure levels, the customer follows the link called “Edit your data” to arrive at the first disclosure level: bronze. The F-shape reading pattern is also used as an auxiliary guide for the implementation of the disclosure levels. In this case, the first horizontal line contains the logo, the link to learning more about privacy and the PLA, the logout link and an image of a medal to indicate the current level. The background colour (bronze, silver or gold), and a highlighted banner in the auxiliary horizontal menu are also used to contextualise the page. The second horizontal line of the F-shape contains an auxiliary menu (tabs) that allows navigationbetween disclosure levels. Finally, the vertical line contains the navigation to the rest of the portal. Regardless of the background colour location reinforcement, the content area is left white to facilitate reading. The implementation of the bronze level is shown in Figure 6.19.



You arrived to the Bronze level

In this level you won't be asked to give any information. This level uses an anonymizer that hides the pages that you have visited previously, and hides your information.

If you want to use Bronze level to buy

You can visit a e-commerce site using the link 'Select & Shop' and select the button next to the letters 'Bronze' in the store that you want to visit.

From this page you can either continue filling your data for the next levels or go directly to 'Select & Shop'

Conditions of Use|Privacy Notice © 2008, Alter-Ego.com or its signatories

Figure 6.19: Alter-Ego Web portal - Disclosure levels - Bronze implementation .

6.4.5 Silver level

The WebML activity diagram, in Figure 6.20, shows the entire process of entering silver level preferences. When the customer enters this page for the first time, all the preference displays are empty and ready to be populated. If the customer does not select preferences, Alter-Ego Web portal uses a default set of preferences.

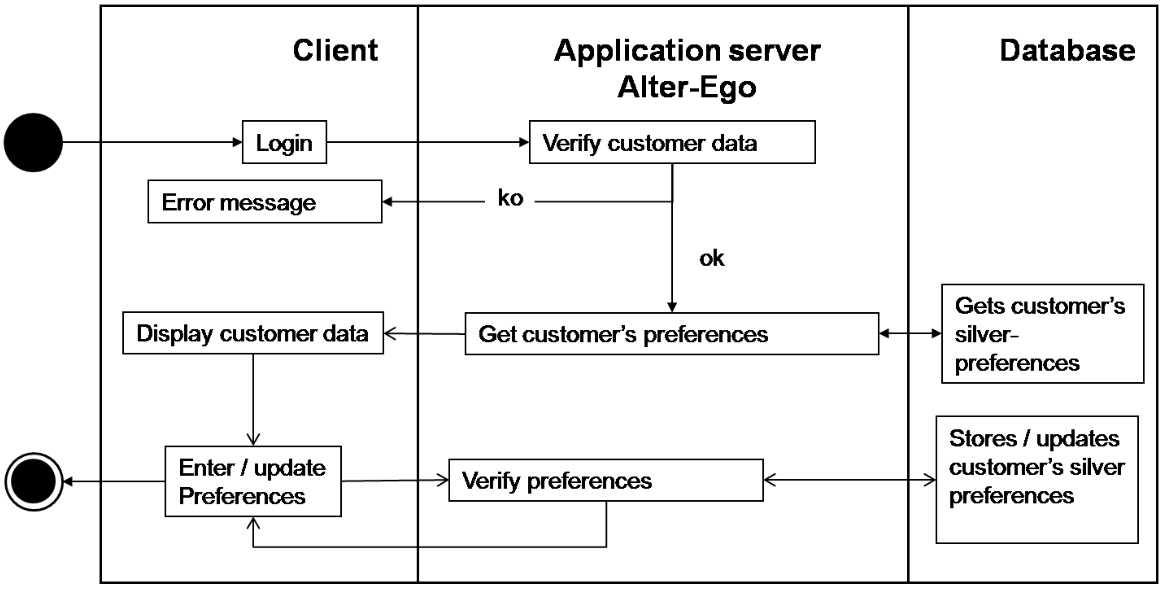

Client
Application server Alter-Ego
Database
Login
Verify customer data
Error message
ko
ok
Display customer data
Get customer's preferences
Gets customer's silver-preferences
Enter / update Preferences
Verify preferences
Stores / updates customer's silver preferences

Figure 6.20: Alter-Ego Web portal - Silver level - WebML activity diagram.

As introduced in Section 5.5.2, five categories are presented to the customers to match their *intensity of preferences*. These categories allow a finer granularity during the selection of the customers' options. The *intensity of preference* categories are: *always, sometimes, maybe, never,* and *don't care*. In the implementation, an icon is assigned to each of these categories to visually assist customers when choosing their selection. The icons are shown in Figure 6.21. At the same time, to increase the Alter-Ego stickiness, these icons were selected to give the Web site a sense of playfulness [82]. The implementation of the silver disclosure level is shown in Figure 6.22.

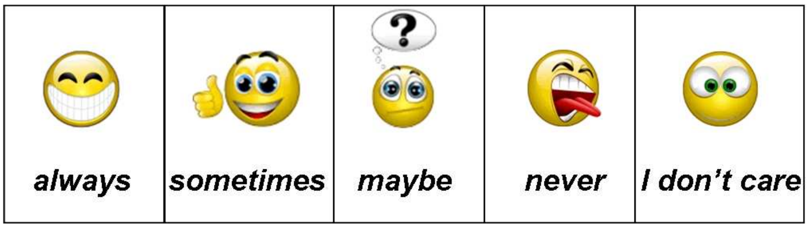
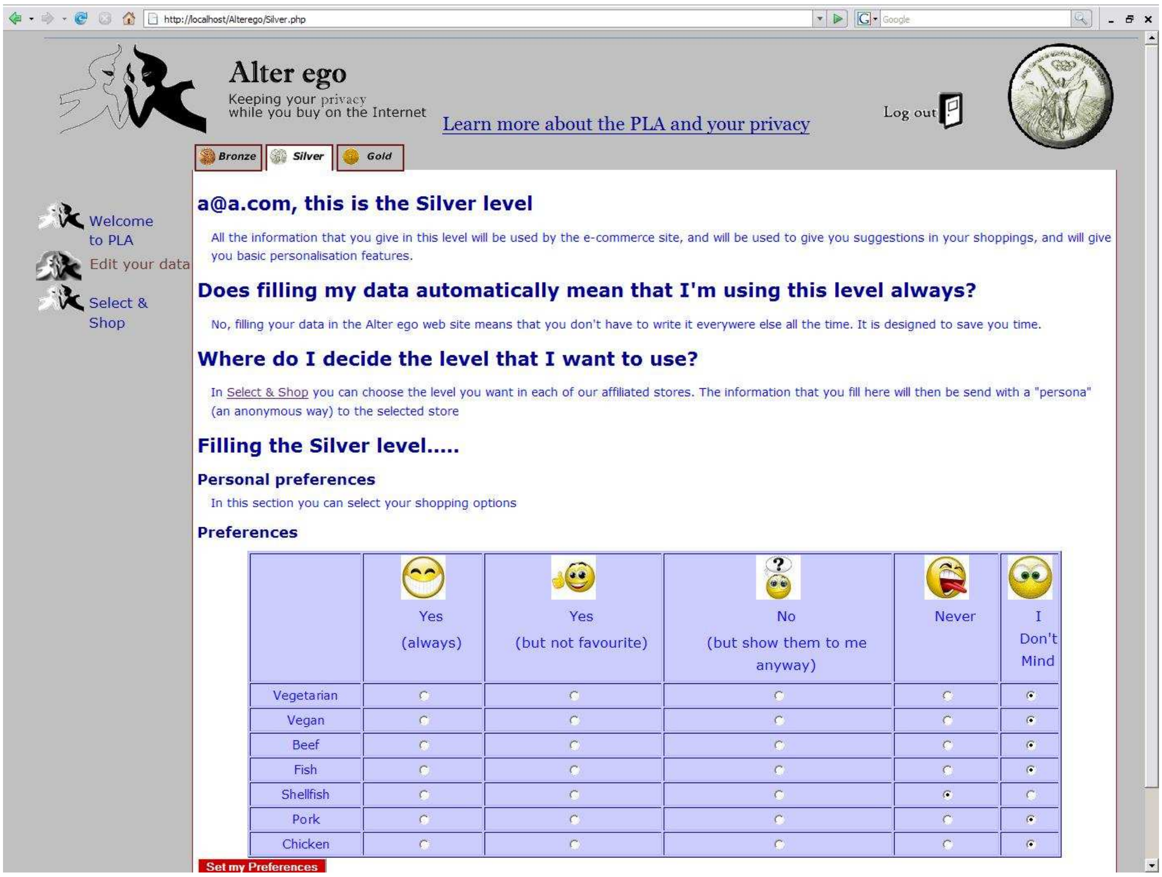


Figure 6.21: Icons used to guide customers while choosing their desired *intensity of preferences* granularity.



http://localhost/Alterego/Silver.php
Google
Alter ego
Keeping your privacy
while you buy on the Internet
Learn more about the PLA and your privacy
Log out
Bronze
Silver
Gold
Welcome to PLA
Edit your data
Select & Shop

a@a.com, this is the Silver level

All the information that you give in this level will be used by the e-commerce site, and will be used to give you suggestions in your shoppings, and will give you basic personalisation features.

Does filling my data automatically mean that I'm using this level always?

No, filling your data in the Alter ego web site means that you don't have to write it everywere else all the time. It is designed to save you time.

Where do I decide the level that I want to use?

In Select & Shop you can choose the level you want in each of our affiliated stores. The information that you fill here will then be send with a "persona" (an anonymous way) to the selected store

Filling the Silver level.....

Personal preferences

In this section you can select your shopping options

Preferences

	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never	I Don't Mind
Vegetarian	○	○	○	○	●
Vegan	○	○	○	○	●
Beef	○	○	○	○	●
Fish	○	○	○	○	●
Shellfish	○	○	○	●	○
Pork	○	○	○	○	●
Chicken	○	○	○	○	●

Set my Preferences

Figure 6.22: Alter-Ego Web portal - Silver level - Implementation.

6.4.6 Gold level

At the gold level, the disclosure options presented to customers can be categorised as sensitive information. As explained in Section 5.5.2, the misuse of this information can have a serious impact on the customer's privacy i.e. reputation. The silver and gold levels layout and transition diagram are shared, as shown in Figure 6.18. However, gold level has an important difference compared with the other levels: the flexibility of introducing new features to be used during the customer's entry of information process.

As shown in Figure 6.23, the gold level process shows customers the sensitive data to be stored in the database. Each of the options contain the granularity provided by the five *intensity of preference* categories. However, the gold level also includes a section where customers can add new specifications to be used by the store to assist the personalisation process based on the customers' options.

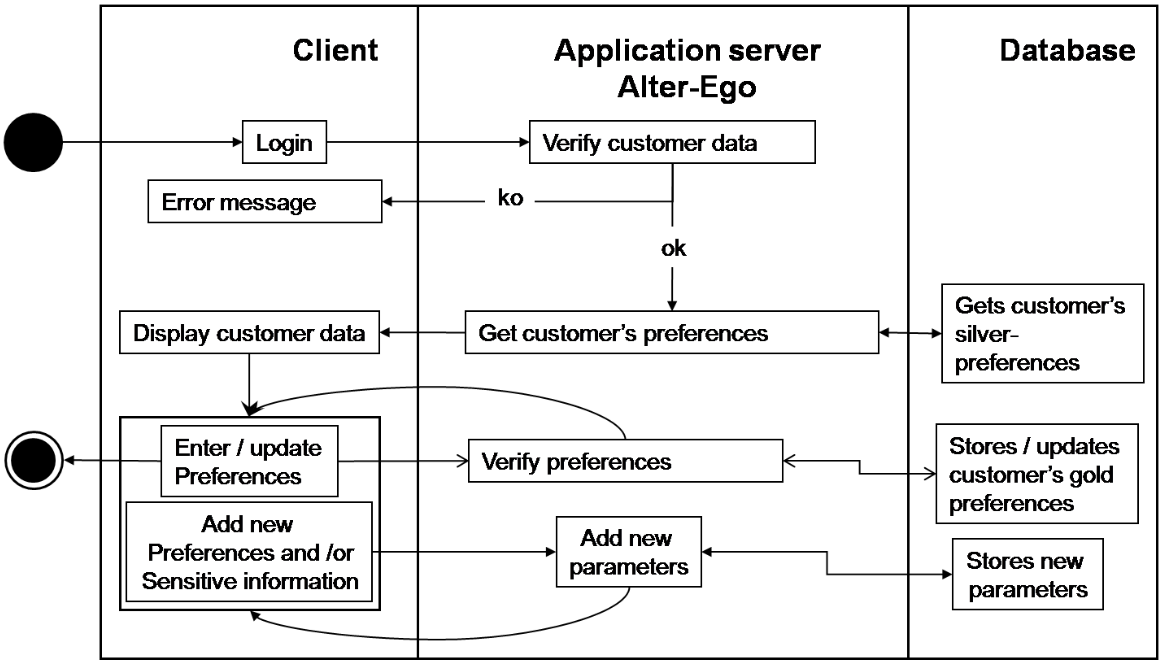
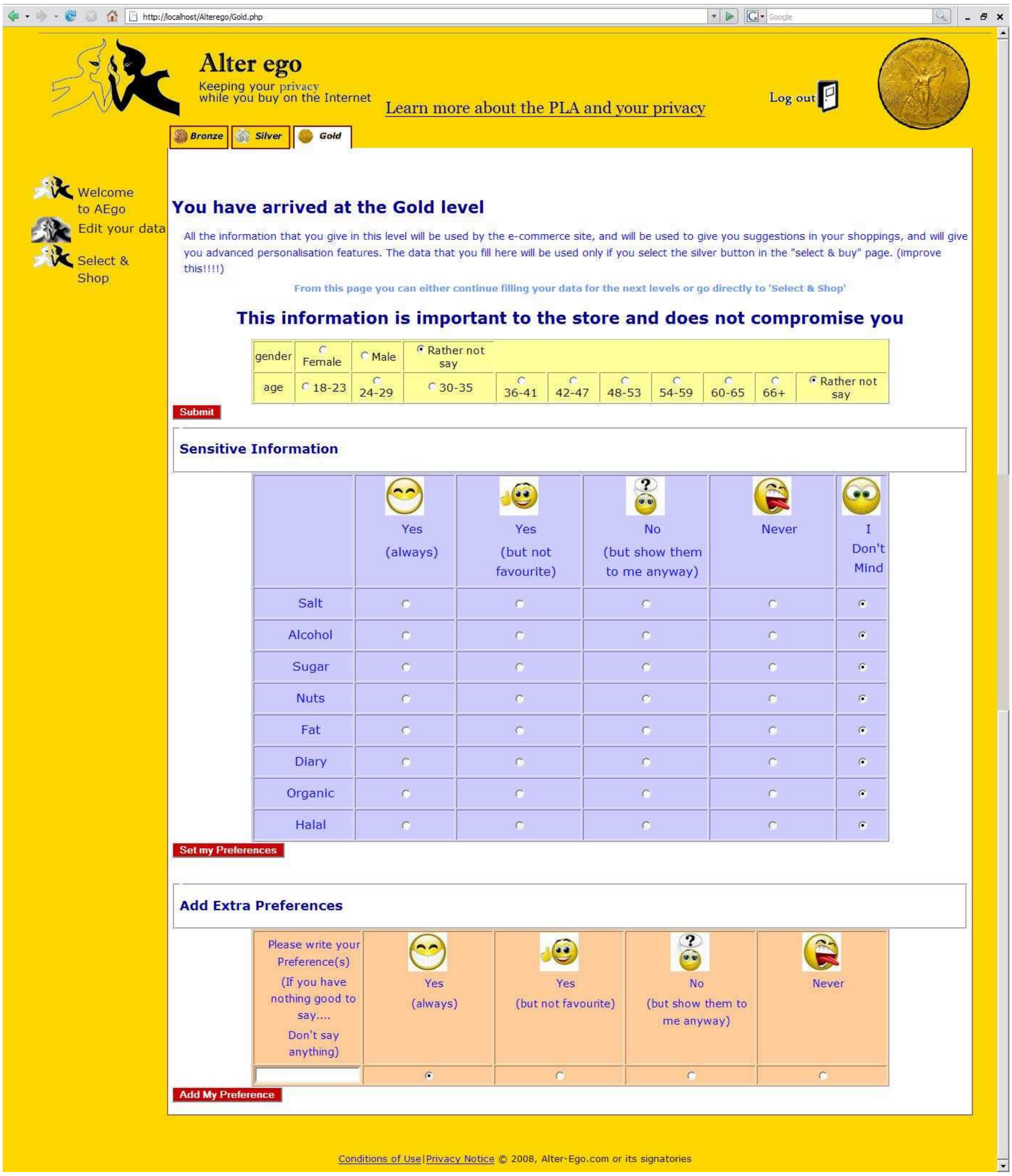

Client
Application server Alter-Ego
Database
Login
Verify customer data
Error message
ko
ok
Display customer data
Get customer's preferences
Gets customer's silver-preferences
Enter / update Preferences
Verify preferences
Stores / updates customer's gold preferences
Add new Preferences and /or Sensitive information
Add new parameters
Stores new parameters

Figure 6.23: Alter-Ego Web portal - Gold level - WebML activity diagram.

The information sent by gold level to the participant stores includes preferences (collected in silver level), sensitive information (collected in gold level), customer information such as demographics, and the new specifications introduced by customers (for instance *chocolate*). This information assists participant stores in their marketing studies.

Hence, in the first section of the gold disclosure level, demographic information is collected, followed by sensitive information, and finally, customers are allowed to add new categories to their preferences. In the implementation, three different colours are used in the tables to indicate the collection difference. At the same time, banners reinforce the identification of the requirements (for colour-blind customers). The Alter-Ego portal gold implementation is shown in Figure 6.24.



http://localhost/Alterego/Gold.php
Alter ego
Keeping your privacy
while you buy on the Internet
Learn more about the PLA and your privacy
Log out
Bronze
Silver
Gold
Welcome to AEgo
Edit your data
Select & Shop

You have arrived at the Gold level

All the information that you give in this level will be used by the e-commerce site, and will be used to give you suggestions in your shoppings, and will give you advanced personalisation features. The data that you fill here will be used only if you select the silver button in the "select & buy" page. (improve this!!!!)

From this page you can either continue filling your data for the next levels or go directly to 'Select & Shop'

This information is important to the store and does not compromise you

gender	○ Female	○ Male	● Rather not say							
age	○ 18-23	○ 24-29	○ 30-35	○ 36-41	○ 42-47	○ 48-53	○ 54-59	○ 60-65	○ 66+	● Rather not say

Submit

Sensitive Information

	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never	I Don't Mind
Salt	○	○	○	○	●
Alcohol	○	○	○	○	●
Sugar	○	○	○	○	●
Nuts	○	○	○	○	●
Fat	○	○	○	○	●
Diary	○	○	○	○	●
Organic	○	○	○	○	●
Halal	○	○	○	○	●

Set my Preferences

Add Extra Preferences

Please write your Preference(s) (If you have nothing good to say.... Don't say anything)	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never
	●	○	○	○

Add My Preference

Conditions of Use | Privacy Notice © 2008, Alter-Ego.com or its signatories

Figure 6.24: Alter-Ego Web portal - Gold level - Implementation.

6.4.7 Select & shop

Finally, the *select & shop* section of the Alter-Ego Web portal is where the customer can decide what disclosure level will be used for shopping with each store. In this same section, customers can give a ranking to the participant store or can be directed to a more detailed feedback section to express their ideas. In the *select & shop* section, customers are presented with the participant stores, so they can choose from the list. The layout and transition diagram are shown in Figure 6.25. In this diagram, the service area is where the customer can add new stores from the collection of participant stores.

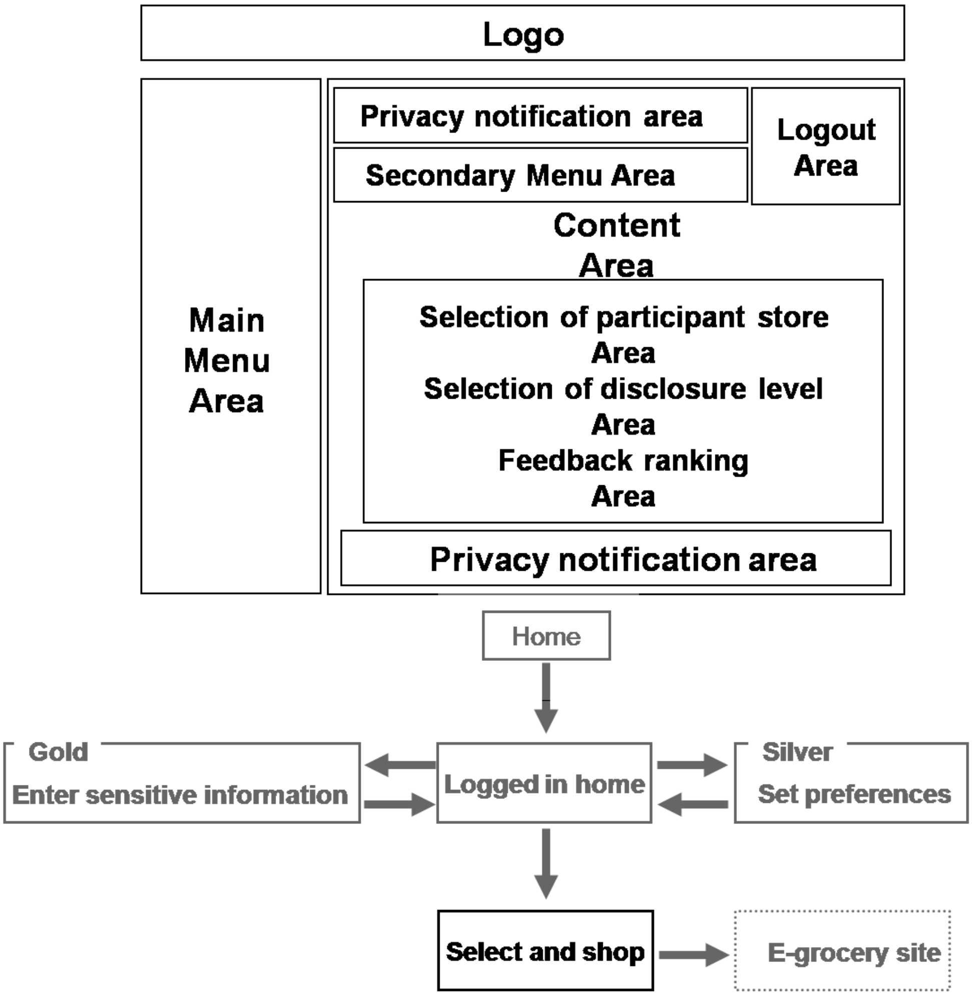


Figure 6.25: Alter-Ego Web portal - Select & shop - Layout and transition diagram.

After selecting the store, the customer decides what disclosure level will be used with that specific store. If no disclosure level is chosen, Alter-Ego Web portal sets *bronze* as the default level. The process of adding a new entry to the customer's set of stores is shown in Figure 6.26.

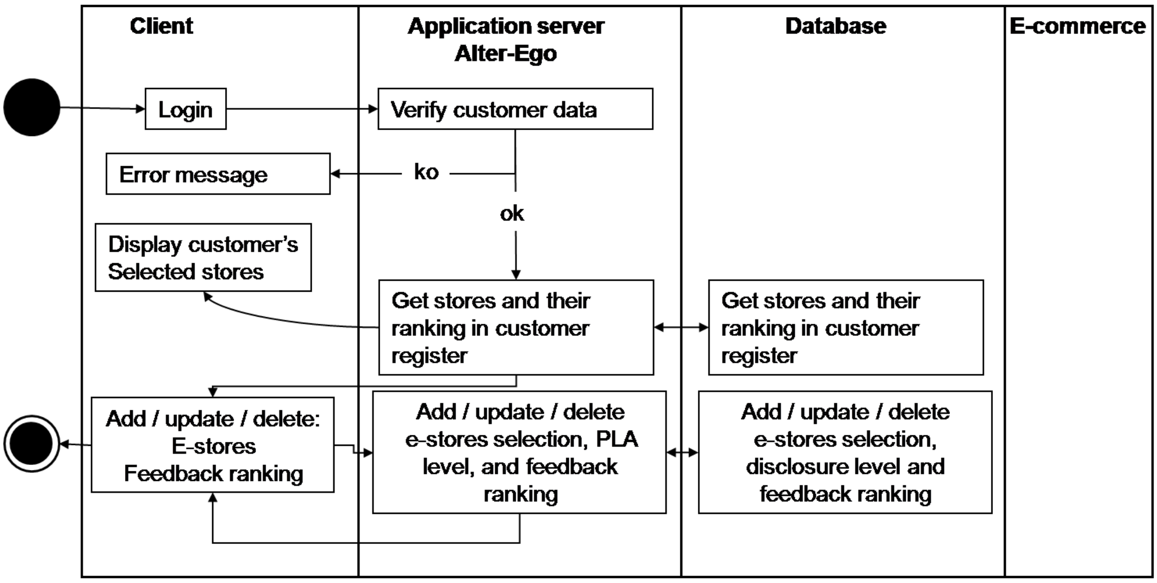

Client
Application server Alter-Ego
Database
E-commerce
Login
Verify customer data
Error message
ko
ok
Display customer's Selected stores
Get stores and their ranking in customer register
Get stores and their ranking in customer register
Add / update / delete: E-stores Feedback ranking
Add / update / delete e-stores selection, PLA level, and feedback ranking
Add / update / delete e-stores selection, disclosure level and feedback ranking

Figure 6.26: Alter-Ego Web portal - Select & shop. WebML activity diagram - Addition of new participant stores.

Once the participant store and disclosure level have been selected, the customer selects a link to be directed to that selected store. The Alter-Ego Web portal generates a random identifier, that, together with the disclosure level, will be used by the store to retrieve the customer's information using a Web service. The process of sending the information to the store is shown in Figure 6.27.

In the case of the bronze disclosure level, no information is retrieved. Preferences are retrieved for silver and both preferences and sensitive information are retrieved in the case of gold. The Web service validates the request to ensure that the store is only provided with the authorised information.

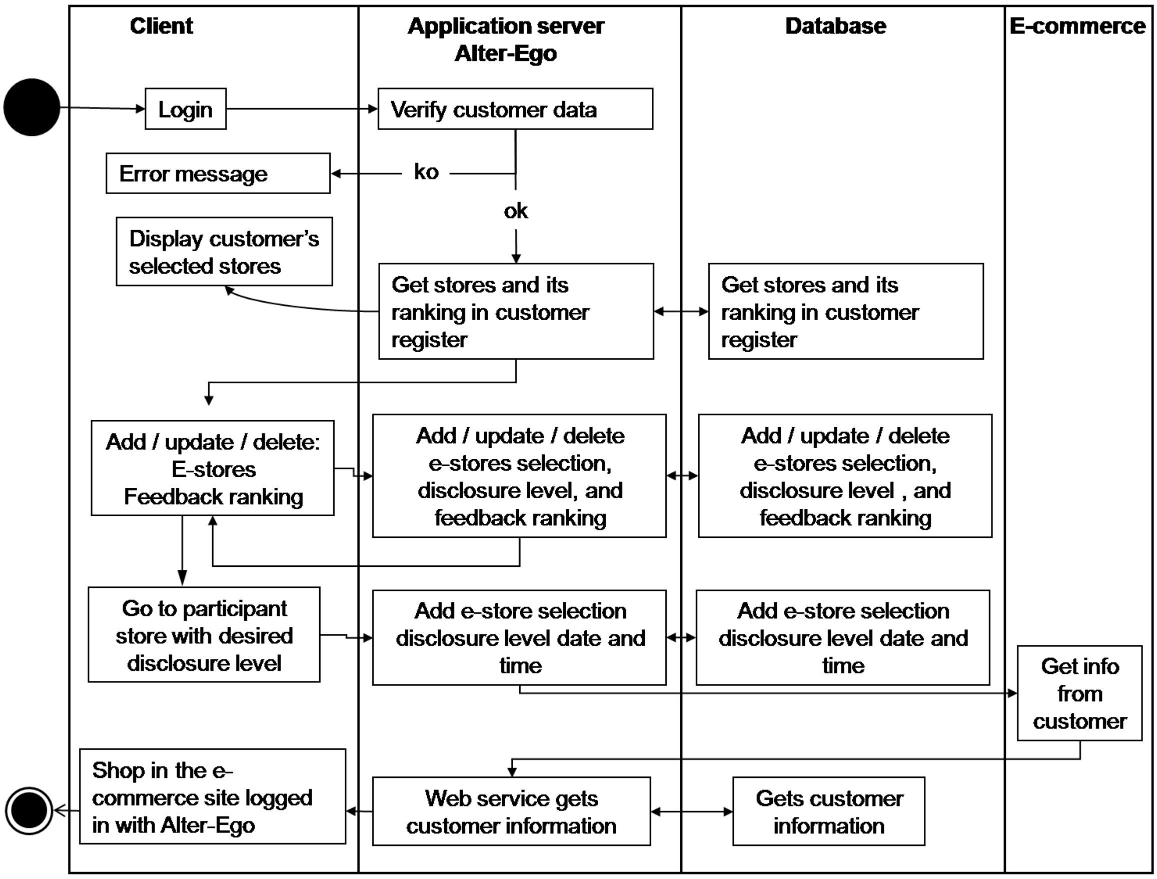
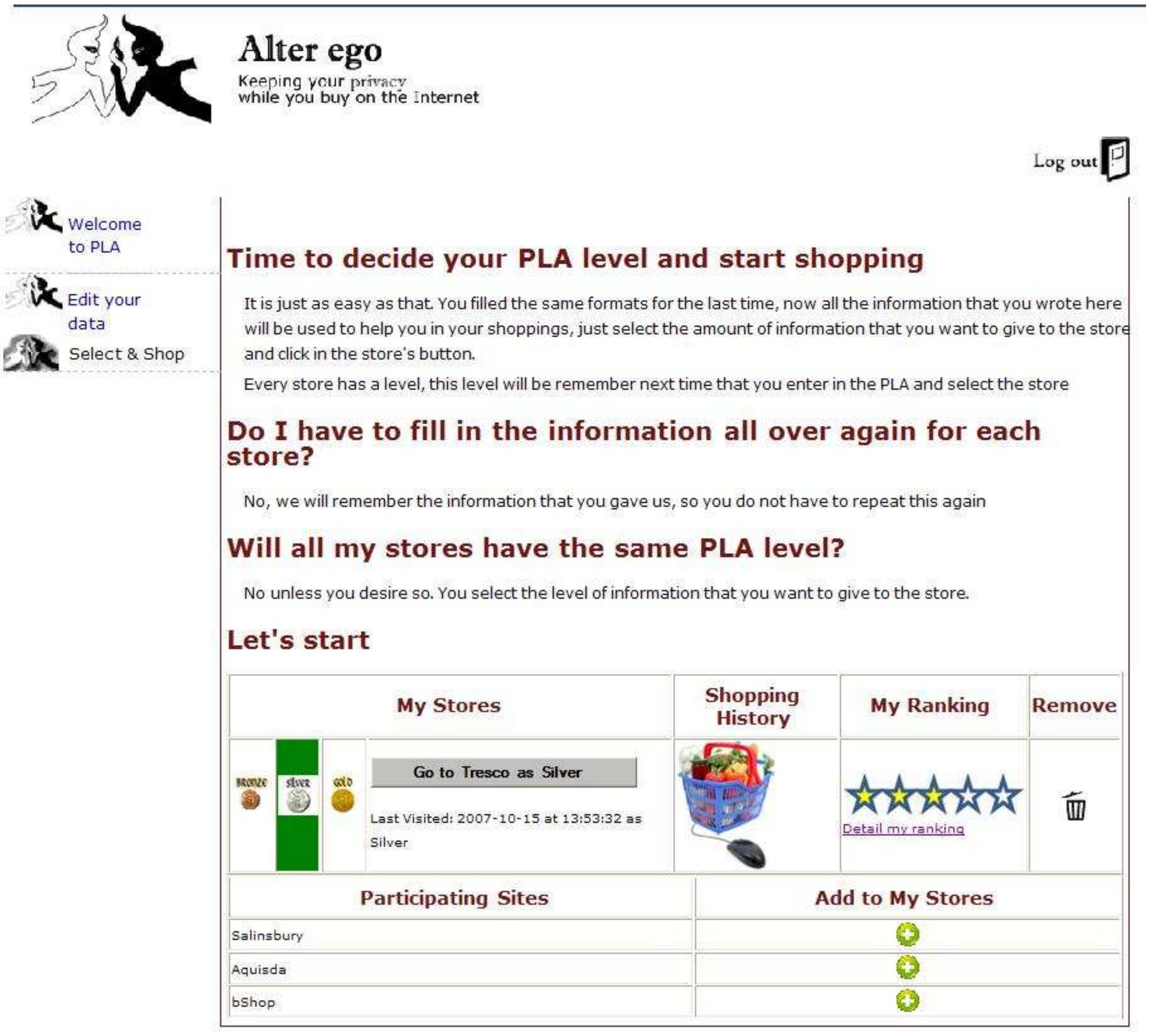

Client
Application server Alter-Ego
Database
E-commerce
Login
Verify customer data
Error message
ko
ok
Display customer's selected stores
Get stores and its ranking in customer register
Get stores and its ranking in customer register
Add / update / delete: E-stores Feedback ranking
Add / update / delete e-stores selection, disclosure level, and feedback ranking
Add / update / delete e-stores selection, disclosure level , and feedback ranking
Go to participant store with desired disclosure level
Add e-store selection disclosure level date and time
Add e-store selection disclosure level date and time
Get info from customer
Shop in the e-commerce site logged in with Alter-Ego
Web service gets customer information
Gets customer information

Figure 6.27: Alter-Ego Web portal - Select & shop. WebML activity diagram - Sending information to store.

The implementation of select & shop, shown in Figure 6.28, includes a list of participant stores to choose from, by clicking an add icon. The selected stores appear with a visual aid on their left hand side, to indicate the disclosure level to be used. To the right hand side, a series of icons (stars) assist visually the customers ranking of the store. The selection of the disclosure level is done by clicking on the icon of the corresponding medal in the left hand side of the selected store's name. Customers are also shown the disclosure level, date and time of their last visit to that particular store, and if they click on the "*Shopping History*" icon, they are presented with a list containing details of their previous visits to that particular participant store. Finally, customers can remove the participant store by clicking on the icon under the "*Remove*" tag.



Time to decide your PLA level and start shopping

It is just as easy as that. You filled the same formats for the last time, now all the information that you wrote here will be used to help you in your shoppings, just select the amount of information that you want to give to the store and click in the store's button.

Every store has a level, this level will be remember next time that you enter in the PLA and select the store

Do I have to fill in the information all over again for each store?

No, we will remember the information that you gave us, so you do not have to repeat this again

Will all my stores have the same PLA level?

No unless you desire so. You select the level of information that you want to give to the store.

Let's start

My Stores	Shopping History	My Ranking	Remove
BRONZE SILVER GOLD Go to Tresco as Silver Last Visited: 2007-10-15 at 13:53:32 as Silver		★★★☆☆ Detail my ranking	

Participating Sites	Add to My Stores
Salinsbury	⊕
Aquisda	⊕
bShop	⊕

Figure 6.28: Alter-Ego Web portal - Select & shop - Implementation.

After the participant store and disclosure level selection is done, the Alter-Ego Web portal sends information to the selected store. With that information, the store uses a Web service to retrieve the customers' preferences or sensitive information. The Web service is presented next.

6.5 Web service

As introduced in Section 6.3.3, a Web service is used to enable the participant store to retrieve the customer's information from the Alter-Ego Web portal. Section 6.4.7, detailed the information to be sent by the Alter-Ego Web portal to the customer's selected participant store, consisting of: a random identifier number and the customer's selected disclosure level. With that information (both data are required), the Web service (named *the dispenser*) retrieves the customer's pertinent information from the Alter-Ego's database, this process is shown in Figure 6.29.

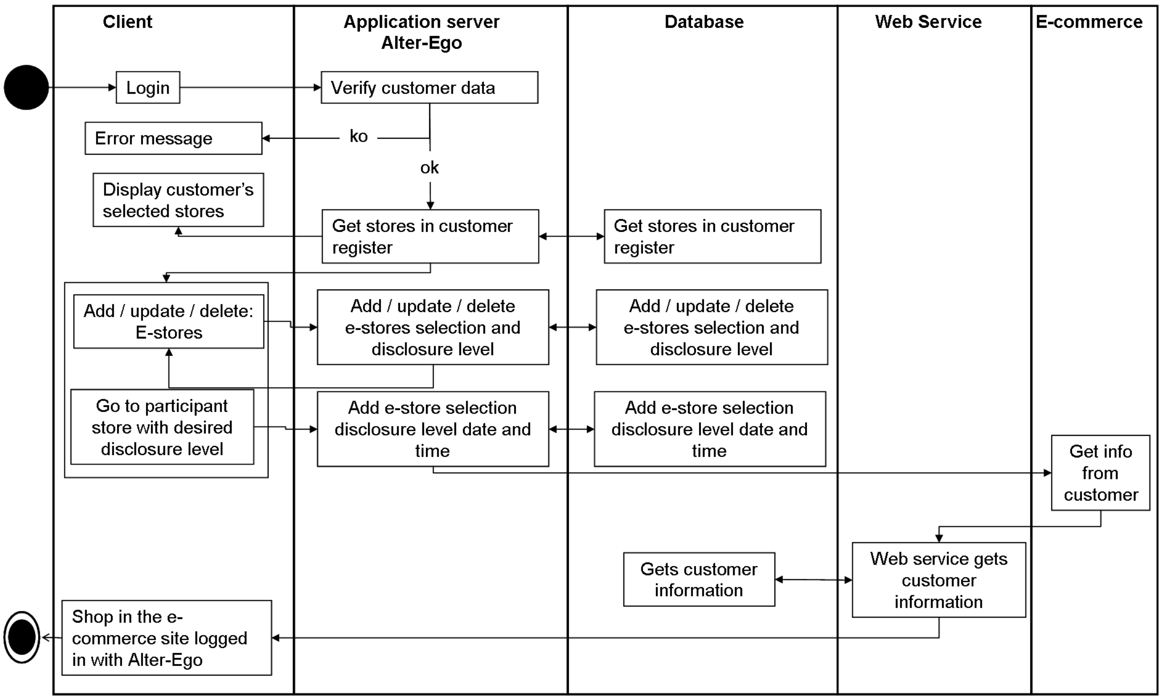

Client
Application server Alter-Ego
Database
Web Service
E-commerce
Login
Verify customer data
Error message
ko
ok
Display customer's selected stores
Get stores in customer register
Get stores in customer register
Add / update / delete: E-stores
Add / update / delete e-stores selection and disclosure level
Add / update / delete e-stores selection and disclosure level
Go to participant store with desired disclosure level
Add e-store selection disclosure level date and time
Add e-store selection disclosure level date and time
Get info from customer
Gets customer information
Web service gets customer information
Shop in the e-commerce site logged in with Alter-Ego

Figure 6.29: WebML activity diagram - Web service of the Alter-Ego Web portal sending information to store.

The Web service (*the dispenser*) has, as shown in Figure 6.30, the following functionalities:

1. The information used by the e-commerce store is verified to check that it does indeed correspond to the customer. This is done by a query to the database that stores customers random numbers and their disclosure levels.

2. If the customer verification is correct, the information corresponding to the customer's disclosure level (preferences, sensitive, or both) is retrieved.

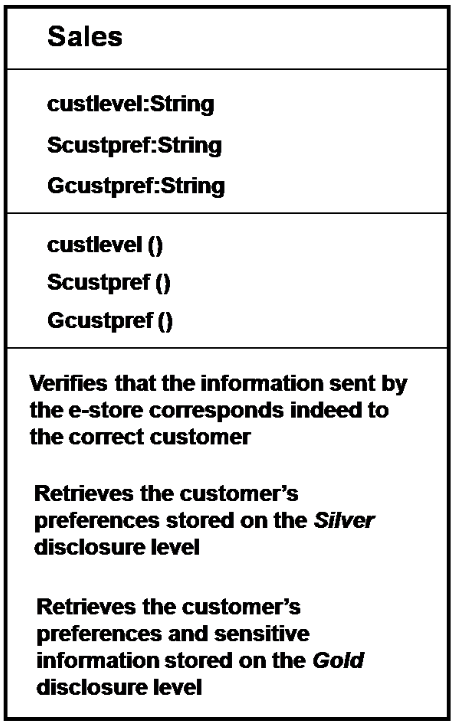

Sales
custlevel:String
Scustpref:String
Gcustpref:String
custlevel ()
Scustpref ()
Gcustpref ()
Verifies that the information sent by the e-store corresponds indeed to the correct customer
Retrieves the customer's preferences stored on the *Silver* disclosure level
Retrieves the customer's preferences and sensitive information stored on the *Gold* disclosure level

Figure 6.30: Web service (*the dispenser*) UML diagram.

The PPSE prototype involves an e-grocery participant store named *bshop*, its design and implementation are presented next.

6.6 E-grocery store - Bshop

As stated in Section 2.2.7, one area of e-commerce that has not experienced quick customer uptake is e-groceries. Furthermore, the data inferred from logging of activities during shopping for groceries could lead to privacy violations. Therefore, an e-grocery store was designed, implemented and developed to complete the environment required for proving the PPSE approach and testing its acceptance. The original e-grocery store, bshop, was developed as a Department of Computing Science third year final students' project. It was then modified to make it comply with the PLA, and become PPSE compatible.

The bshop was adapted to work in two different modalities: *stand-alone* and *Alter-Ego Web portal authenticated.* The stand-alone mode allows the customer to browse, select and shop for items using the bshop privacy policy. The transition diagram, presented in Figure 6.31, shows how customers authenticate themselves according to the bshop registration (number 1 in the diagram), and are provided with the store's personalisation features (number 2 in the diagram) and purchases. After customers fill the shopping basket with their selections (number 3 in diagram), they proceed to checkout (number 4 in the diagram). Since it was an experimental model, the checkout only presents a list of the purchased items, the total amount of the purchase and a single notice informing the customer that the transaction has been successful. No credit card details are requested due to ethical concerns. The implementation of the stand-alone functionality of the bshop is shown in Figure 6.32.

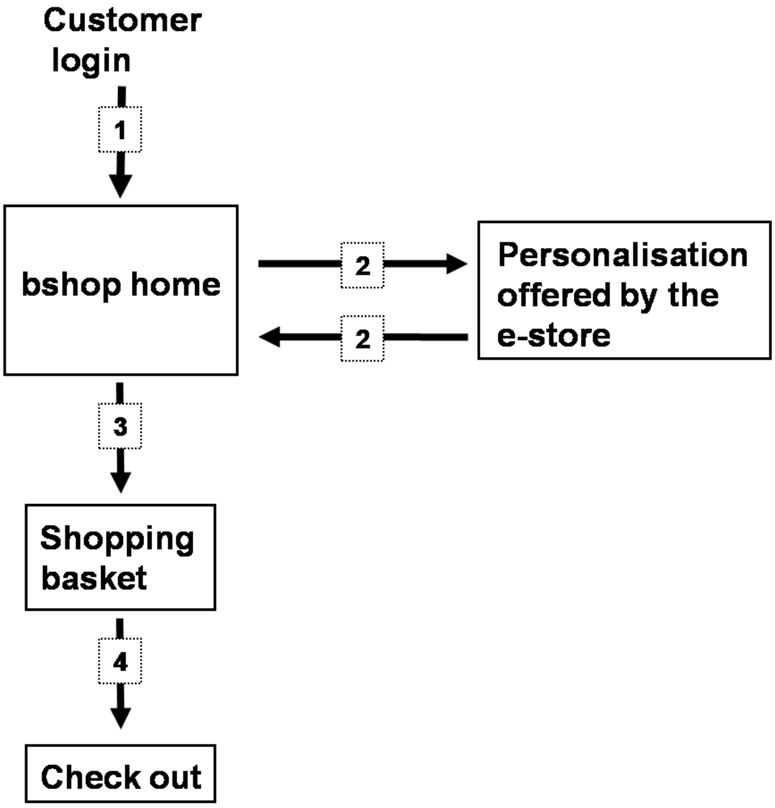
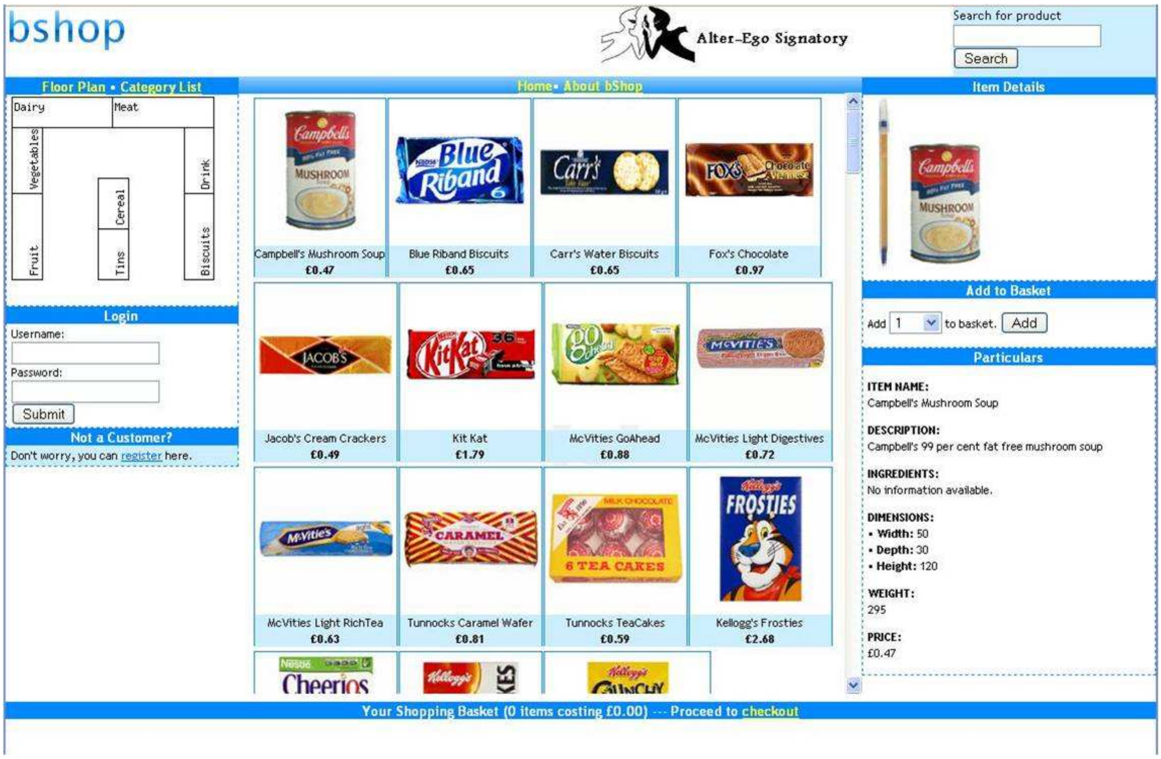


Figure 6.31: Transition diagram of bshop as stand alone.



Don't worry, you can register here.

ITEM NAME:
Campbell's Mushroom Soup

DESCRIPTION:
Campbell's 99 per cent fat free mushroom soup

INGREDIENTS:
No information available.

DIMENSIONS:
- Width: 50
- Depth: 30
- Height: 120

WEIGHT:
295

PRICE:
£0.47

Your Shopping Basket (0 items costing £0.00) --- Proceed to checkout

Figure 6.32: Bshop as stand alone - Implementation.

The Alter-Ego Web portal authenticated modality of the bshop, shown in Figure 6.33, presents the interaction between Alter-Ego Web portal and the bshop. As shown in the diagram, the information that is provided by the Alter-Ego portal (number 1) is used to retrieve the customer's preferences and sensitive information from an associated Web service (numbers 2,3,4 and 5 in the diagram). The personalisation and checkout characteristics of the bshop vary according to the requested disclosure level (number 6 in the diagram). After the goods are chosen and introduced in the shopping basket (number 7 in the diagram), the check out process and the information stored by the e-grocery store correspond to the disclosure levels (number 8 in the diagram).

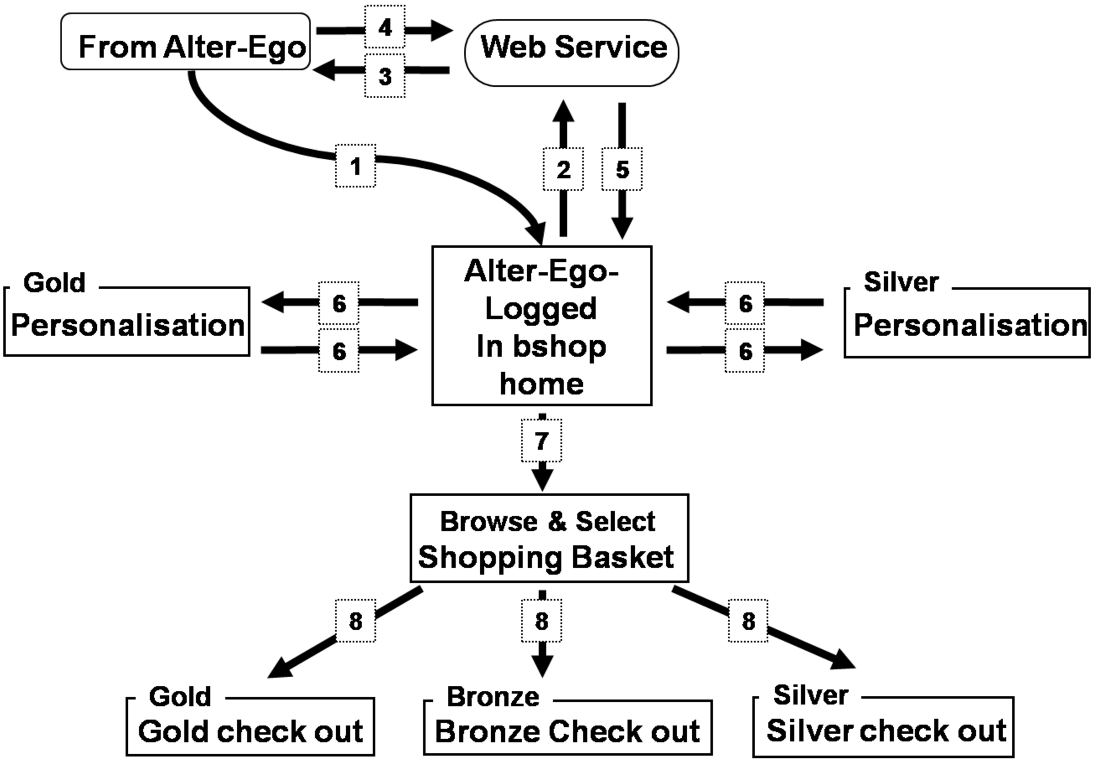


Figure 6.33: Transition diagram of Alter-Ego Web portal and bshop - Interaction.

Hence, when the customer is directed to the bshop from the Alter-Ego portal, it behaves according to the disclosure level used to access it. The different behaviours according to the disclosed level are presented next.

6.6.1 Bshop and Alter-Ego Web portal - *Bronze*

Alter-Ego's default level is bronze. However the customer can also select bronze as the disclosure level. The bronze level anonymises the customer's information and discloses no information to the store. Since the store obtains no benefit from this customer's information, no personalisation is provided. The layout of the bshop at bronze level is shown in Figure 6.34 and the implementation is shown in Figure 6.35.

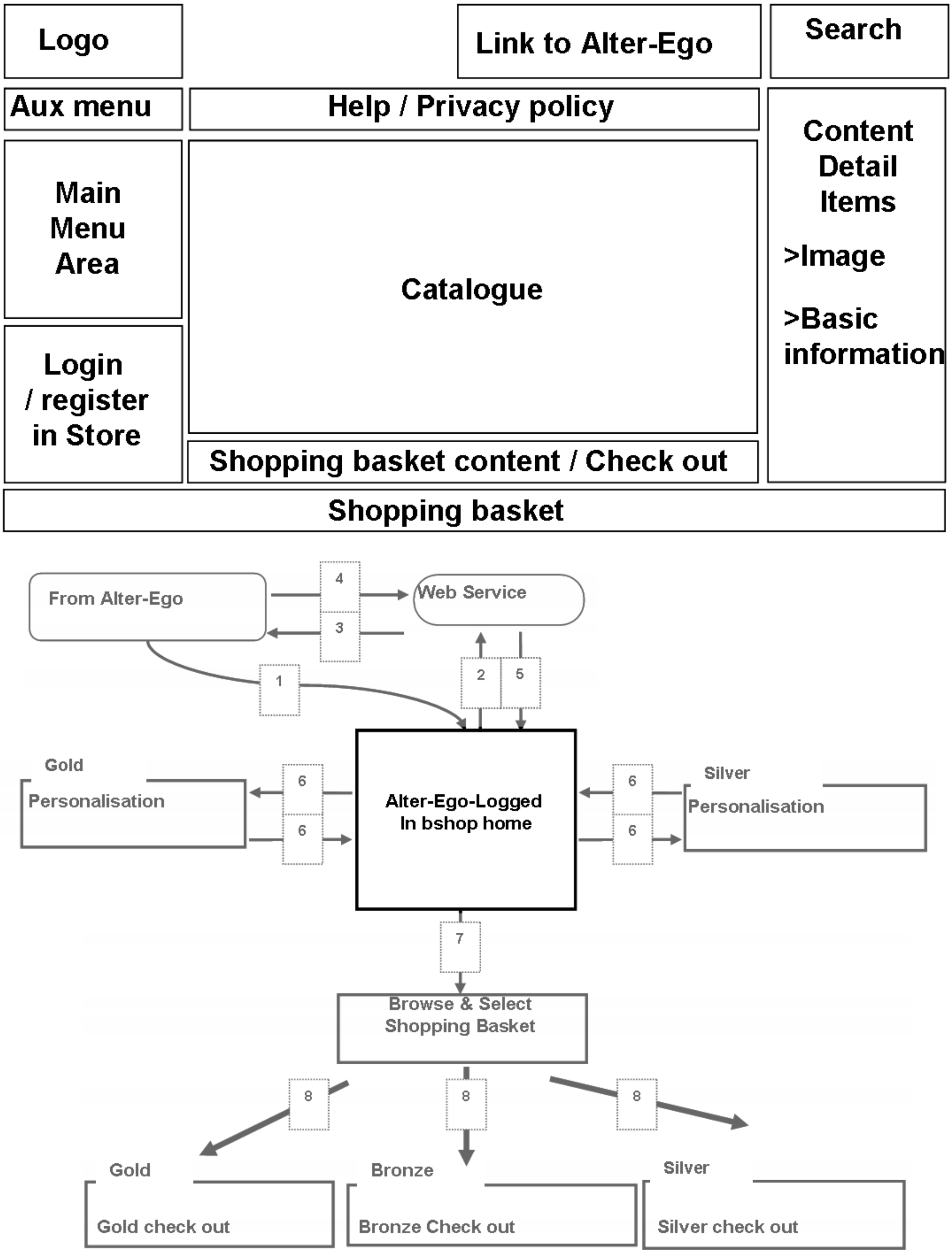
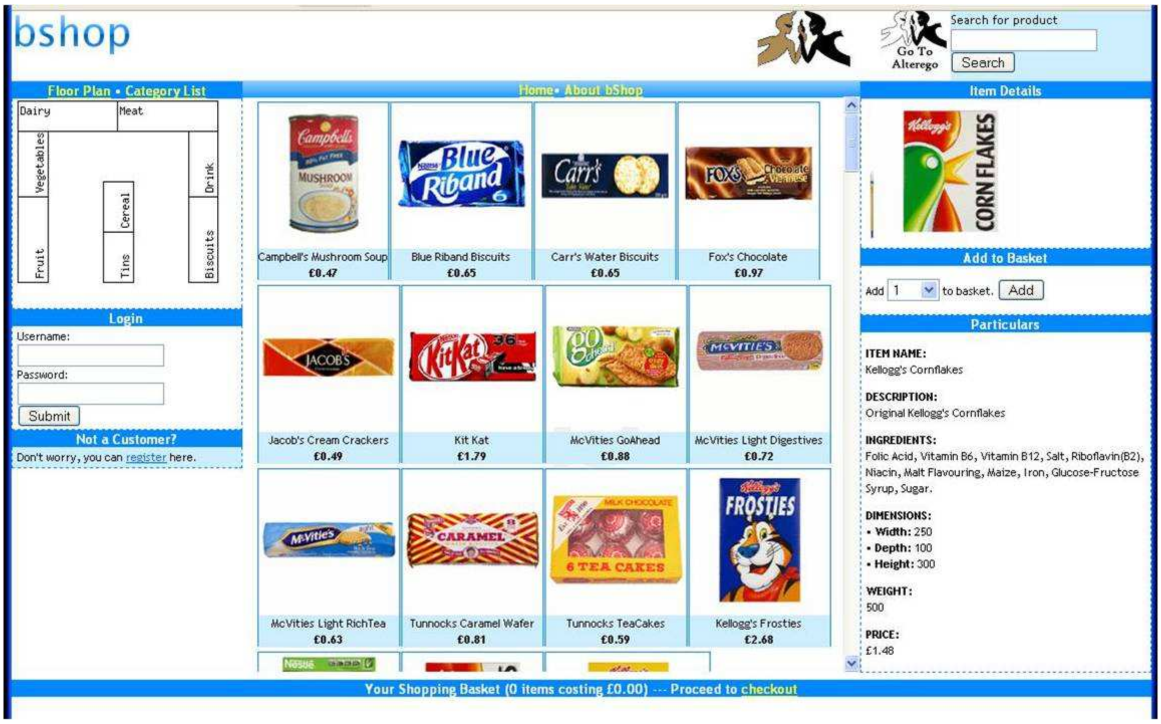


Figure 6.34: Bshop logged in Alter-Ego bronze level - Layout and transition diagram.

In the implementation, to indicate that the customer was directed from the Alter-Ego

Web portal, two Alter-Ego icons appear in the store's top right hand corner. The icon on the left hand side has a bronze colour and if the customer moves the cursor over the icon, a banner indicating "Bronze level" appears. The icon on the right opens, in a new window, the Alter-Ego Web portal. However, the changes implemented in the Alter-Ego Web portal opened in the new window, will take effect in a different (new) shopping session. To alter the settings in the current shopping session, changes can be executed, for silver and gold levels only, by using the window that the (silver or gold) icon at the left hand side opens.


bshop
Go To Alterego
Search for product
Search
Floor Plan • Category List
Dairy
Meat
Vegetables
Drink
Cereal
Fruit
Tins
Biscuits
Login
Username:
Password:
Submit
Not a Customer?
Don't worry, you can register here.
Home • About bShop
Campbell's Mushroom Soup £0.47
Blue Riband Biscuits £0.65
Carr's Water Biscuits £0.65
Fox's Chocolate £0.97
Jacob's Cream Crackers £0.49
Kit Kat £1.79
McVities GoAhead £0.88
McVities Light Digestives £0.72
McVities Light RichTea £0.63
Tunnocks Caramel Wafer £0.81
Tunnocks TeaCakes £0.59
Kellogg's Frosties £2.68
Item Details
Add to Basket
Add 1 to basket. Add
Particulars
Your Shopping Basket (0 items costing £0.00) --- Proceed to checkout

ITEM NAME:
Kellogg's Cornflakes

DESCRIPTION:
Original Kellogg's Cornflakes

INGREDIENTS:
Folic Acid, Vitamin B6, Vitamin B12, Salt, Riboflavin(B2), Niacin, Malt Flavouring, Maize, Iron, Glucose-Fructose Syrup, Sugar.

DIMENSIONS:
- **Width:** 250
- **Depth:** 100
- **Height:** 300

WEIGHT:
500

PRICE:
£1.48

Figure 6.35: Bshop logged in Alter-Ego bronze level - Implementation.

During the browsing and selection of goods, the bshop operates as the normal store does outside the PPSE. After the checkout, the customer is presented with a privacy notice stating that no information was stored or linked to a personal record. The layout and transition diagram of the bronze checkout are shown in Figure 6.36, and the implementation is shown in Figure 6.37.

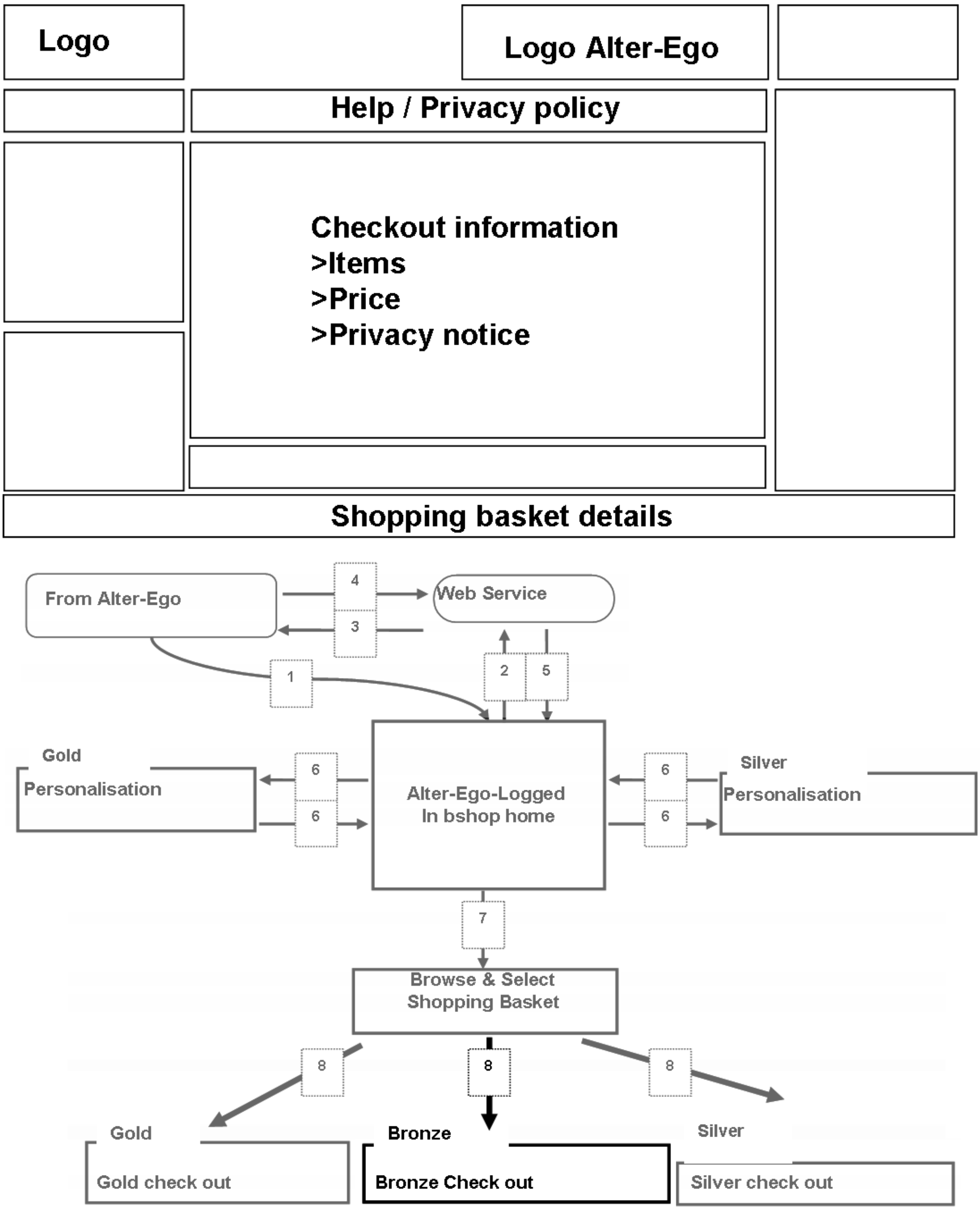
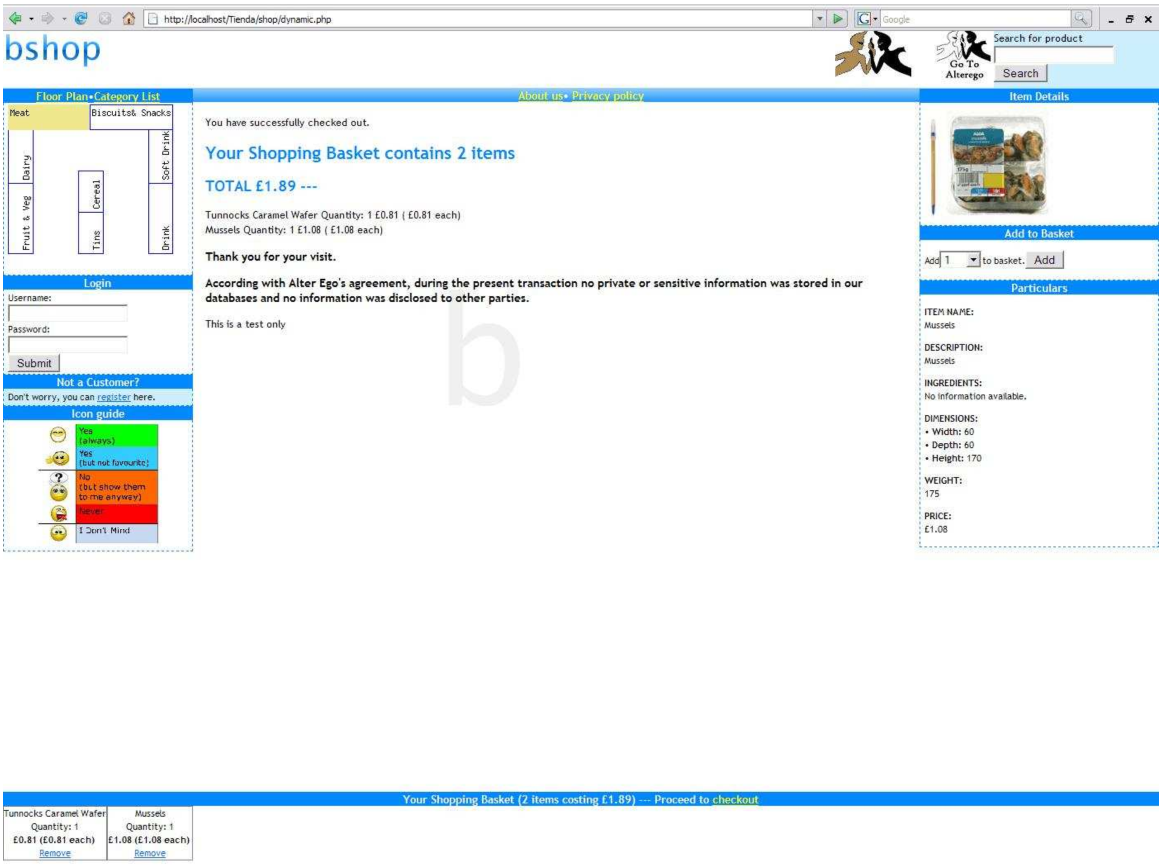


Figure 6.36: Bshop checkout when logged in Alter-Ego bronze level - Layout and transition diagram.



Don't worry, you can register here.

You have successfully checked out.

Your Shopping Basket contains 2 items

TOTAL £1.89 ---

Tunnocks Caramel Wafer Quantity: 1 £0.81 (£0.81 each)
Mussels Quantity: 1 £1.08 (£1.08 each)

Thank you for your visit.

According with Alter Ego's agreement, during the present transaction no private or sensitive information was stored in our databases and no information was disclosed to other parties.

This is a test only

ITEM NAME:
Mussels

DESCRIPTION:
Mussels

INGREDIENTS:
No information available.

DIMENSIONS:
- Width: 60
- Depth: 60
- Height: 170

WEIGHT:
175

PRICE:
£1.08

Your Shopping Basket (2 items costing £1.89) --- Proceed to checkout

Tunnocks Caramel Wafer Quantity: 1 £0.81 (£0.81 each) Remove	Mussels Quantity: 1 £1.08 (£1.08 each) Remove

Figure 6.37: Bshop checkout when logged in Alter-Ego bronze level - Implementation.

6.6.2 Bshop and Alter-Ego Web portal - *Silver*

When the customer selects the silver level, the Web service (named the *dispenser*) retrieves the customer preferences and therefore the store can use those customer preferences to offer customisation of the shopping experience. Figure 6.38 shows the layout of the bshop when the customer chooses the silver level. In the central part, the catalogue presented to the customer is personalised, based on the customer's preferences (provided by the customer to the Alter-Ego Web portal and collected by the store using the *dispenser*).

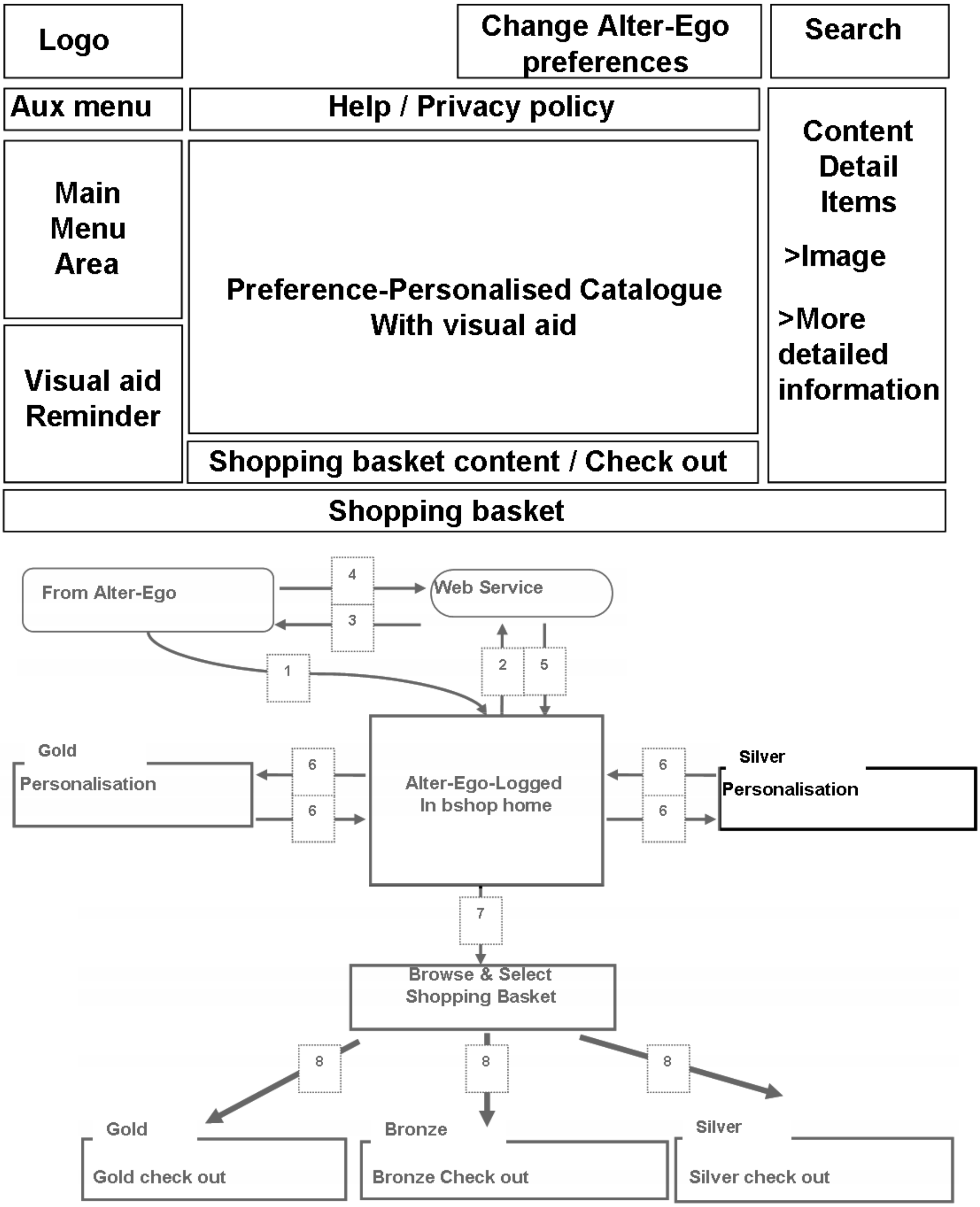
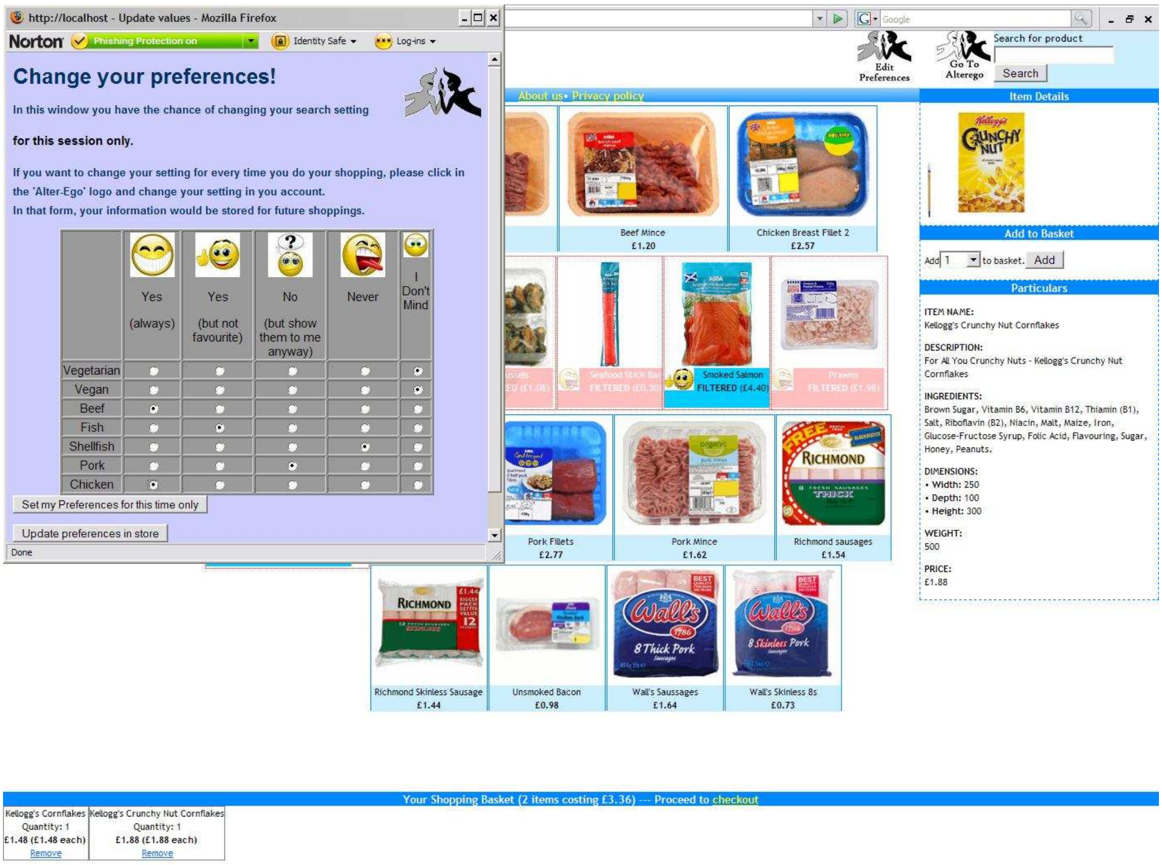


Figure 6.38: Bshop logged in Alter-Ego silver level - Layout and transition diagram.

The store rewards the customer's disclosure of information. In this case, extra information and customisation are offered. To exemplify this, the right hand side shows the items with more detail than the bronze level. At the same time, in the bottom left part of the window, a set of auxiliary icons ("Visual aid reminder" in Figure 6.38) act as a visual guide to the customer in recalling their specified preferences (implementation is shown in Figure 6.37). The customer can change the preferences by using the auxiliary window opened with the silver Alter-Ego icon situated in the top right of the window. This option, named *"Change Alter-Ego preferences"*, allows the customer to modify the preferences only for the current shopping session. The implementation of this feature is shown in Figure 6.39. Future versions can include an option to allow the customers to store the changes made during the shopping session into their Alter-Ego Web portal account.



http://localhost - Update values - Mozilla Firefox

Change your preferences!

In this window you have the chance of changing your search setting

for this session only.

If you want to change your setting for every time you do your shopping, please click in the 'Alter-Ego' logo and change your setting in you account.
In that form, your information would be stored for future shoppings.

	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never	I Don't Mind
Vegetarian	○	○	○	○	●
Vegan	○	○	○	○	●
Beef	●	○	○	○	○
Fish	○	●	○	○	○
Shellfish	○	○	○	●	○
Pork	○	○	●	○	○
Chicken	●	○	○	○	○

Set my Preferences for this time only

Update preferences in store

Beef Mince £1.20

Chicken Breast Fillet 2 £2.57

ussels ED (£1.08)

Seafood Stick Bar FILTERED (£0.30)

Smoked Salmon FILTERED (£4.40)

Prawns FILTERED (£1.98)

Pork Fillets £2.77

Pork Mince £1.62

Richmond sausages £1.54

Richmond Skinless Sausage £1.44

Unsmoked Bacon £0.98

Wall's Saussages £1.64

Wall's Skinless 8s £0.73

Item Details

Add to Basket

Add 1 to basket. Add

Particulars

ITEM NAME:
Kellogg's Crunchy Nut Cornflakes

DESCRIPTION:
For All You Crunchy Nuts - Kellogg's Crunchy Nut Cornflakes

INGREDIENTS:
Brown Sugar, Vitamin B6, Vitamin B12, Thiamin (B1), Salt, Riboflavin (B2), Niacin, Malt, Maize, Iron, Glucose-Fructose Syrup, Folic Acid, Flavouring, Sugar, Honey, Peanuts.

DIMENSIONS:
- Width: 250
- Depth: 100
- Height: 300

WEIGHT:
500

PRICE:
£1.88

Your Shopping Basket (2 items costing £3.36) --- Proceed to checkout

Kellogg's Cornflakes Quantity: 1 £1.48 (£1.48 each) Remove	Kellogg's Crunchy Nut Cornflakes Quantity: 1 £1.88 (£1.88 each) Remove

Figure 6.39: Bshop logged in Alter-Ego silver level - Change of preferences - Implementation.

The checkout for the bshop when the Alter-Ego Web portal is logged in as silver is shown in Figure 6.40, and the implementation is shown in Figure 6.41.

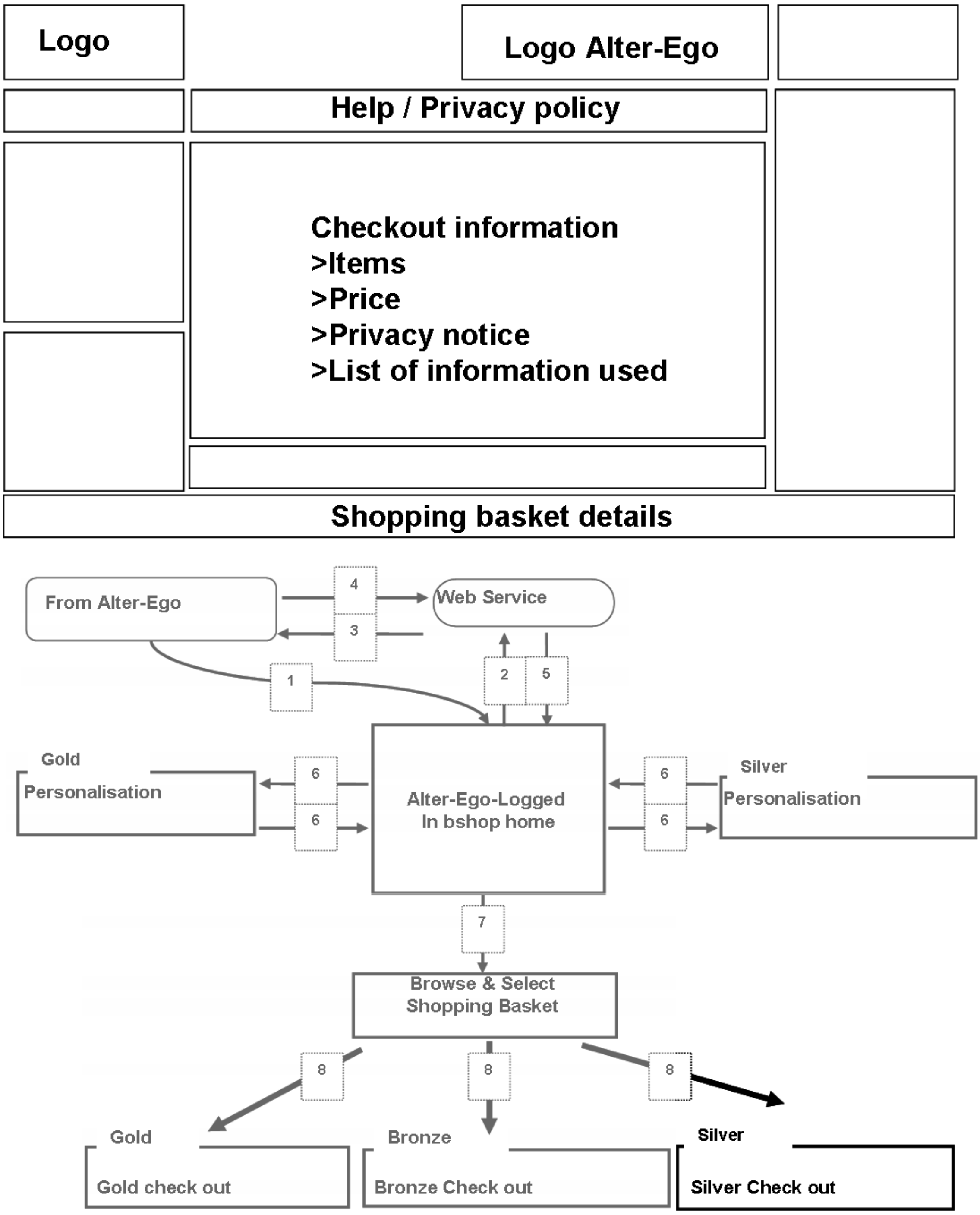
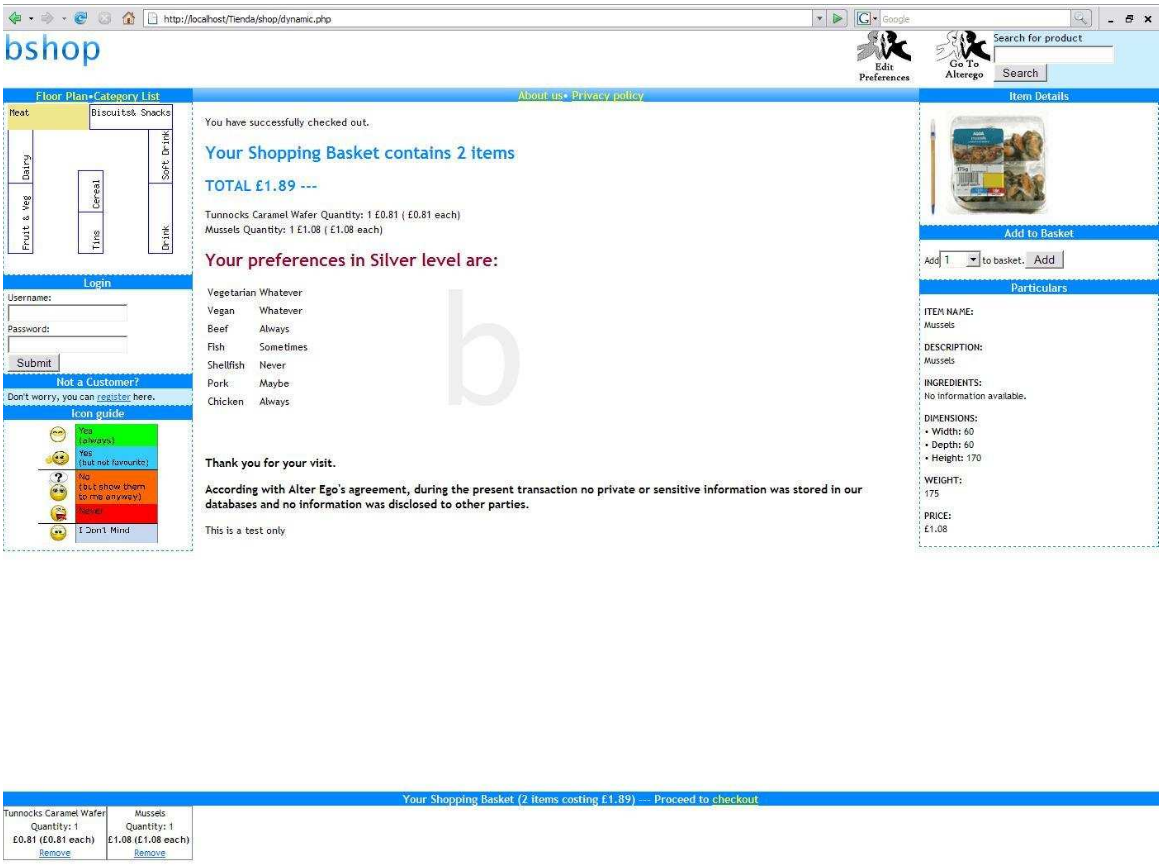


Figure 6.40: Bshop checkout when logged in as silver - Layout and transition diagram.



Don't worry, you can register here.

You have successfully checked out.

Your Shopping Basket contains 2 items

TOTAL £1.89 ---

Tunnocks Caramel Wafer Quantity: 1 £0.81 (£0.81 each)
Mussels Quantity: 1 £1.08 (£1.08 each)

Your preferences in Silver level are:

Vegetarian	Whatever
Vegan	Whatever
Beef	Always
Fish	Sometimes
Shellfish	Never
Pork	Maybe
Chicken	Always

Thank you for your visit.

According with Alter Ego's agreement, during the present transaction no private or sensitive information was stored in our databases and no information was disclosed to other parties.

This is a test only

ITEM NAME:
Mussels

DESCRIPTION:
Mussels

INGREDIENTS:
No information available.

DIMENSIONS:
- Width: 60
- Depth: 60
- Height: 170

WEIGHT:
175

PRICE:
£1.08

Your Shopping Basket (2 items costing £1.89) --- Proceed to checkout

Tunnocks Caramel Wafer	Mussels
Quantity: 1	Quantity: 1
£0.81 (£0.81 each)	£1.08 (£1.08 each)
Remove	Remove

Figure 6.41: Bshop checkout when logged in as silver - Implementation.

6.6.3 Bshop and Alter-Ego Web portal - *Gold*

Finally, when the Alter-Ego customer decides to use the gold level, the bshop is able to retrieve all the customer's preferences and sensitive information using the Web service (the *dispenser*). It is important to notice that customers using the gold disclosure level have the facility of managing both their preferences and sensitive information at the same time, unlike silver disclosure level that only allowed the managing of their preferences.

The participant store, the bshop, offers the full extent of facilities as a way of rewarding the customer's information disclosure. The layout and transition diagram are shown in Figure 6.42.

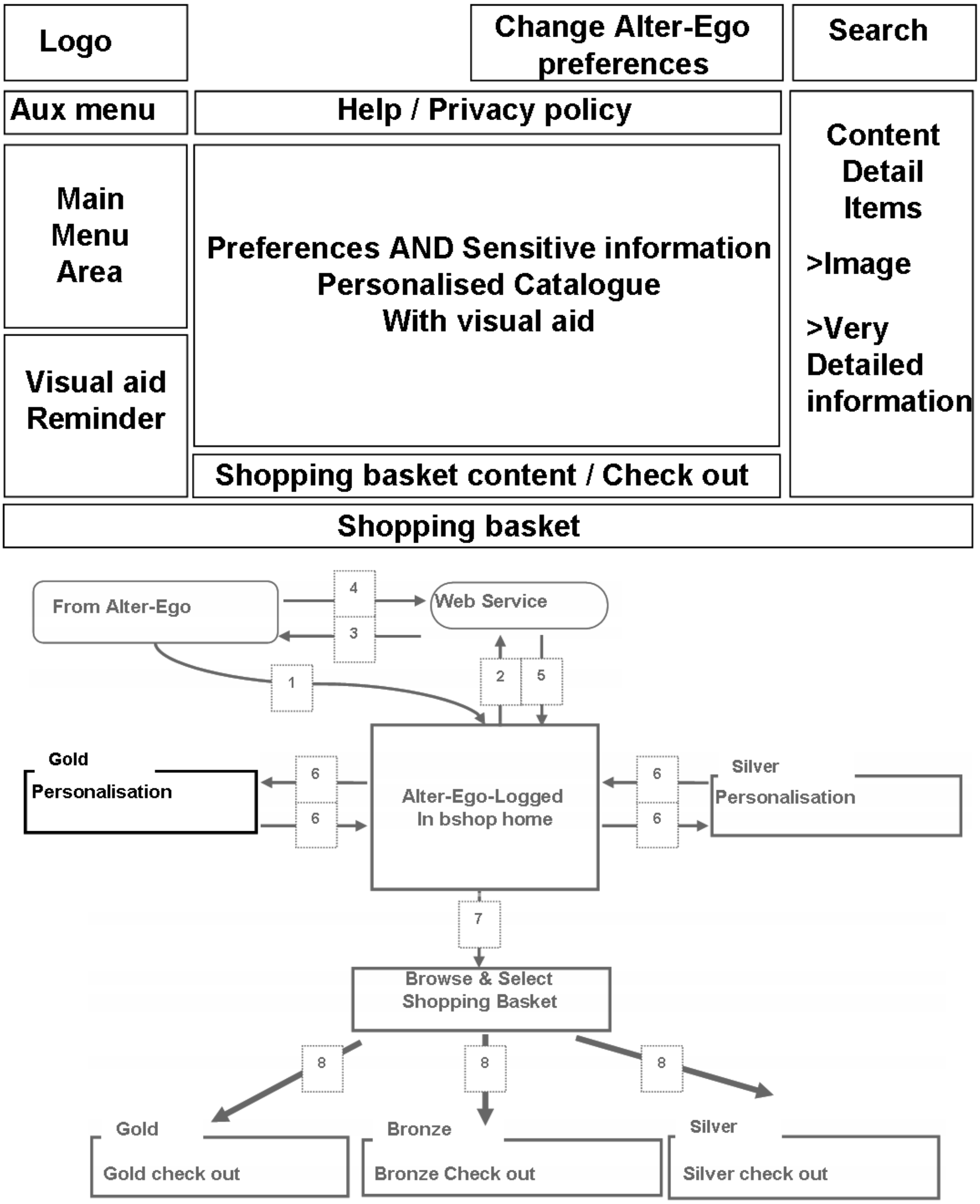
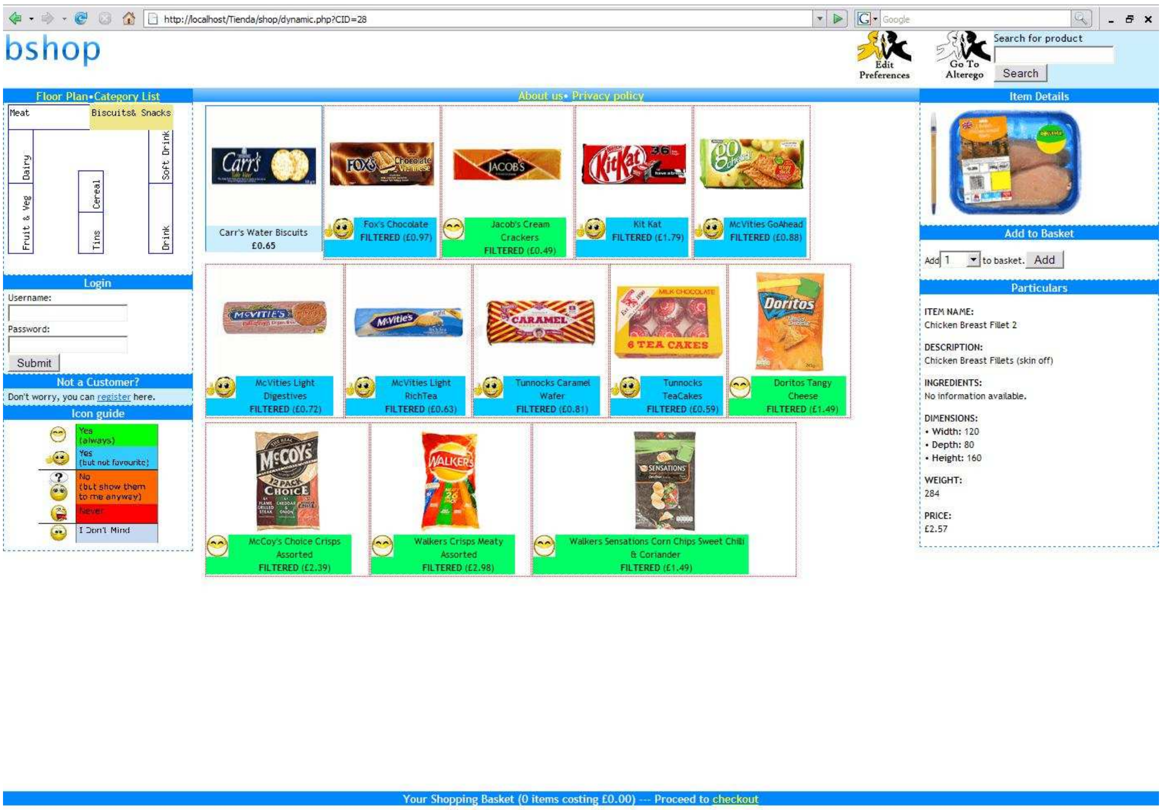


Figure 6.42: Bshop logged in Alter-Ego gold level - Layout and transition diagram.

Figure 6.43 shows the bshop's appearance when the user is directed from the Alter-Ego Web portal via the gold disclosure level. A more detailed item content is shown on the right. The catalogue, in the centre, is customised using the customer's preferences and sensitive information.



http://localhost/Tienda/shop/dynamic.php?CID=28

Carr's Water Biscuits £0.65
Fox's Chocolate FILTERED (£0.97)
Jacob's Cream Crackers FILTERED (£0.49)
Kit Kat FILTERED (£1.79)
McVities GoAhead FILTERED (£0.88)
McVities Light Digestives FILTERED (£0.72)
McVities Light RichTea FILTERED (£0.63)
Tunnocks Caramel Wafer FILTERED (£0.81)
Tunnocks TeaCakes FILTERED (£0.59)
Doritos Tangy Cheese FILTERED (£1.49)
McCoy's Choice Crisps Assorted FILTERED (£2.39)
Walkers Crisps Meaty Assorted FILTERED (£2.98)
Walkers Sensations Corn Chips Sweet Chilli & Coriander FILTERED (£1.49)

Not a Customer?
Don't worry, you can register here.

Icon guide

😁	Yes (always)
👍😀	Yes (but not favourite)
? 😐	No (but show them to me anyway)
😝	Never
😐	I Don't Mind

Item Details

Add to Basket
Add 1 to basket. Add

Particulars

ITEM NAME:
Chicken Breast Fillet 2

DESCRIPTION:
Chicken Breast Fillets (skin off)

INGREDIENTS:
No information available.

DIMENSIONS:
• Width: 120
• Depth: 80
• Height: 160

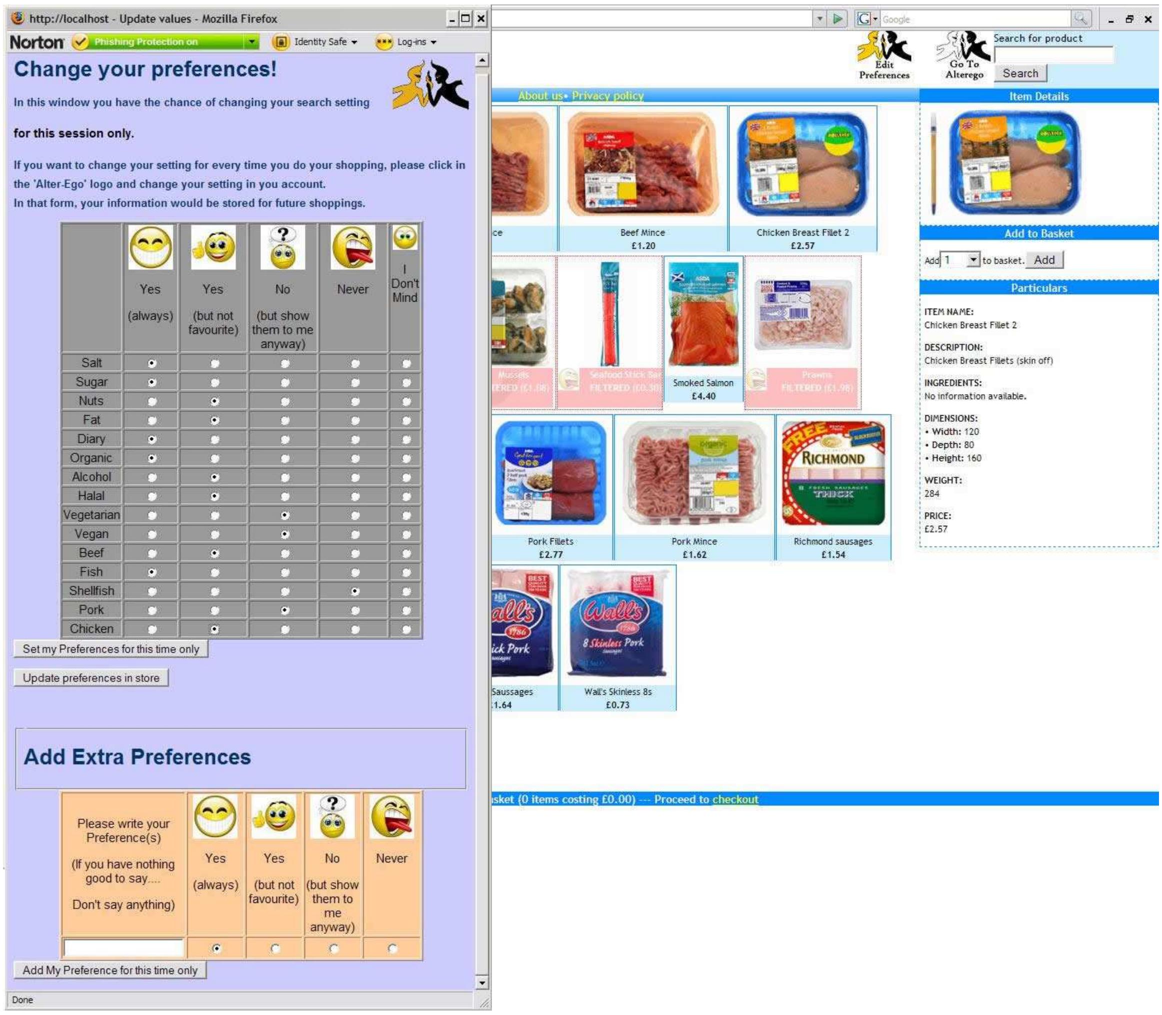
WEIGHT:
284

PRICE:
£2.57

Your Shopping Basket (0 items costing £0.00) --- Proceed to checkout

Figure 6.43: Bshop logged in Alter-Ego gold level - Implementation.

The customer is also allowed to change the preferences and the sensitive information for that particular session using an auxiliary window opened with the golden Alter-Ego icon. This option is named *Change Alter-Ego preferences.* In the extended functionality offered by the gold disclosure level, customers can, besides from modifying that shopping session's preferences, add new search items and its preferences. Customers are reminded that any change of preferences or sensitive information or any newly added features are only effective during the current shopping session. If customers decide to change the records permanently, the changes would need to be done directly in the Alter-Ego Web portal. This facility allows dynamic updating and adding of preferences while shopping. The implementation of this feature is shown in Figure 6.44.



http://localhost - Update values - Mozilla Firefox

Change your preferences!

In this window you have the chance of changing your search setting

for this session only.

If you want to change your setting for every time you do your shopping, please click in the 'Alter-Ego' logo and change your setting in you account.
In that form, your information would be stored for future shoppings.

	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never	I Don't Mind
Salt	●	○	○	○	○
Sugar	●	○	○	○	○
Nuts	○	●	○	○	○
Fat	○	●	○	○	○
Diary	●	○	○	○	○
Organic	●	○	○	○	○
Alcohol	○	●	○	○	○
Halal	○	●	○	○	○
Vegetarian	○	○	●	○	○
Vegan	○	○	●	○	○
Beef	○	●	○	○	○
Fish	●	○	○	○	○
Shellfish	○	○	○	●	○
Pork	○	○	●	○	○
Chicken	○	●	○	○	○

Set my Preferences for this time only

Update preferences in store

Add Extra Preferences

Please write your Preference(s) (If you have nothing good to say.... Don't say anything)	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never
	●	○	○	○

Add My Preference for this time only

Done

Item Details

Add 1 to basket. Add

Particulars

ITEM NAME:
Chicken Breast Fillet 2

DESCRIPTION:
Chicken Breast Fillets (skin off)

INGREDIENTS:
No information available.

DIMENSIONS:
- Width: 120
- Depth: 80
- Height: 160

WEIGHT:
284

PRICE:
£2.57

sket (0 items costing £0.00) --- Proceed to checkout

Figure 6.44: Bshop logged in Alter-Ego gold level - Changes and addition of preferences for the shopping session - Implementation.

During the checkout process, as Figure 6.45 shows, customers are shown a list of the used information as well as the standard information and the list of preferences and sensitive information that will be added to their record and be used on future occasions to assist their shopping. This list can be edited. Therefore, if customers decide that the options shown in this list do not reflect their requirements, they can amend or remove them. This feature informs and empowers the customer. The implementation is shown in Figure 6.46.

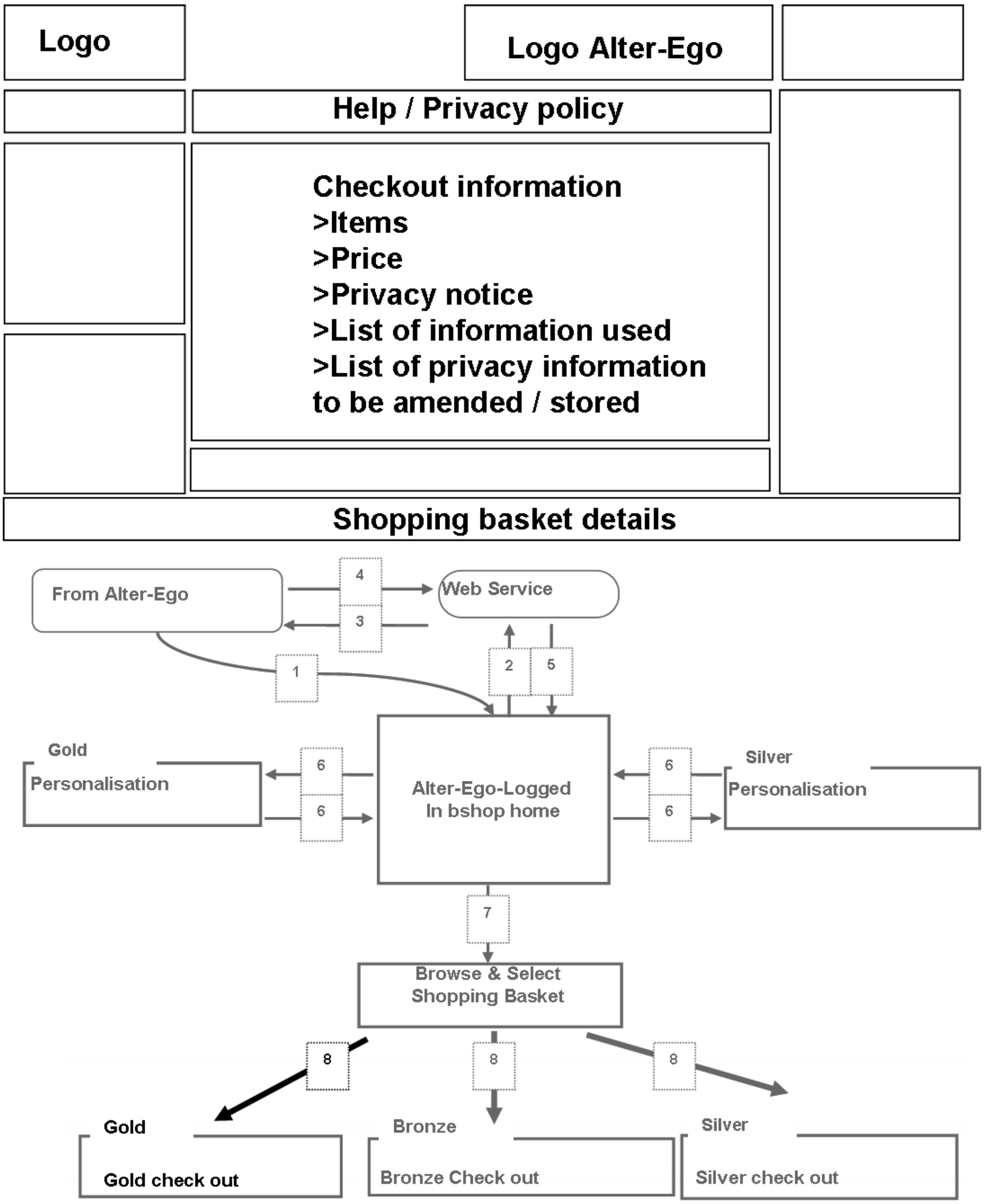
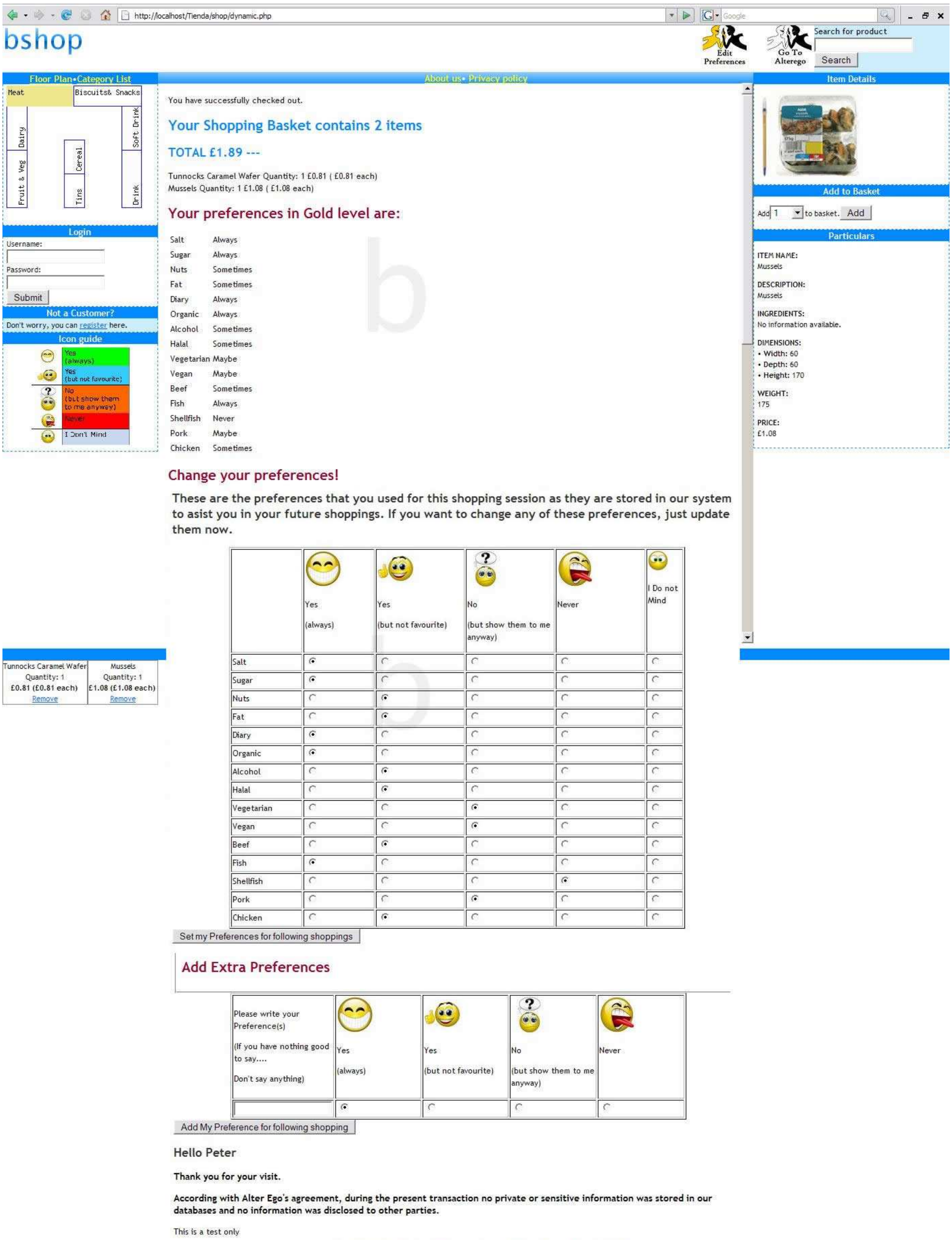


Figure 6.45: Bshop checkout when logged in Alter-Ego gold level - Layout and transition diagram.



You have successfully checked out.

Your Shopping Basket contains 2 items

TOTAL £1.89 ---

Tunnocks Caramel Wafer Quantity: 1 £0.81 (£0.81 each)
Mussels Quantity: 1 £1.08 (£1.08 each)

Your preferences in Gold level are:

Salt	Always
Sugar	Always
Nuts	Sometimes
Fat	Sometimes
Diary	Always
Organic	Always
Alcohol	Sometimes
Halal	Sometimes
Vegetarian	Maybe
Vegan	Maybe
Beef	Sometimes
Fish	Always
Shellfish	Never
Pork	Maybe
Chicken	Sometimes

Change your preferences!

These are the preferences that you used for this shopping session as they are stored in our system to asist you in your future shoppings. If you want to change any of these preferences, just update them now.

	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never	I Do not Mind
Salt	●	○	○	○	○
Sugar	●	○	○	○	○
Nuts	○	●	○	○	○
Fat	○	●	○	○	○
Diary	●	○	○	○	○
Organic	●	○	○	○	○
Alcohol	○	●	○	○	○
Halal	○	●	○	○	○
Vegetarian	○	○	●	○	○
Vegan	○	○	●	○	○
Beef	○	●	○	○	○
Fish	●	○	○	○	○
Shellfish	○	○	○	●	○
Pork	○	○	●	○	○
Chicken	○	●	○	○	○

Set my Preferences for following shoppings

Add Extra Preferences

Please write your Preference(s) (If you have nothing good to say.... Don't say anything)	Yes (always)	Yes (but not favourite)	No (but show them to me anyway)	Never
	●	○	○	○

Add My Preference for following shopping

Hello Peter

Thank you for your visit.

According with Alter Ego's agreement, during the present transaction no private or sensitive information was stored in our databases and no information was disclosed to other parties.

This is a test only

Tunnocks Caramel Wafer Quantity: 1 £0.81 (£0.81 each) Remove	Mussels Quantity: 1 £1.08 (£1.08 each) Remove

ITEM NAME:
Mussels

DESCRIPTION:
Mussels

INGREDIENTS:
No information available.

DIMENSIONS:
- Width: 60
- Depth: 60
- Height: 170

WEIGHT:
175

PRICE:
£1.08

Figure 6.46: Bshop checkout when logged in Alter-Ego gold level - Implementation.

6.7 Conclusions

The thesis statement proposes the creation of a privacy preserved shopping (PPSE) environment to assist customers in their shopping while ensuring that their privacy is preserved. This proposal considers the needs and value that information presents for both customers and e-tailers. Supporting the thesis statement divided it into two parts. First; the creation of a prototype of the PPSE environment and second; the performance of a user test to evaluate customer satisfaction and possible loyalty.

A Web portal was selected to implement the mediating third party within the controlled environment, where customers can disclose their information while remaining in control of their information. Therefore, the created prototype environment, and the details of its design and implementation were presented in foregoing Chapter 6. It can be concluded then, that the approach is feasible, supporting the first part of the thesis statement.

To prove the second part of the thesis statement, an evaluation of user satisfaction as indicator of customer loyalty in the use of the PPSE was carried out. The user satisfaction evaluation is reported in the next chapter.

Chapter 7

Evaluation and Results

Thesis statement

> *It is possible to develop a privacy preserving shopping environment (PPSE), which respects the customer's privacy needs while allowing the company to gather and use sufficient reliable customer-specified data to achieve a level of personalisation which can be used to encourage customer loyalty.*

7.1 Introduction

This chapter presents the validation of the second part of the thesis statement relied on a user test to determine potential customers' satisfaction as an reinforcer of customer loyalty [133]. In this chapter, three hypotheses were explored:

Hypothesis 1 People have privacy needs.

Hypothesis 2 The PPSE can satisfy these privacy needs.

Hypothesis 3 Users were satisfied with the PPSE.

Hypotheses 1 and 2 explore privacy needs (a fundamental part of the first part of the thesis statement) while hypothesis 3 explores directly the second part of the thesis statement; customer loyalty.

In the organisation of this chapter, Section 7.2 outlines the evaluation, its objectives, method context and the tasks performed by participants. The selection of participants is discussed in Section 7.3, while the analysis tools are introduced in Section 7.4. The exploratory analysis of the customers' privacy needs (hypothesis 1) is presented in Sections

7.5 and 7.7, while Section 7.8 presents the PPSE evaluation emphasising the customers' privacy needs. The results of the PPSE evaluation in regard to customer satisfaction and customer loyalty (hypothesis 2 and 3) are presented in Section 7.9. This chapter ends with a discussion of the results in Section 7.10

7.2 Evaluation

Evaluation, a measure of the quality of some other attribute of a system against a standard or scale, is used in this research to determine the usability of the system, e.g. how easy to use the system is, or the quality of the user experience when interacting with the system, e.g. how satisfying the interaction is [134]. The evaluation of a system is needed, as Sharp *et al.*[134] explain, "to check that users can use the product and that they like it, particularly if the design concept is new" [134][p586].

In Section 3.2.1, attitudinal factors were introduced. According to Shang *et al.*[131], attitudinal factors, also called intrinsic motives (those with subjective orientation), have a bigger influence on the customer's decisions while shopping, than do extrinsic motives (those with a practical orientation). Furthermore, surveys have shown that intrinsic motivations (such as satisfaction and perceived enjoyment) have a bigger role within the participant's reasons for shopping on-line over extrinsic motivations (such as perceived usefulness and ease of use) [131]. Since satisfaction is considered as "the sum of one's feelings or attitudes toward a variety of factors affecting the situation" [79][p192], and intrinsic factors have a high relevance in the user's decisions [131], we assume that the customer satisfaction would influence the likelihood of customers using the PPSE, and this can be considered a fairly reliable predictor of customer loyalty. Therefore, the evaluation of the second part of the thesis statement, *"which can be used to encourage customer loyalty"*, will focus on determining the customer's satisfaction towards the PPSE.

Satisfaction, on the other hand, has also been related to *quality*, specially since *user perceived quality* is defined as "the combination of product attributes which provide the greatest satisfaction to a specified user" [17][p116]. Furthermore, *quality of use* can be defined as "the extent to which a product satisfies stated and implied needs when used under stated conditions" [17][p116]. Therefore, if a system satisfies the user's needs under stated conditions, and if the user-perceived quality is related to the measure of user's satisfaction [17], the validation process focusing on measuring the customer satisfaction within the quality of use can be used as indicator to validate the PPSE and its quality.

In order to evaluate user's satisfaction, an adaptation of Bevan's *quality of use measures determined by the context of use* [17], shown in Figure 7.1, was used. In *the quality of use measures*, effectiveness, efficiency and satisfaction are obtained as a "result of the interaction between the user and product while carrying out a task in a technical, physical, social and organisational environment" [17][p119], and can be used to "evaluate the suitability of a product for use in a particular context" [17][p119]. Therefore, to evaluate satisfaction within the context of the PPSE environment quality of use, the main elements to measure satisfaction using Bevan's adapted approach are:

- the creation of an evaluation context,
- the definition of tasks, and
- the user interaction with those tasks.

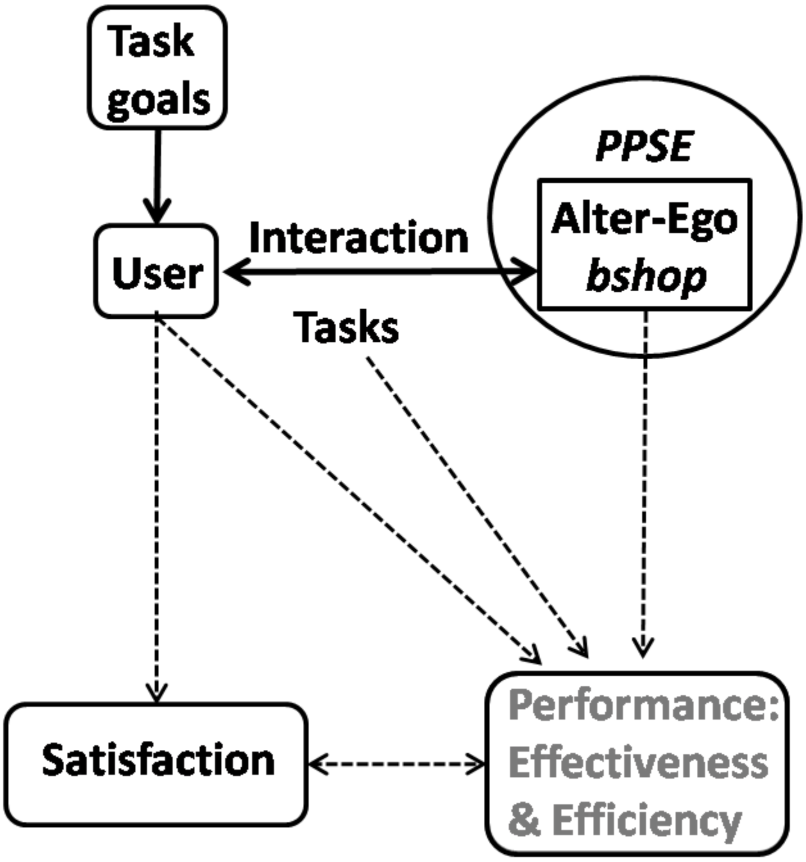


Figure 7.1: Evaluation of the PPSE using the quality of use measures determined by the context of use. *Adapted from [17]*

Figure 7.1 shows the determination of satisfaction and performance (the quality of use measures) as a result of the interaction between the user and the PPSE prototype while carrying out tasks in a context that contains groups of participants with different privacy perceptions.

From the three main elements measured by the *quality of use measures*, effectiveness relates to the percentage obtained from the measure of user's amount of completed tasks (also called quantity) times the degree to which the output achieves the task goals (also called quality)[17].

$$\text{task effectiveness} = 1/100(\text{quantity x quality})\ (\%)[17]$$

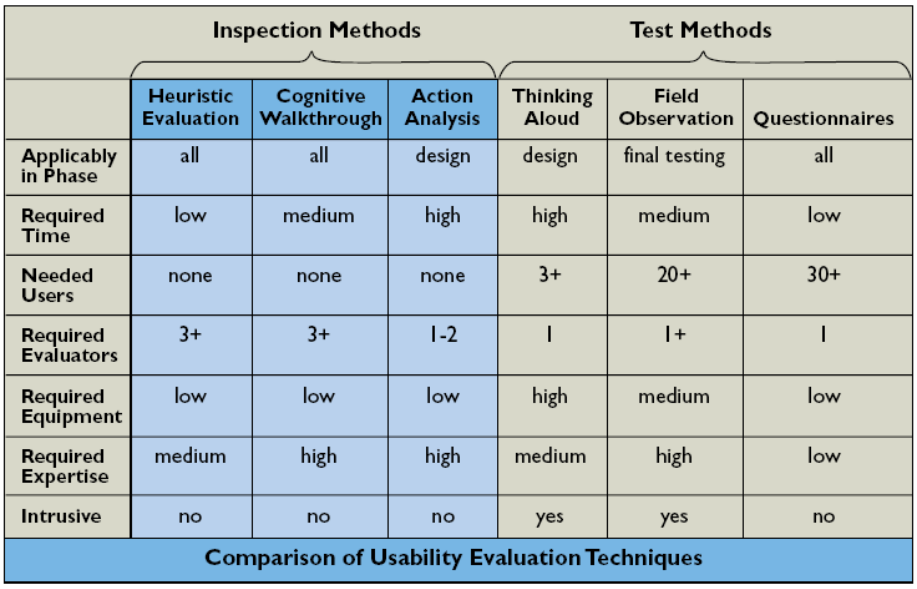
For instance, lets take the case where the effectiveness and efficiency of the grammar option of a text editor is evaluated. Task effectiveness would measure if the user could write, for example, 50 words using the grammar option offered by the text editor at least three times during the text writing. Efficiency, on the other hand, is the level of effectiveness achieved relative to the expenditure of resources. For instance, in the same example of the grammar option of the text editor, a temporal efficiency would measure that writing the same 50 words would be completed within a certain time, e.g. 5 minutes. However, since the goal of this evaluation is to measure the user's satisfaction, an extra effort was made to isolate satisfaction. This was achieved by keeping effectiveness and efficiency as constant as possible. The method used to evaluate satisfaction is outlined in the following section.

7.2.1 Evaluation method

Usability evaluations have been conducted by using different methods tailored according to the objective of the evaluation. These approaches, summarised in Figure 7.2, include [59]:

- Inspection methods: those which do not require participation of the end user, such as; heuristic evaluation, cognitive walkthroughs, and action analysis; and
- Test methods: those that involve the end user's participation, such as; thinking aloud, field observation and questionnaires.

Since the PPSE evaluation aims to assess user satisfaction, the users' participation is essential, therefore a *test method* was selected. From the *test methods* group, questionnaires (an evaluation tool) provide an indirect way of collecting user opinions, and are useful in studying the end user's interaction with the system and their preferred features [59]. Therefore, questionnaires were selected in order to evaluate the PPSE end user's satisfaction, their perception towards privacy violations and how susceptible they were towards invasion of their privacy. Likert scales were provided to elicit responses, due to the fact that these scales are used for "measuring opinions, attitudes, and beliefs, and consequently they are widely used for evaluating user satisfaction with products" [134][p314]. The questionnaires can be found in Appendix A.



	Inspection Methods			Test Methods		
	Heuristic Evaluation	**Cognitive Walkthrough**	**Action Analysis**	**Thinking Aloud**	**Field Observation**	**Questionnaires**
Applicably in Phase	all	all	design	design	final testing	all
Required Time	low	medium	high	high	medium	low
Needed Users	none	none	none	3+	20+	30+
Required Evaluators	3+	3+	1-2	1	1+	1
Required Equipment	low	low	low	high	medium	low
Required Expertise	medium	high	high	medium	high	low
Intrusive	no	no	no	yes	yes	no

Comparison of Usability Evaluation Techniques

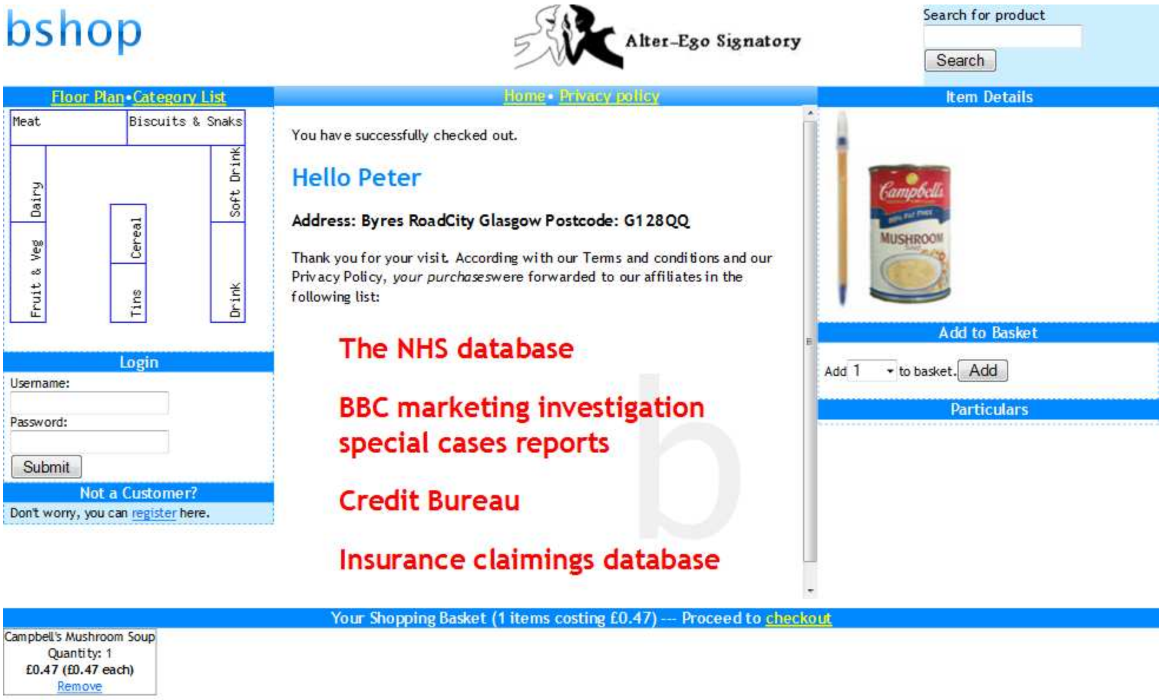
Figure 7.2: Comparison of usability evaluation techniques [59].

On the other hand, to evaluate the end user's satisfaction, an attempt was made to maintain a constant and positive performance (effectiveness and efficiency) during the execution of their tasks. This was done by means of the use of scenarios, and the provision of direct guidance from the evaluator at evaluation time (while participants followed tasks related to the scenarios) to overcome any difficulty that participants might experience. Scenarios describe human activities or tasks in an informal and narrative way. These scenario-based descriptions allow an exploration and discussion of contexts, needs and requirements [134], making them particularly suitable for evaluating the PPSE.

7.2.2 Evaluation context

In order to support hypothesis 1 (people have privacy needs), and evaluate the users' satisfaction obtained from the use of the PPSE (hypotheses 2 and 3), a shopping e-groceries scenario was designed. This scenario provided the context where the three privacy groupings (fundamentalists, pragmatic and unconcerned) [75] were shopping e-groceries. Participants did their shopping in a privacy protected environment (using the PPSE) and in a non-privacy protected environment, allowing them to compare both situations. Therefore two scenarios were designed to give participants the elements to compare different shopping environments.

The evaluation of the *privacy preserving shopping environment* (PPSE), required privacy violations to be explored. In this particular case a message, shown in Figure 7.3, informing participants that the information was disclosed to third parties without prompting for their specific previous consent was presented.



You have successfully checked out.

Hello Peter

Address: Byres RoadCity Glasgow Postcode: G128QQ

Thank you for your visit. According with our Terms and conditions and our Privacy Policy, *your purchases*were forwarded to our affiliates in the following list:

The NHS database

BBC marketing investigation special cases reports

Credit Bureau

Insurance claimings database

Figure 7.3: Privacy violation message.

In order to avoid ethical issues and to protect the participant's privacy, a *persona* ("rich description of typical user of the product under development" [134] [p481]) was used as the scenarios' principal actor. No credit card numbers were collected and the scenarios provided a fictitious address.

Both scenarios introduced "Peter", a *persona* with certain privacy requirements due to health problems, and his need to purchase groceries according to a shopping list with elements that, if misused, could impact his personal privacy. The scenarios can be found in Appendix B.

7.2.3 Definition of tasks

Since satisfaction and resulting customer loyalty were the main objectives of the evaluation, the definition of tasks had to be carefully designed so that effectiveness and efficiency were kept constant, or at least not problematical. To achieve this, participants were shown how to perform the tasks during a training session. After basic training, participants were given scenarios that contained lists of tasks to perform on behalf of "*Peter*". The experimental scenarios asked participants to perform tasks which involved the use of the Alter-Ego Web portal and bshop, such as:

- *Alter-Ego Web portal*
  - Task 1: Registration.
  - Task 2: Login.
  - Task 3: Provide *Peter's* preferences and sensitive information.
  - Task 4: Select the *disclosure level.*
  - Task 5: Select the *participant store.*
- *bshop*
  - Task 1: Select products from the scenario's *shopping list.*
  - Task 2: Checkout.
  - Task 3: Introduce *Peter's* checkout details.

7.2.4 Interaction with tasks

In order to support the hypotheses, participants were presented with a comparative context where privacy was *preserved* or *not preserved.* Therefore, to facilitate comparison, participants were required to use and comment on both environments in a random way. The approaches are shown in Figure 7.4

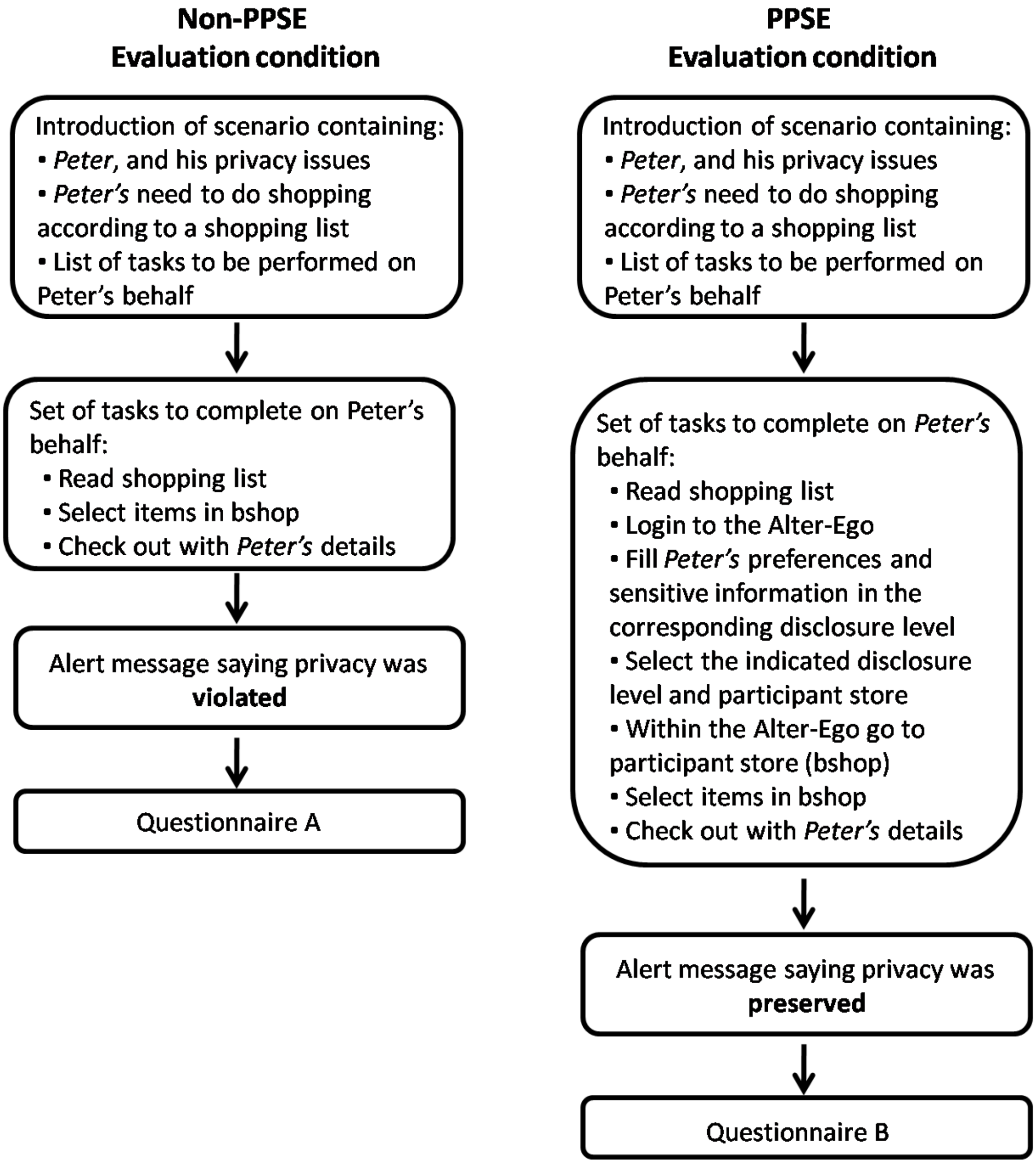

Non-PPSE
Evaluation condition
Introduction of scenario containing:
• *Peter*, and his privacy issues
• *Peter's* need to do shopping according to a shopping list
• List of tasks to be performed on Peter's behalf
Set of tasks to complete on Peter's behalf:
• Read shopping list
• Select items in bshop
• Check out with *Peter's* details
Alert message saying privacy was **violated**
Questionnaire A
PPSE
Evaluation condition
Introduction of scenario containing:
• *Peter*, and his privacy issues
• *Peter's* need to do shopping according to a shopping list
• List of tasks to be performed on Peter's behalf
Set of tasks to complete on *Peter's* behalf:
• Read shopping list
• Login to the Alter-Ego
• Fill *Peter's* preferences and sensitive information in the corresponding disclosure level
• Select the indicated disclosure level and participant store
• Within the Alter-Ego go to participant store (bshop)
• Select items in bshop
• Check out with *Peter's* details
Alert message saying privacy was **preserved**
Questionnaire B

Figure 7.4: Evaluation environments.

However, the results of the evaluation would be biased if participants were asked to use only the environments in one order (first the PPSE and second the non-PPSE, or first the non-PPSE and second the PPSE). Hence, to avoid influencing the outcome of the evaluation, the order of the use of the two environments was randomised. Two approaches were used: one with participants using the PPSE first and then the non-PPSE environment, and another in which they would use non-PPSE first and then the PPSE environment. Therefore, participants were randomly assigned to one of these two evaluation options, as shown in Figure 7.5.

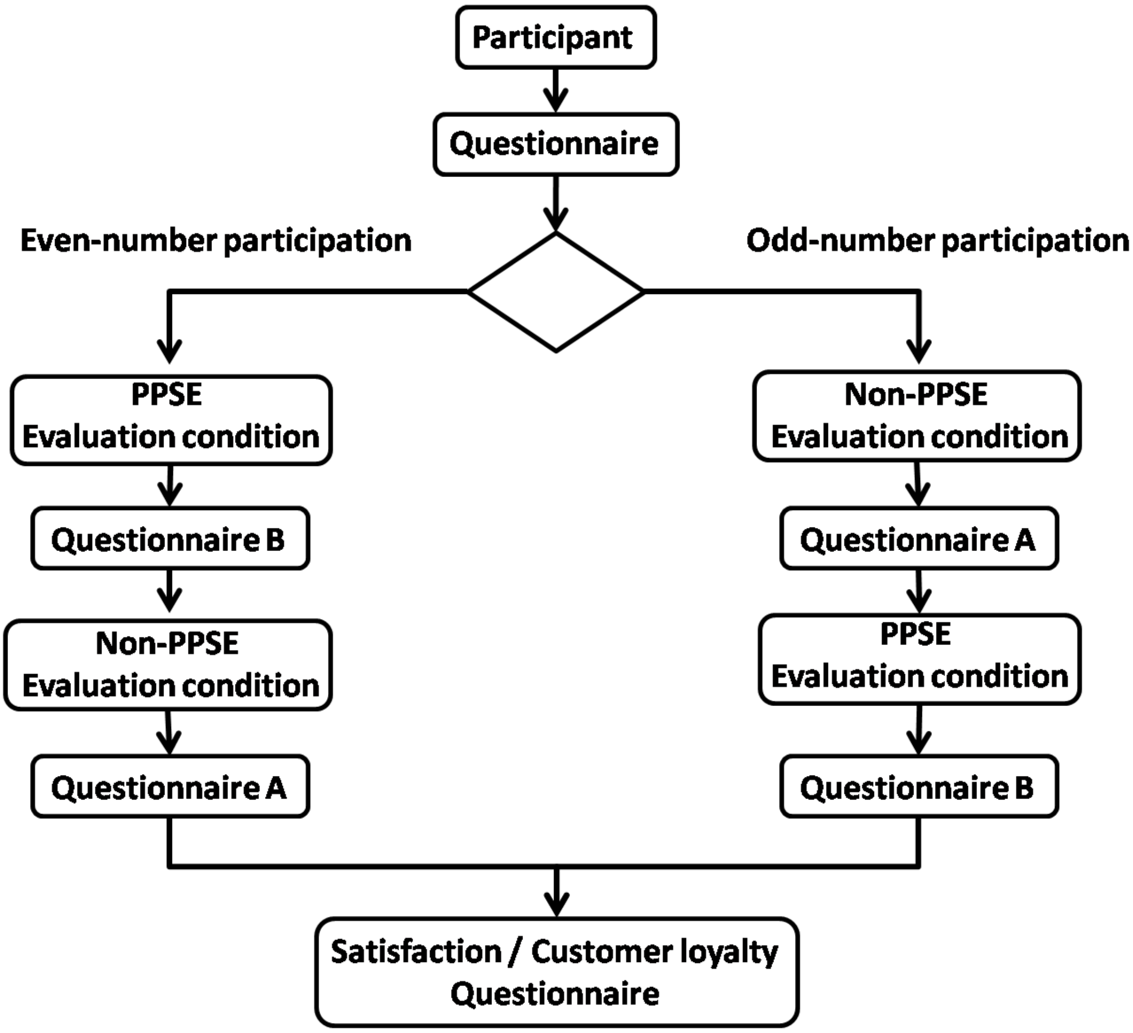

Participant
Questionnaire
Even-number participation
Odd-number participation
PPSE Evaluation condition
Questionnaire B
Non-PPSE Evaluation condition
Questionnaire A
Non-PPSE Evaluation condition
Questionnaire A
PPSE Evaluation condition
Questionnaire B
Satisfaction / Customer loyalty Questionnaire

Figure 7.5: Order of environment usage - Two evaluation conditions.

7.3 Participants

The recruitment of participants can be divided in two main categories the nature and the number of participants:

Nature - Since the time required for the evaluation was 45 minutes, the participants needed to have certain time flexibility to receive the basic training, and perform the tasks without a major disruption of their activities. At the same time, an open mind attitude towards new proposals and a basic knowledge of technology was required, and, since there was no monetary compensation[1], voluntarily participation in evaluations was required. Therefore, the target of the PPSE evaluation was directed to university students and staff.

Number - A call for volunteers was sent by Internet and by placing posters in strategic places, 41 participants that answered the call were recruited.

7.4 Statistical analysis

The participant's opinion collected in the questionnaires had the form of categorical data. According to Field [44, 43], when categorical data is collected, each person contributes once to each category and the results can be expressed in frequencies. To determine if there is a relationship between two variables expressed in categorical data, the analysis is performed using Chi square (χ^2) test [44, 43].

7.5 Results

The first questionnaire presented to the participants contained three main sections:

1. Demographics,
2. Participants' privacy perceptions and
3. Participants' privacy awareness.

The results of the evaluation are presented next.

7.6 Demographics

From the 41 participants, their main occupations were full time PhD students 19 (46%) followed by academic staff 7 (17%) full time undergraduate students 5 (12%) and other 10 (25%).

[1] Participants had previous knowledge that no final gratification was given, but chocolates were given after the test in gratitude to their participation.

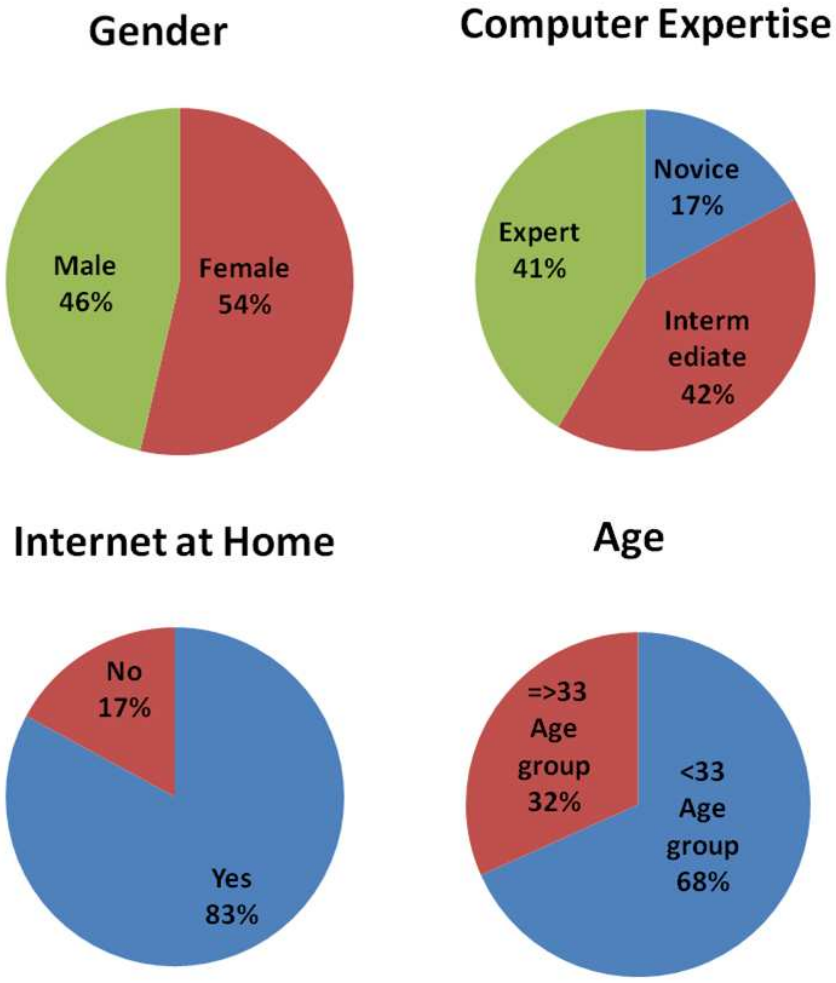


Figure 7.6: Main demographic results.

As Figure 7.6 shows, 22 participants (54%) were female and 19 (46%) male. When asked about their computer expertise using Internet, the participants considered themselves to be: expert 17 (42%), intermediate 17 (42%) and novice together with the option ‘I don’t know’ 7 (16%). 34 (83%) of the participants had Internet access at home. Ages varied from 18 to 62 years of age, where 28-32 had the most participants 13 (32%) followed by 18-22 with 9 (22%), and 23-27 with 6 (15%) participants.

As introduced in Section 2.2.6, the OECD sets 1995 as the start of e-commerce, therefore 1995 can be used to divide generations of shoppers. While younger generations would be raised with the existence of e-commerce as an every-day occurrence, older generations would perceive it as a novelty, having to go under an adaptation stage to incorporate e-commerce to their every day activities. Therefore, the participants' age groups have been merged into two groups. One group includes participants that were considered adults (21 years old) at the time e-commerce started and the other group contains participants younger 21 years at the time of the e-commerce's launch. Hence, one group is formed by participants younger than 33 years of age (< 33), 28 participants (68%), and the other includes participants over 33 years of age (=> 33), 13 participants (32%).

7.7 Privacy perceptions

In the first questionnaire, the participants' privacy perceptions were collected against the three different parts of the privacy categories, illustrated in Figure 7.7, were collected.

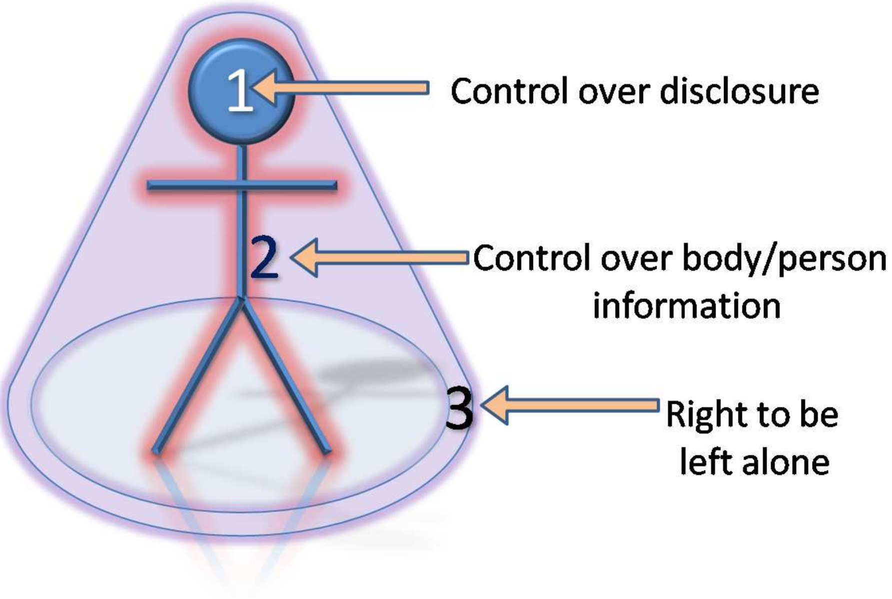


Figure 7.7: Privacy categories

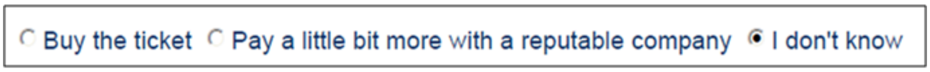
7.7.1 Control over disclosure

To determine the participants' opinion towards *control over disclosure*, a scenario-based question was presented. In this question participants had to choose whether they would take a risk and disclose their information, or not. The question was:

Question:

It has been a long day looking for cheap trips to visit New Zealand. Carol does *not have a lot of money to spend, but has agreed to be a bridesmaid and she has to do the trip. Suddenly, an unknown web site appears with the cheapest fare so far. She has no knowledge of that web site; she is tired and this will save some money. She has heard about Internet fraud, and she does not know what to do. What do you think Carol should do?*

Answer options:



○ Buy the ticket ○ Pay a little bit more with a reputable company ● I don't know

The results of this question are shown in Figure 7.8 and summarised on Table 7.1.

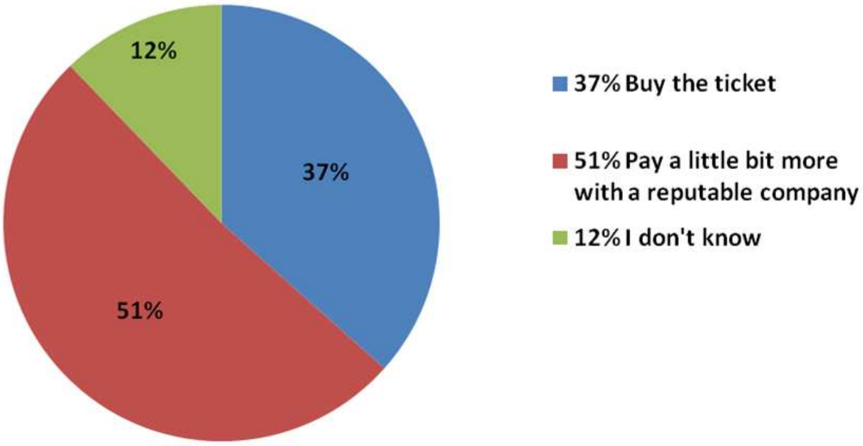


Figure 7.8: Privacy perceptions - Control over disclosure.

	Buy the ticket	Pay a little bit more with a reputable company	I don't know
Total			
	++ 37%	51%	12%
Age			
< 33	39%	54%	7%
=> 33	31%	46%	23%
Gender			
Male	42%	53%	5%
Female	32%	50%	18%
Internet expertise			
Expert	35%	53%	12%
Intermediate	35%	59%	6%
Novice	43%	29%	28%

Table 7.1: Privacy perceptions according to privacy definition - Control over disclosure

As shown in Table 7.1 most of the participants chose reliability over price. This suggests that the majority of the participants were not willing to take risks with unknown companies. The majority of participants that selected to buy the ticket or pay a little bit more for the ticket were from the < 33 group of age, male and considered themselves to have an intermediate and expert computer expertise. The group that selected the *"I don't know"* option had a majority of participants in the => 33 group of age, females and considered themselves to be novices in their computer expertise. This suggests that younger, male participants that consider their computer expertise to be expert or intermediate have a distinct opinion over their online shopping, and elder, female participants, that consider their computer expertise to be novice are not decided whether taking the risk of their online shopping or not.

7.7.2 Control over body / personal information

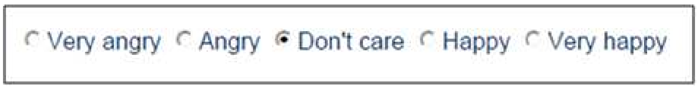
To determine the participants' opinion towards *control over body / personal information*, participants were presented with a scenario-based question involving an identity fraud attack due to lack of control over the disclosure of personal information.

Question:

> *Last year, Peter went to Rome on holiday. He was very careful with his credit cards, but one was copied and certain purchases were carried out on his behalf. It took him one year to solve the problem, but now he wants those purchases to*

be erased from his record and the bureau of credit does not want to do that, how do you think Peter feels?

Answer options:



○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

The results of this question are shown in Figure 7.9 and summarised in Table 7.2.

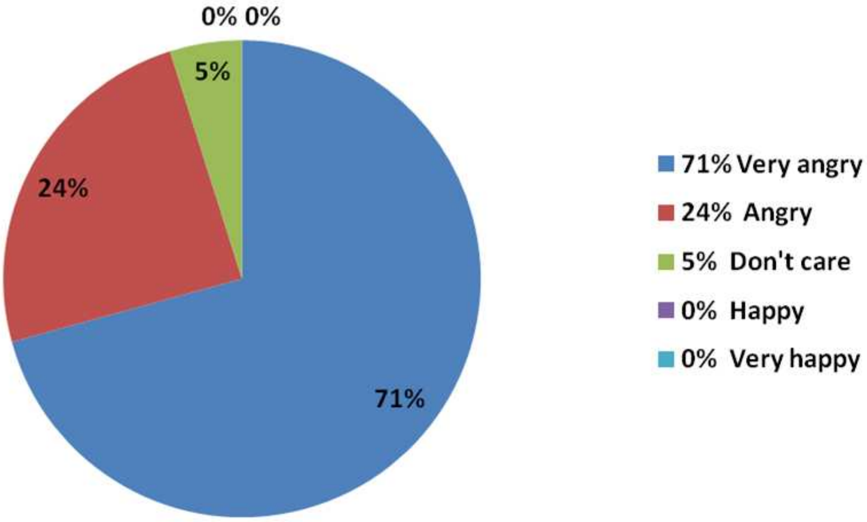


Figure 7.9: Privacy perceptions - Control over body / personal information.

Total					
	Very angry	Angry	Don't care	Happy	Very happy
	71%	24%	5%	0%	0%
Age					
	Very angry	Angry	Don't care	Happy	Very happy
< 33	68%	28%	4%	0%	0%
=> 33	77%	15%	8%	0%	0%
Gender					
	Very angry	Angry	Don't care	Happy	Very happy
Male	68%	32%	0%	0%	0%
Female	73%	18%	9%	0%	0%
Internet expertise					
	Very angry	Angry	Don't care	Happy	Very happy
Expert	65%	35%	0%	0%	0%
Intermediate	82%	18%	0%	0%	0%
Novice	57%	14%	29%	0%	0%

Table 7.2: Privacy perceptions according to privacy definition - Control over body / personal information

As shown in Table 7.2, the participants with a stronger reactions were from => 33 age group, female and considered their computer expertise to be intermediated, and the participants that selected *"Don't care"* were from age groups < 33 and => 33, females and considered their computer expertise as novice. This results suggest that, from a gender perspective, female novice computer users did not considered the recovery of their data as much as females with a higher computer expertise.

It is important to note that this question offered the options 'Happy' and 'Very happy' but these options were not chosen by any participant. It can be suggested that participants' reaction to this question was a combination of the lack of control over the disclosure and the increased awareness of the time required to recover from fraud.

7.7.3 Right to be left alone

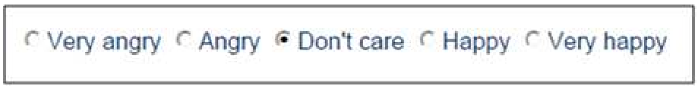
In this question, designed to explore the participants' attitudes towards the *right to be left alone and setting boundaries*, a third scenario-based question was presented.

Question:

> *John has very poor eye sight, so he needs to use bigger fonts on his computer. Marc, on the other hand, sits behind John and has very good eyesight, so he*

often reads over John's shoulder and can read the content of John's e-mail. If you were John, how would you feel?

Answer options:



○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

The results of this question are shown in Figure 7.10 and summarised in Table 7.3.

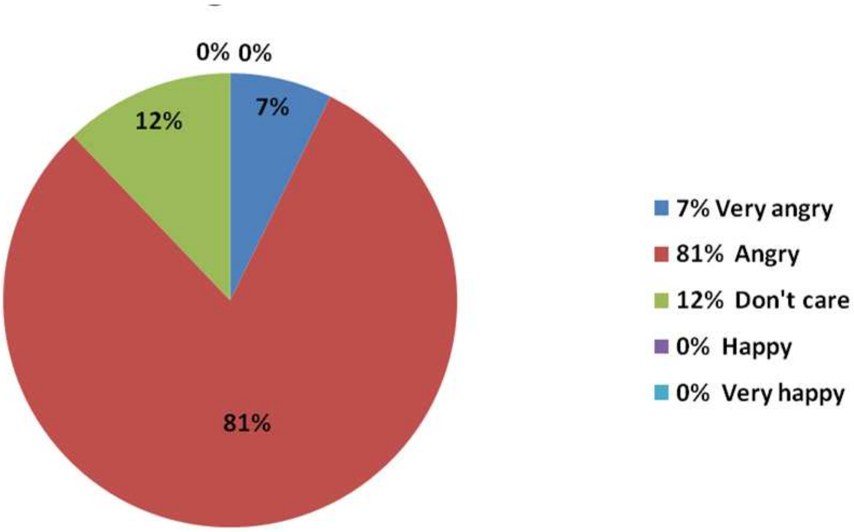


Figure 7.10: Privacy perceptions - Right to be left alone.

Total					
	Very angry	Angry	Don't care	Happy	Very happy
	7%	81%	12%	0%	0%

Age					
	Very angry	Angry	Don't care	Happy	Very happy
< 33	11%	82%	7%	0%	0%
=> 33	0%	77%	23%	0%	0%

Gender					
	Very angry	Angry	Don't care	Happy	Very happy
Male	11%	78%	11%	0%	0%
Female	4%	82%	14%	0%	0%

Internet expertise					
	Very angry	Angry	Don't care	Happy	Very happy
Expert	12%	82%	6%	0%	0%
Intermediate	6%	76%	18%	0%	0%
Novice	0%	86%	14%	0%	0%

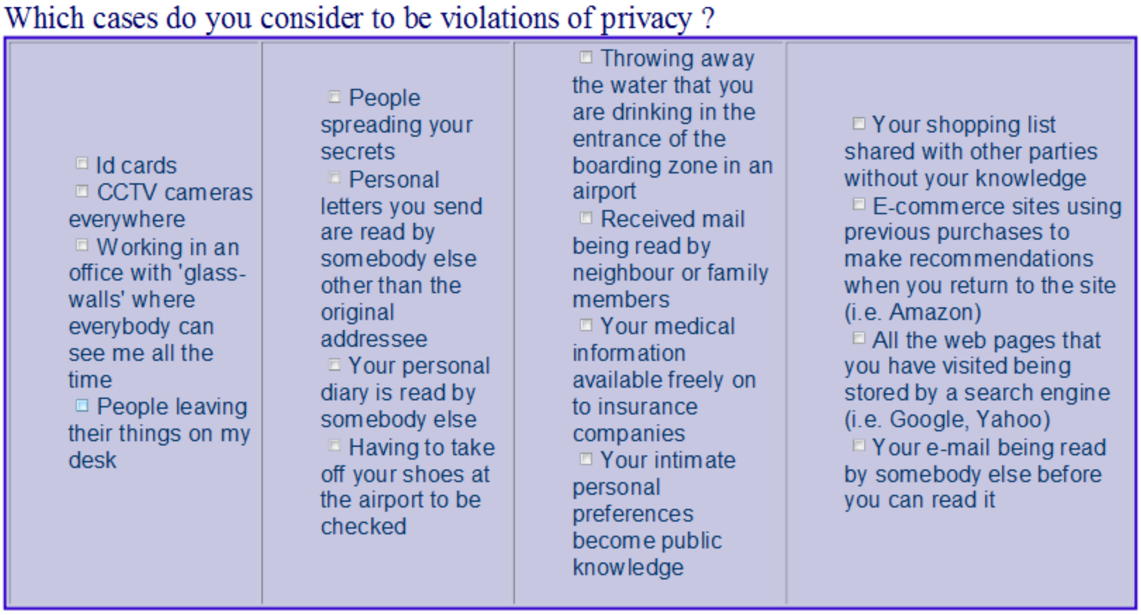
Table 7.3: Privacy perceptions according to privacy definition - Right to be left alone

From the results shown in Table 7.3, the participants with stronger reactions were from the < 33 age group, female and considered their computer expertise as novice, whereas the majority of participants that selected *"Don't care"* were from the => 33 age group, female that considered their computer expertise as intermediate. At the same time, the results of this question show that the majority of participants selected the option *'Angry'* for this option followed by *'Don't care'* and *'Very angry'*. These results suggest that whether participants are affected by the set of boundaries, their negative reaction is not the strongest, especially in the case of participants from the => 33 age group and participants that considered their computer expertise as novices.

7.7.4 Privacy perceptions according to privacy violations

To explore the participant's perceptions of privacy violations, several examples were obtained from newspapers, Internet news and every-day occurrences of situations that could be considered privacy violations, such as: the presence of CCTV cameras, personal diaries read by somebody else or intimate personal preferences becoming public knowledge.

Participants were presented with 16 different options, shown in Figure 7.11, that corresponded to one of the three parts of the privacy definition ('*control over disclosure*", "*control over body / personal information*" and "*the right to be left alone (set boundaries)*").



Which cases do you consider to be violations of privacy ?

□ Id cards
□ CCTV cameras everywhere
□ Working in an office with 'glass-walls' where everybody can see me all the time
□ People leaving their things on my desk

□ People spreading your secrets
□ Personal letters you send are read by somebody else other than the original addressee
□ Your personal diary is read by somebody else
□ Having to take off your shoes at the airport to be checked

□ Throwing away the water that you are drinking in the entrance of the boarding zone in an airport
□ Received mail being read by neighbour or family members
□ Your medical information available freely on to insurance companies
□ Your intimate personal preferences become public knowledge

□ Your shopping list shared with other parties without your knowledge
□ E-commerce sites using previous purchases to make recommendations when you return to the site (i.e. Amazon)
□ All the web pages that you have visited being stored by a search engine (i.e. Google, Yahoo)
□ Your e-mail being read by somebody else before you can read it

Figure 7.11: Privacy perceptions - 16 awareness options.

These options were merged in the following three categories; *clear perception of violation to privacy*(more than 11 occurrences), *violation to privacy* (between 5 and 10 occurrences), and *not-perceived violation* (less than 5 occurrences). Since the selection of these options reflect the participant's perceptions towards privacy, Westin's classification (*fundamentalist*, *pragmatic* and *unconcerned* [75], introduced on Section 4.4.1) can be used to assist the analysis according to the following association:

- Fundamentalist - participants selecting "*clear perception of violation to privacy*",
- Pragmatic - participants selecting "*violation to privacy*"
- Unconcerned - participants selecting "*not-perceived violation*".

7.7.4.1 Control over disclosure

The results for control over disclosure according to privacy perception, are shown in Figure 7.12 and summarised in Table 7.4.

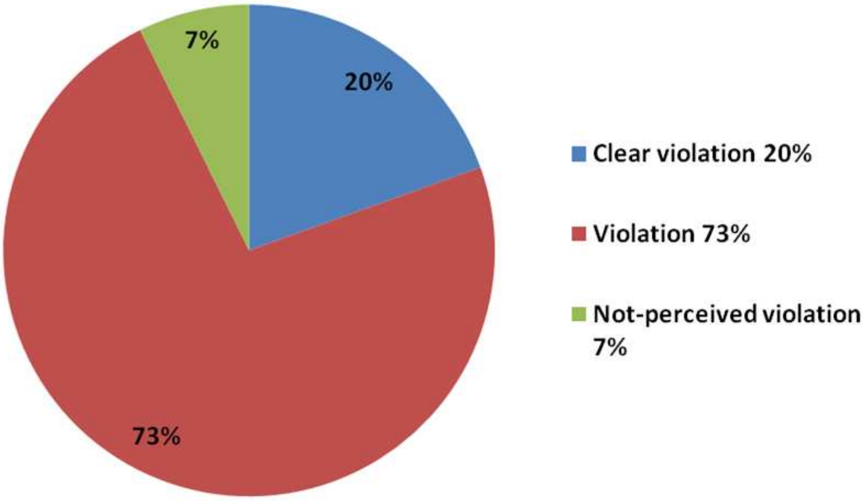


Figure 7.12: Privacy perceptions according to privacy violations - Control over disclosure.

Control over disclosure - General			
	Clear violation	Violation	Not-perceived violation
Total	20%	73%	7%

Age			
	Clear violation	Violation	Not-perceived violation
< 33	21%	75%	4%
=> 33	15%	70%	15%

Gender			
	Clear violation	Violation	Not-perceived violation
Male	5%	84%	11%
Female	32%	63%	5%

Internet Expertise			
	Clear violation	Violation	Not-perceived violation
Expert	12%	82%	6%
Intermediate	29%	65%	6%
Novice	15%	71%	14%

Table 7.4: Privacy perceptions according to privacy violations - Control over disclosure

From the results shown in Table 7.4, it can be seen that the majority of the participants selected the option "*Violation*". Participants selecting this option are associated with the *pragmatic* category in Westin's privacy indices.The participants that perceived violations to privacy were a majority of the $<$ 33 age group, female and considered their computer expertise as intermediate, whereas participants that perceived less violations to privacy were a majority of the => 33 age group, male and considered their computer expertise as novices. This results suggest that younger, female participants that considered themselves having an intermediate computer expertise are more sensitive towards violations related to control over disclosure than elder male participants that considered themselves having a novice computer expertise.

7.7.4.2 Control of disclosure over body/person information

The results for control of disclosure over body/person information according to privacy perception, are shown in Figure 7.13 and summarised in Table 7.5.

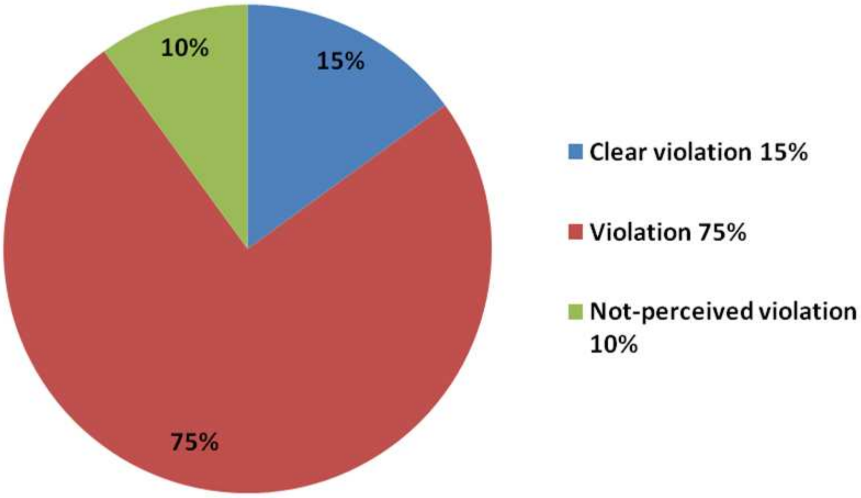


Figure 7.13: Privacy perceptions according to privacy violations - Control of disclosure over body/person information.

Control over body / personal information - General			
	Clear violation	Violation	Not-perceived violation
Total	15%	75%	10%

Age			
	Clear violation	Violation	Not-perceived violation
< 33	21%	79%	0%
=> 33	0%	69%	31%

Gender			
	Clear violation	Violation	Not-perceived violation
Male	11%	84%	5%
Female	18%	68%	14%

Internet Expertise			
	Clear violation	Violation	Not-perceived violation
Expert	18%	76%	6%
Intermediate	6%	88%	6%
Novice	28%	43%	29%

Table 7.5: Privacy perceptions according to privacy violations - Control of disclosure over body/person information

From the results shown in Table 7.5, it can be noticed that the majority of participants selected the option "*Violation*", associated with the *pragmatic* group. The participants that perceived more violations to privacy were a majority of the < 33 age group, male and considered their computer expertise as experts, whereas participants that perceived less violations to privacy were a majority of the => 33 age group, female and considered their computer expertise as novices. This results suggest that younger, male participants that considered themselves having an expert computer expertise are more sensitive towards violations related to control of disclosure over body/person information than elder female participants that considered themselves having a novice computer expertise.

7.7.4.3 The right to be left alone (set boundaries)

The category *The right to be left alone (set boundaries)*, presented an interesting change. While the majority of participants selected, again, the option "*Violation*", the second most selected option was "*Not-perceived violation*". The results are shown in Figure 7.14 and summarised in Table 7.6.

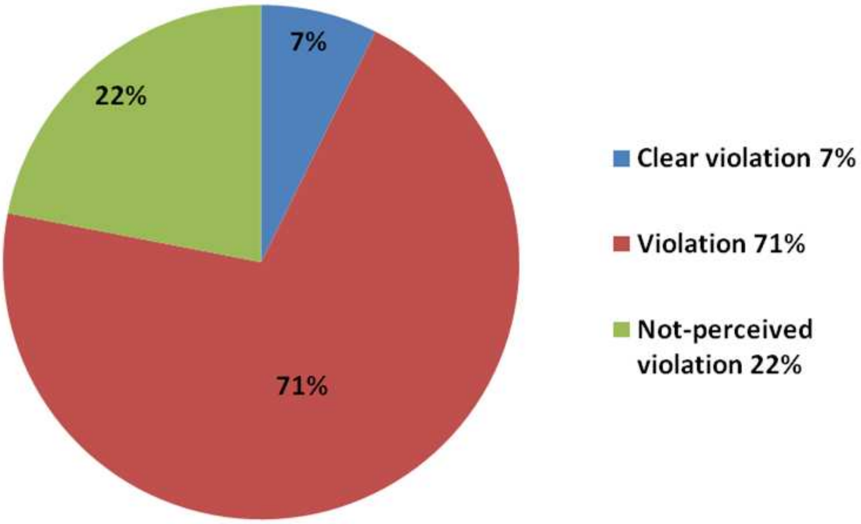


Figure 7.14: Privacy perceptions according to privacy violations - The right to be left alone (set boundaries).

The right to be left alone (set boundaries) - General			
	Clear violation	Violation	Not-perceived violation
Total	7%	71%	22%
Age			
	Clear violation	Violation	Not-perceived violation
< 33	4%	75%	21%
=> 33	15%	62%	23%
Gender			
	Clear violation	Violation	Not-perceived violation
Male	5%	69%	26%
Female	9%	73%	18%
Internet Expertise			
	Clear violation	Violation	Not-perceived violation
Expert	12%	70%	18%
Intermediate	6%	76%	18%
Novice	0%	57%	43%

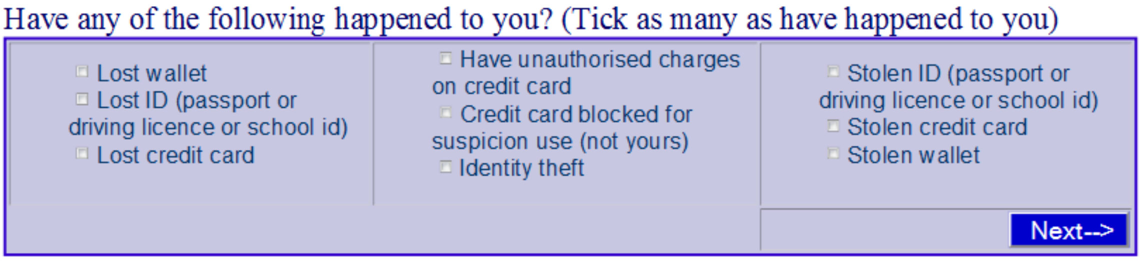
Table 7.6: Privacy perceptions according to privacy violations - The right to be left alone (set boundaries)

From the results shown in Table 7.6, it can be noticed that the majority of participants selected the option "*Violation*", associated to the *pragmatic* group.

The participants that perceived more violations to privacy were a majority of female and considered their computer expertise as expert, whereas participants that perceived less violations to privacy were a majority of male and considered their computer expertise as novices. This results suggest that female participants that considered themselves having an expert computer expertise are more sensitive towards violations related to the right to be left alone (set boundaries) than male participants that considered themselves having a novice computer expertise.

7.7.5 Awareness-based privacy perceptions

The objective of this questionnaire was to identify to what extent privacy violations have affected participants. Awareness can be created by previous experience, presentation of information or education, and previous privacy violations occurrences would raise privacy awareness as well. The questionnaire presented participants with nine options, shown in Figure 7.15.



Have any of the following happened to you? (Tick as many as have happened to you)

□ Lost wallet
□ Lost ID (passport or driving licence or school id)
□ Lost credit card

□ Have unauthorised charges on credit card
□ Credit card blocked for suspicion use (not yours)
□ Identity theft

□ Stolen ID (passport or driving licence or school id)
□ Stolen credit card
□ Stolen wallet

Figure 7.15: Awareness privacy perceptions - 9 privacy violations occurrences.

From the privacy violations occurrences (such as lost wallet, identity theft or stolen passport) the selection of more than four occurrences were considered as *high occurrences*, from 2 to 4 occurrences were considered *medium occurrences*, and 1 or 0 occurrences were considered *no occurrences*. The results are shown in Figure 7.16 and summarised in Table 7.7.

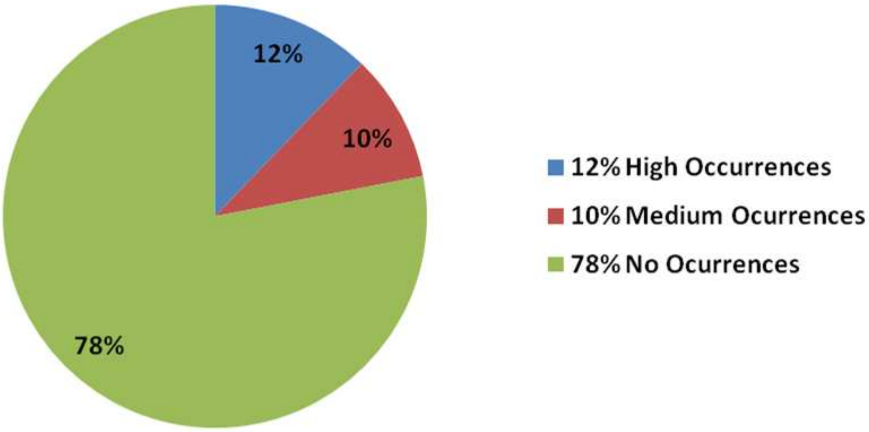


Figure 7.16: Awareness-based privacy perceptions.

Awareness-based privacy perceptions - General			
	High occurrences	Medium occurrences	No occurrences
Total	12%	10%	78%
Age			
	High occurrences	Medium occurrences	No occurrences
< 33	11%	11%	78%
=> 33	15%	8%	77%
Gender			
	High occurrences	Medium occurrences	No occurrences
Male	16%	10%	74%
Female	9%	9%	82%
Internet Expertise			
	High occurrences	Medium occurrences	No occurrences
Expert	18%	0%	82%
Intermediate	6%	23%	71%
Novice	14%	0%	86%

Table 7.7: Awareness-based privacy perceptions

From the results shown in Table 7.7, it can be noticed that the majority of participants have experienced few, if any, privacy violations. The participants with more experience in privacy violations were male that considered their computer expertise as expert, whereas female participants that considered their computer expertise as intermediate presented less privacy violations. This results show that male participants have been more in contact with privacy violations experiences than females.

7.8 PPSE evaluation

After the initial questionnaire, participants were directed to the PPSE evaluation according to their participation order (as shown in Figure 7.5). After performing the first set of tasks, participants were presented with questionnaires to obtain their opinion of the PPSE according to their perception of privacy and how it was preserved. The results of each questionnaire, are presented divided into two groups: group 1 (following the non-PPSE - PPSE order) and group 2 (following the PPSE - non-PPSE order). To provide a clear view of the results, only the options from the five point Likert scales selected by participants are presented.

7.8.1 Questionnaire A

To determine the participants' privacy needs, their privacy perceptions were collected using three questions after using the bshop. The questions all relate to the scenario presented in appendix B which concerns Peter, a persona with health problems who has made a number of purchases for a party for his own consumption. It is important to note that for group 1, this questionnaire was the first opinion gathered after performing their task and being presented with a message revealing that their information was being disclosed to third parties. Participants in group 2, on the other hand, had already used the Alter-Ego to store the scenario persona's preferences and this was the second time that they were using bshop. The first use of bshop by group 2 did not include the privacy violation message (see Figure 7.5).

7.8.1.1 Control over disclosure

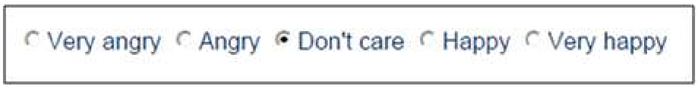
The first question aimed to obtain the participant's privacy perception of *control over disclosure*, after using the non-PPSE-bshop. The question is as follows:

Question:

> *There is no way for Peter to inform them that his shopping was not for him but for a party. If you were Peter, how do you think you would feel when you finished doing your shopping and saw that the information was being reported*

to the NHS database, BBC marketing investigation special cases reports or the Bureau of Credit Insurance claiming database?

Answer options:



○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

The responses are shown in Figure 7.17 and summarised in Table 7.8.

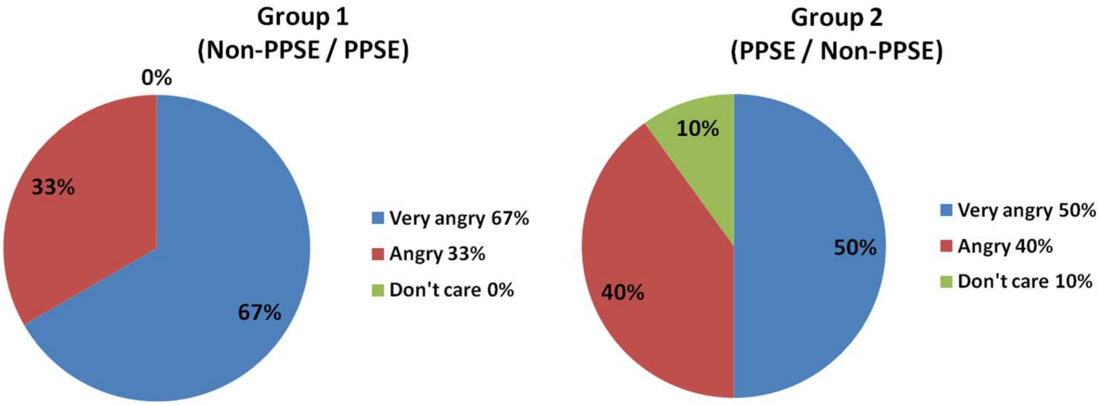


Figure 7.17: Questionnaire A - Control over disclosure.

Group 1			
	Very angry	Angry	Don't care
Total	67 %	33 %	0 %
Age			
< 33	54 %	46 %	0 %
=> 33	88 %	12 %	0 %
Gender			
Male	73 %	27 %	0 %
Female	60 %	40 %	0 %
Internet Expertise			
Expert	50 %	50 %	0 %
Intermediate	80 %	20 %	0 %
Novice	67 %	33 %	0 %
Group 2			
	Very angry	Angry	Don't care
Total	50 %	40 %	10 %
Age			
< 33	53 %	34 %	13 %
=> 33	40 %	60 %	0 %
Gender			
Male	50 %	25 %	25 %
Female	50 %	50 %	0 %
Internet Expertise			
Expert	56 %	33 %	11 %
Intermediate	29 %	71 %	0 %
Novice	75 %	0 %	25 %

Table 7.8: Questionnaire A, presented after performing tasks in the bshop. Group 1 and group 2 - Control over disclosure

The frequency of "*Very angry*, *Angry*, and *Don't care*" from group 1 and group 2 is not statistically different ($\chi^2 = 4.176$; P=0.124). Therefore the order of undertaking the scenario did not affect the participant's privacy perception in relation to *control over disclosure.*

It is important to notice that while the option *'Don't care'* was not selected by any participant of group 1, it was selected by 2 (10%) participants of group 2, both from the < 33 age group, male and one considered his computer expertise as novice while the other considered himself as expert. At the same time, more participants of group 1 selected the *'Very angry'* option over the *'Angry'* option, than the difference between options *'Very angry'* and *'Angry'* in group 2. Therefore, participants that undertook the non-PPSE - PPSE approach and were presented with a privacy violation message had a stronger opinion than those that were presented with the message after using the PPSE.

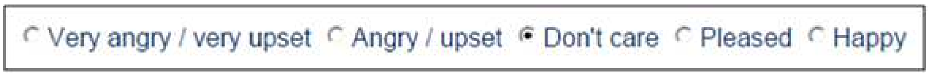
7.8.1.2 Control over body / personal information disclosure

The second question aimed to obtain the participant's privacy perception of *control over body / personal information disclosure* after using the non-PPSE-bshop. The question is as follows:

Question:

> *The message containing a list of Peter's purchases will go to his GP. The GP will assume that Peter has had a relapse and it was all Peter's fault. If you were Peter, how would you feel about it?*

Answer options:



○ Very angry / very upset ○ Angry / upset ● Don't care ○ Pleased ○ Happy

This question reflected the lack of control that participants had over the information that is disclosed to third parties, the results of the evaluation are summarised in Table 7.9 and shown in Figure 7.18.

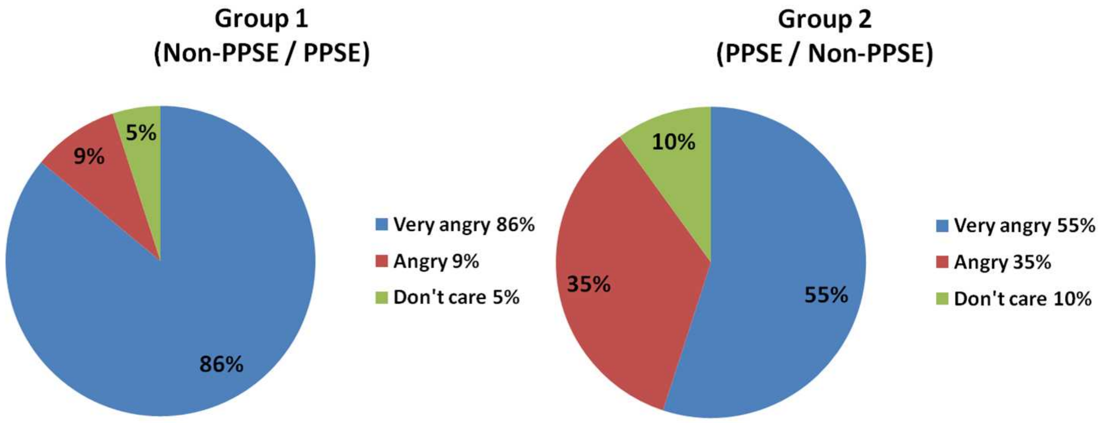


Figure 7.18: Questionnaire A - Control over body / personal information disclosure.

Group 1			
	Very angry	Angry	Don't care
Total	86 %	9 %	5 %
Age			
< 33	85 %	7 %	8 %
=> 33	87 %	13 %	0 %
Gender			
Male	100 %	0 %	0 %
Female	70 %	20 %	10 %
Internet Expertise			
Expert	88 %	12 %	0 %
Intermediate	100 %	0 %	0 %
Novice	33 %	34 %	33 %
Group 2			
	Very angry	Angry	Don't care
Total	55 %	35 %	10 %
Age			
< 33	53 %	33 %	14 %
=> 33	60 %	40 %	0 %
Gender			
Male	63 %	25 %	12 %
Female	50 %	42 %	8%
Internet Expertise			
Expert	56 %	33 %	11 %
Intermediate	43 %	43 %	14 %
Novice	75 %	25 %	0 %

Table 7.9: Questionnaire A, presented after performing tasks in the bshop. Group 1 and group 2 - Control over body / personal information disclosure

The frequency of "*Very angry*, *Angry*, and *Don't care*" from group 1 and group 2 is not statistically different ($\chi^2 = 2.445$; P=0.655). Therefore the order of undertaking the scenario did not affect the participant's privacy perception in relation to *control over body / personal information disclosure*. From group 1, a female participant from age group < 33 that considered herself as novice in computer expertise selected *"Don't care"* while participants from group 2 that selected the same option were from the same age group (< 33), this suggests that only younger participants were not concerned about their information being sent to their GP or what would the GP considered.

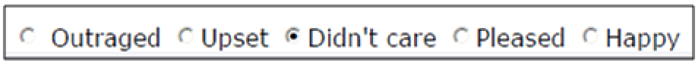
7.8.1.3 Right to be left alone (set boundaries)

The third question was aimed to explore the participants' perception of the *right to be left alone and set boundaries*, in relation to the activities performed in the non-PPSE-bshop after the privacy-violation message. Participants were asked to answer the following question:

Question:

If you were Peter, how would you feel when you realise that you agreed by default to the e-stores terms and conditions that permitted your information to be disclosed?

Answer options:



○ Outraged ○ Upset ● Didn't care ○ Pleased ○ Happy

The results are shown in Figure 7.19 and summarised in Table 7.10.

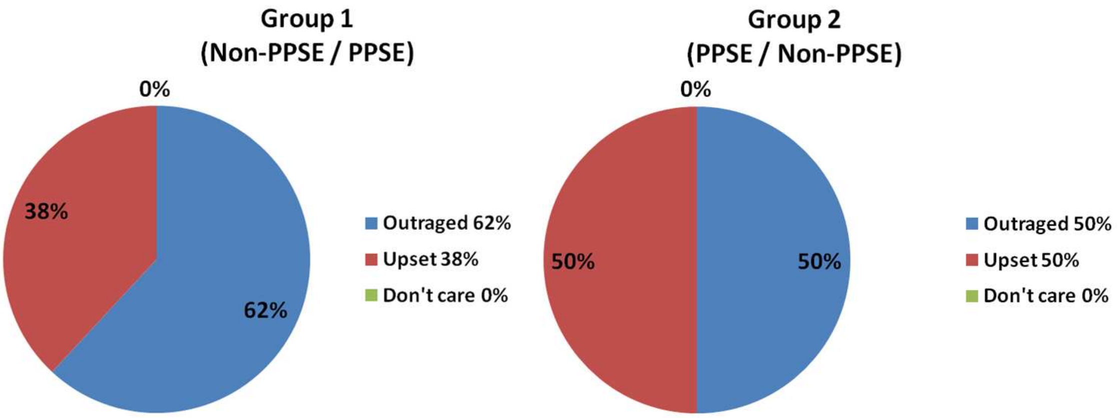


Figure 7.19: Questionnaire A - Right to be left alone (set boundaries).

Group 1			
	Outraged	Upset	Don't care
Total	62 %	38 %	0 %
Age			
< 33	54 %	46 %	0 %
=> 33	75 %	25 %	0 %
Gender			
Male	73 %	27 %	0 %
Female	80 %	20 %	0 %
Internet Expertise			
Expert	50 %	50 %	0 %
Intermediate	70 %	30 %	0 %
Novice	67 %	33 %	0 %
Group 2			
	Outraged	Upset	Don't care
Total	50 %	50 %	0 %
Age			
< 33	47 %	53 %	0 %
=> 33	60 %	40 %	0 %
Gender			
Male	37 %	63 %	0 %
Female	58 %	42 %	0 %
Internet Expertise			
Expert	56 %	44 %	0 %
Intermediate	29 %	71 %	0 %
Novice	75 %	25 %	0 %

Table 7.10: Questionnaire A, presented after performing tasks in the bshop. Group 1 and group 2 - Right to be left alone (set boundaries)

The frequency of "*Outraged*, *Upset* and *Don't care*" from group 1 and group 2 is not statistically different ($\chi^2 = 0$; P=0.675). Therefore the order of undertaking the scenario did not affect the participant's privacy perception in relation to *right to be left alone (set boundaries)*. The option "*Outraged*" was selected from more female participants from group 1, group age => 33 that considered their computer expertise as intermediate and a majority of female participants from group 2, group age => 33 that considered their computer expertise as novices.

7.8.2 Questionnaire B

After performing the tasks involving the PPSE, involving the use of the Alter-Ego portal and shopping on the bshop, the participants were presented with a questionnaire that aimed to evaluate their opinion against the three levels of privacy; "*control over disclosure*", "*control over body / personal information*" and "*the right to be left alone (set boundaries)*". In the case of participants in group 2, this was their first activity. The results are presented next.

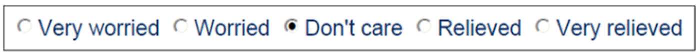
7.8.2.1 Control over disclosure

The question to obtain the participant's opinion after using the PPSE in relation with their perception of *control over disclosure* was:

Question:

> *Peter could do his personal shopping and the Christmas list shopping on separate PLA levels (silver level). His preferences would not be mixed with the Christmas ones. If you were Peter, how would you feel?*

Answer options:



○ Very worried ○ Worried ● Don't care ○ Relieved ○ Very relieved

The results are shown in Figure 7.20 and summarised in Table 7.11.

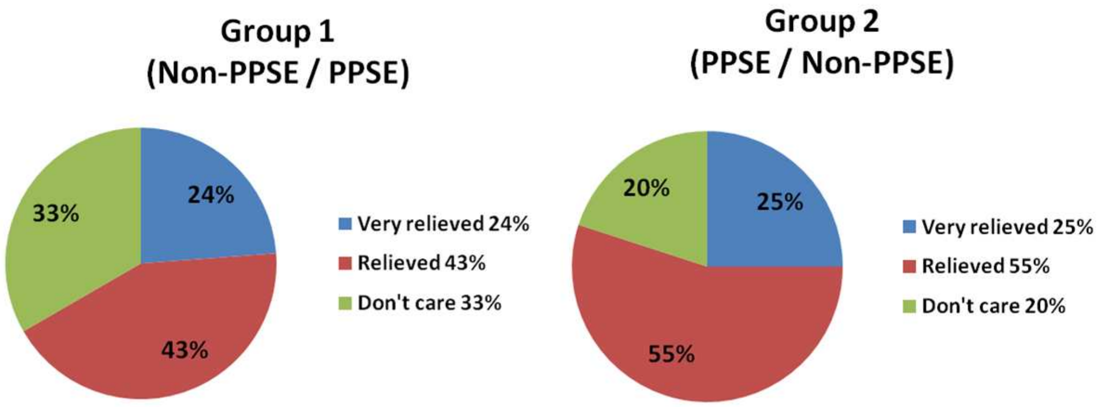


Figure 7.20: Questionnaire B - Control over disclosure..

Group 1			
	Very relieved	Relieved	Don't care
Total	24 %	43 %	33 %
Age			
< 33	31 %	38 %	31 %
=> 33	12 %	50 %	38 %
Gender			
Male	27 %	37 %	36 %
Female	20 %	50 %	30 %
Internet Expertise			
Expert	25 %	37 %	38 %
Intermediate	30 %	40 %	30 %
Novice	0 %	67 %	33 %
Group 2			
	Very relieved	Relieved	Don't care
Total	25 %	55 %	20 %
Age			
< 33	20 %	67 %	13 %
=> 33	40 %	20 %	40 %
Gender			
Male	13 %	63 %	24 %
Female	33 %	50 %	17 %
Internet Expertise			
Expert	33 %	56 %	11 %
Intermediate	29 %	42 %	29 %
Novice	0 %	75 %	25 %

Table 7.11: Questionnaire B, presented after performing tasks using PPSE. Group 1 and group 2 - Control over disclosure

The frequency of “*Very relieved*, *Relieved* and *Don't care*” from group 1 and group 2 is not statistically different ($\chi^2 = 2.5$; P=0.645). Therefore the order of undertaking the scenario did not affect the participant's privacy perception in relation to *control over disclosure*.

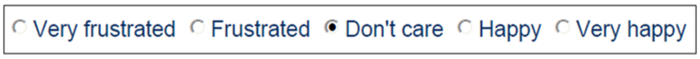
7.8.2.2 Control over body / personal information disclosure

To evaluate the participant's perspective in *control over body / personal information* disclosure after the use of the PPSE, participants were presented with the following question;

Question:

> *Peter could enter his preferences and his sensitive information in the Alter-Ego to be passed to any participant store. If you were Peter, how would you feel when you did not have to provide the preferences to every one of the sites that you do your shopping with?*

Answer options:



○ Very frustrated ○ Frustrated ● Don't care ○ Happy ○ Very happy

The results are shown in Figure 7.21 and summarised in Table 7.12.

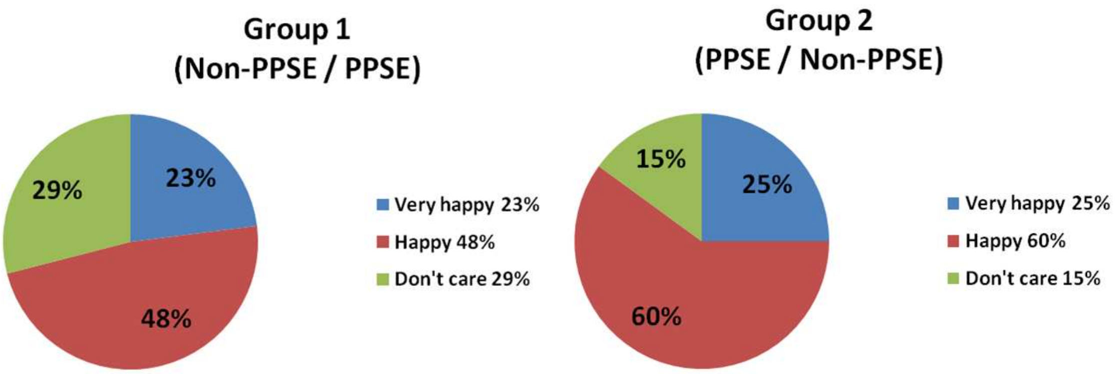


Figure 7.21: Questionnaire B - Control over body / personal information disclosure.

Group 1			
	Very happy	Happy	Don’t care
Total	23 %	48 %	29 %
Age			
< 33	31%	46 %	23 %
=> 33	13 %	50 %	37 %
Gender			
Male	27 %	36 %	37 %
Female	20 %	60 %	20 %
Internet Expertise			
Expert	25 %	50 %	25 %
Intermediate	30 %	30 %	40 %
Novice	0 %	100 %	0 %
Group 2			
	Very happy	Happy	Don’t care
Total	25 %	60 %	15 %
Age			
< 33	27%	60 %	13 %
=> 33	20 %	60 %	20 %
Gender			
Male	25 %	50 %	25 %
Female	25 %	67 %	8 %
Internet Expertise			
Expert	44 %	44 %	12 %
Intermediate	14 %	71 %	15 %
Novice	0 %	75 %	25 %

Table 7.12: Questionnaire B, presented after performing tasks using PPSE. Group 1 and group 2 - Control over body / personal information disclosure

The frequency of “*Very happy*, *Happy* and *Don’t care*” from group 1 and group 2 is not statistically different (χ^2 = 5.73; P=0.220). Therefore the order of undertaking the scenario did not affect the participant’s privacy perception in relation to *control over body / personal information disclosure.* It is important to note that all participants from group 1 that considered their computer expertise as novice selected the option “*Happy*” as well as a majority of novices of group 2, this suggests that novices appreciated not having to introduce their preferences in every e-store.

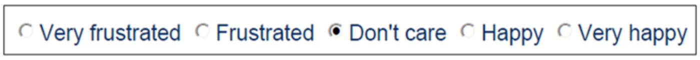
7.8.2.3 Right to be left alone (set boundaries)

Finally, the question presented to explore the participant's perception in relation to the *right to be left alone (set boundaries)* was:

Question:

Peter shops at bshop, which is known to have an affiliated insurance company bshopMed. However, by using Alter-Ego he has prevented them from sending details of the Christmas party purchases to bshopMed where he is insured. If you were Peter, how would you feel?

Answer options:



○ Very frustrated ○ Frustrated ● Don't care ○ Happy ○ Very happy

The results are shown in Figure 7.22, and summarised in Table 7.13.

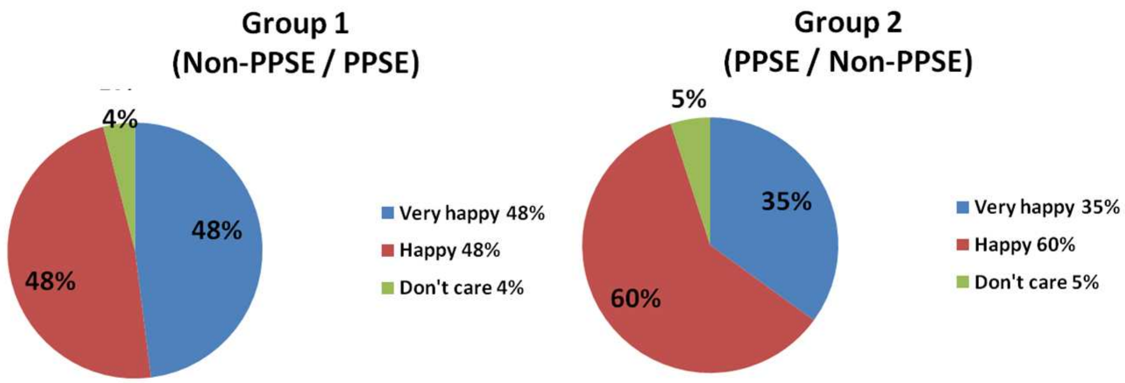


Figure 7.22: Questionnaire B - Right to be left alone (set boundaries).

Group 1			
	Very happy	Happy	Don't care
Total	48 %	48 %	4 %
Age			
< 33	54 %	38 %	8 %
=> 33	38 %	62 %	0 %
Gender			
Male	40 %	60 %	0 %
Female	55 %	36 %	9%
Internet Expertise			
Expert	50 %	50 %	0 %
Intermediate	50 %	50 %	0 %
Novice	33 %	34 %	33 %
Group 2			
	Very happy	Happy	Don't care
Total	35 %	60 %	5 %
Age			
< 33	40 %	60 %	0 %
=> 33	20 %	60 %	20 %
Gender			
Male	33 %	59 %	8 %
Female	38 %	62 %	0 %
Internet Expertise			
Expert	56 %	44 %	0 %
Intermediate	0 %	100 %	0 %
Novice	50 %	25 %	25 %

Table 7.13: Questionnaire B, presented after performing tasks using PPSE. Group 1 and group 2 - Right to be left alone (set boundaries)

The frequency of "*Very happy*, *Happy* and *Don't care*" from group 1 and group 2 is not statistically different ($\chi^2 = 3.135$; P=0.536). Therefore the order of undertaking the scenario did not affect the participant's privacy perception in relation to *right to be left alone (set boundaries)*. It is important to notice that all participants from group 2 that considered their computer expertise as intermediate, selected the option "*Happy*", this suggest that these participants valued that their information was not send to third parties.

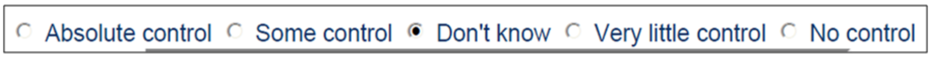
7.8.2.4 Control over privacy

To isolate the participant's perception about the control over their privacy when using PPSE, their opinion was directly asked using the following question:

Question:

Do you think that Peter has control over his privacy?

Answer options:



○ Absolute control ○ Some control ● Don't know ○ Very little control ○ No control

The results are shown in Figure 7.23 and summarised in Table 7.14.

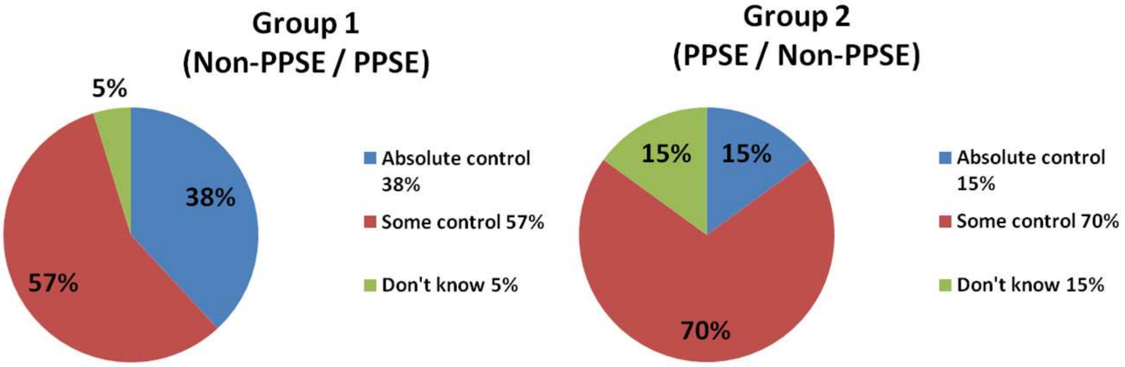


Figure 7.23: Questionnaire B - Control over privacy.

Group 1			
	Absolute control	Some control	Don’t know
Total	38 %	57 %	5 %
Age			
< 33	54 %	46 %	0 %
=> 33	13 %	75 %	12 %
Gender			
Male	30 %	60 %	10 %
Female	45 %	55 %	0 %
Internet Expertise			
Expert	25 %	75 %	0 %
Intermediate	50 %	40 %	10 %
Novice	33 %	67 %	0 %
Group 2			
	Absolute control	Some control	Don’t know
Total	15 %	70 %	15 %
Age			
< 33	13 %	80 %	7 %
=> 33	20 %	40 %	40 %
Gender			
Male	25 %	50 %	25 %
Female	0 %	100 %	0%
Internet Expertise			
Expert	22 %	78 %	0 %
Intermediate	0 %	71 %	29 %
Novice	25 %	50 %	25 %

Table 7.14: Questionnaire B, presented after performing tasks using PPSE. Group 1 and group 2 - Question 4

The frequency of “*Absolute control*, *Some control* and *Don’t know*” from group 1 and group 2 is not statistically different (χ^2 = 6.094; P=0.192). Therefore the order of undertaking the scenario did not affect the participant’s perception in relation to *control over privacy*. It is important to note that all female participants from group 2 selected the option “*Some control*” while a majority of females from group 1 selected the same option. This suggests that whether both groups of female participants perceived to have control over privacy, the group that used the Non-PPSE environment first (group 1), perceived to have more control over privacy than participants from group 2.

7.9 Final questionnaire

As introduced in Section 7.2.4 - Figure 7.5, after performing the tasks contained in the two scenarios, and answering the corresponding questionnaires, participants were asked to answer a general questionnaire. When participants arrived to this questionnaire, they already had the experience of shopping with and without the PPSE, therefore, the objective of this final questionnaire was to explore their overall opinion, their satisfaction and their possible customer loyalty. The details of the results are presented next.

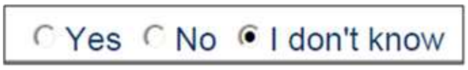
7.9.1 Recommending the PPSE

The first two questions of the final questionnaire aimed to explore if participants would recommend the use of PPSE. Since the scenario's persona had certain privacy requirements, the participant's recommendation would reflect their perception of the use of PPSE in cases with specific privacy needs. The first question to explore the participants' recommendations was:

Question:

Would you suggest that Peter should use the Alter-Ego to assist his shopping?

Answer options:



○ Yes ○ No ● I don't know

The results to the question are shown in Figure 7.24 and summarised in Table 7.15.

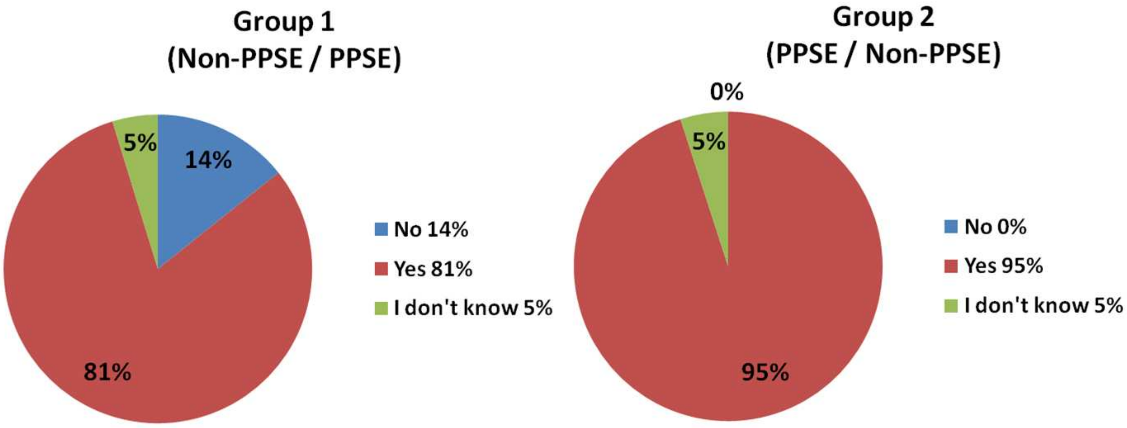


Figure 7.24: Final questionnaire - Recommendations to *Peter*.

Group 1			
	No	Yes	Don't know
Total	14 %	81 %	5 %
Age			
< 33	8 %	92 %	0 %
=> 33	25 %	62 %	13 %
Gender			
Male	0 %	91 %	9 %
Female	30 %	70 %	0 %
Internet Expertise			
Expert	13 %	74 %	13 %
Intermediate	10 %	90 %	0 %
Novice	33 %	67 %	0 %
Group 2			
	No	Yes	Don't know
Total	0 %	95 %	5 %
Age			
< 33	0 %	93 %	7 %
=> 33	0 %	100 %	0 %
Gender			
Male	0 %	88 %	12 %
Female	0 %	100 %	0 %
Internet Expertise			
Expert	0 %	100 %	0 %
Intermediate	0 %	86 %	14 %
Novice	0 %	100 %	0 %

Table 7.15: Final questionnaire - Recommendations to *Peter*

The frequency of "*No*, *Yes* and *Don't know*" from group 1 and group 2 is not statistically different ($\chi^2 = 0.263$; P=0.877). Therefore the order of undertaking the scenario did not affect the participant's decision in recommending the PPSE to "*Peter*". It is important to note that whether a group of female participants in group 1 selected to suggest that Peter should not use the Alter-Ego to assist his shopping, no participants from group 2 selected this option. On the contrary, all female participants from group 2 suggested Peter to use the Alter-Ego to assist his shopping.

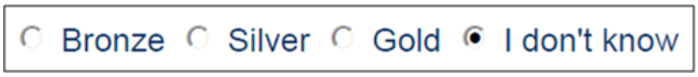
7.9.2 Recommending the PPSE - Part 2

As introduced previously, the majority of participants recommended the scenario's persona, the use of the PPSE. However, to expand the answer, participants were asked about disclosure level:

Question:

If you were Peter and were using the Alter-Ego, what level would you use?

Answer options:



○ Bronze ○ Silver ○ Gold ● I don't know

The answers are shown in Figure 7.25 and summarised in Table 7.16.

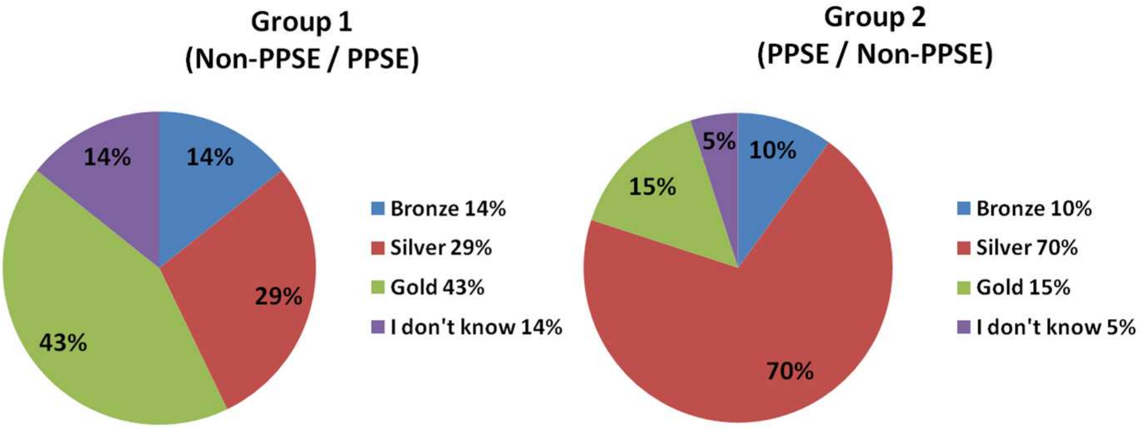


Figure 7.25: Final questionnaire - Recommending disclosure levels to *Peter*

.

Group 1				
	Bronze	Silver	Gold	Don't know
Total	14 %	29 %	43 %	14 %
Age				
< 33	15 %	31 %	46 %	8 %
=> 33	12 %	25 %	38 %	25 %
Gender				
Male	28 %	27 %	27 %	18 %
Female	0 %	30 %	60 %	10 %
Internet Expertise				
Expert	25 %	25 %	25 %	25 %
Intermediate	0 %	40 %	50 %	10 %
Novice	33 %	0 %	67 %	0 %
Group 2				
	Bronze	Silver	Gold	Don't know
Total	10 %	70 %	15 %	5 %
Age				
< 33	13 %	67 %	13 %	7 %
=> 33	0 %	80 %	20 %	0 %
Gender				
Male	12 %	75 %	13 %	0 %
Female	8 %	67 %	17 %	8 %
Internet Expertise				
Expert	11 %	78 %	11 %	0 %
Intermediate	14 %	58 %	14 %	14 %
Novice	0 %	75 %	25 %	0 %

Table 7.16: Final questionnaire - Recommending disclosure levels to *Peter*

The frequency of "*Bronze*, *Silver*, *Gold* and *Don't know*" from group 1 and group 2 is not statistically different (χ^2 = 12.083; P=0.209). Therefore the order of undertaking the scenario did not affect the participant's decision of recommending a particular disclosure level to "*Peter*". It is important to note that whether the options selected from group 1 were (in descendant order) gold, silver and bronze, for group 2 the options selected were (in descendant order) silver, gold and bronze. The only change is registered in male participants from group 1 where the options selected (in descendant order) were bronze silver gold.

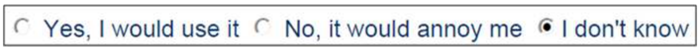
7.9.3 Participant's satisfaction

Since the use of the PPSE involves an extra step in the traditional shopping, the objective of this question was to evaluate if participants were satisfied enough with the use of the PPSE to use it. The question used to perceive the participant's satisfaction was:

Question:

> *Doing your shopping assisted by the Alter-Ego represents an extra step in everyday shopping. Do you think that this extra step to maintain your privacy would be warranted?*

Answer options:



○ Yes, I would use it ○ No, it would annoy me ● I don't know

The responses are shown in Figure 7.26, and summarised in Table 7.17.

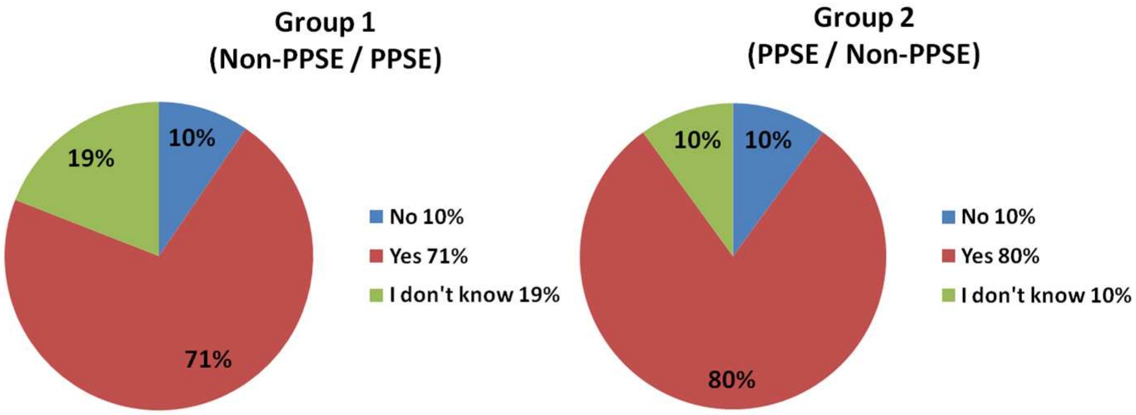


Figure 7.26: Final questionnaire - Customer satisfaction

.

Group 1			
	No	Yes	Don't know
Total	10 %	71 %	19 %
Age			
< 33	8%	77 %	15 %
=> 33	12 %	63 %	25 %
Gender			
Male	18 %	64 %	18 %
Female	0 %	80 %	20 %
Internet Expertise			
Expert	25 %	50 %	25 %
Intermediate	0 %	80 %	20 %
Novice	0 %	100 %	0 %
Group 2			
	No	Yes	Don't know
Total	10 %	80 %	10 %
Age			
< 33	6 %	87 %	7 %
=> 33	20 %	60 %	20 %
Gender			
Male	12 %	75 %	13 %
Female	9 %	83 %	8 %
Internet Expertise			
Expert	11 %	78 %	11 %
Intermediate	0 %	86 %	14 %
Novice	25 %	75 %	0 %

Table 7.17: Final questionnaire - Customer satisfaction

The frequency of "*No*, *Yes* and *Don't know*" from group 1 and group 2 is not statistically different ($\chi^2 = 2.143$; P=0.710). Therefore the order of undertaking the scenario did not affect the participant's satisfaction. It is important to note that all participants from group 1 that considered their computer expertise to be novice, majority of females than males, considered that the effort of using the Alter-Ego in their everyday shopping was warranted.

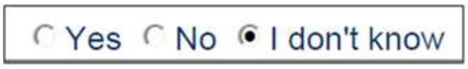
7.9.4 Participant's customer loyalty

In the last section of the experiment to support the thesis statement, the encouragement of customer loyalty is explored. To evaluate the participants' satisfaction as an indicator of customer loyalty. The following question was presented:

Question:

Would you use it?

Answer options:



○ Yes ○ No ● I don't know

The results are shown in Figure 7.27 and summarised in Table 7.18.

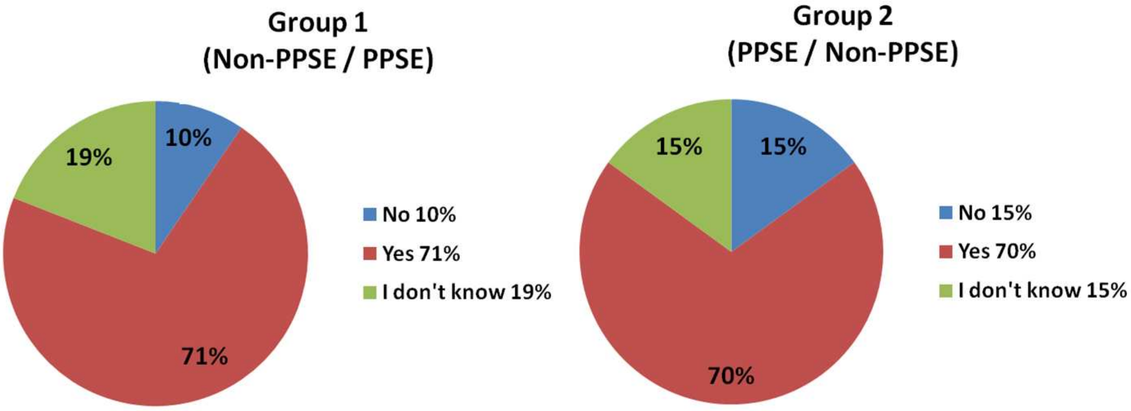


Figure 7.27: Final questionnaire - Customer loyalty / satisfaction

.

Group 1			
	No	Yes	Don't know
Total	10 %	71 %	19 %
Age			
< 33	8%	69 %	23 %
$=> 33$	13 %	75 %	12 %
Gender			
Male	0 %	73 %	27 %
Female	20 %	70 %	10%
Internet Expertise			
Expert	12 %	63 %	25 %
Intermediate	10 %	70 %	20 %
Novice	0 %	100 %	0 %
Group 2			
	No	Yes	Don't know
Total	15 %	70 %	15 %
Age			
< 33	13 %	67 %	20 %
$=> 33$	20 %	80 %	0 %
Gender			
Male	25 %	50 %	25 %
Female	9 %	83 %	8%
Internet Expertise			
Expert	22 %	67 %	11 %
Intermediate	14 %	57 %	29 %
Novice	0 %	100 %	0 %

Table 7.18: Final questionnaire - Customer loyalty

The frequency of "*No*, *Yes* and *Don't know*" from group 1 and group 2 is not statistically different ($\chi^2 = 1.633$; P=0.803). Therefore the order of undertaking the scenario did not affect the participant's decision of using the PPSE. It is important to notice that all participants from group 1 and 2 that considered their computer expertise as novice would use the Alter-Ego. In group 2, however, a larger group of males selected the options "*No* and *Don't know*" than the females from group 2 selecting the same options.

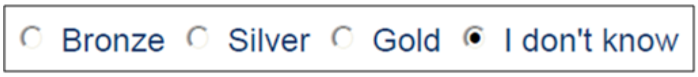
7.9.5 Participant's use of PPSE

As shown in Table 7.18, the majority of participants would use the PPSE. However, to expand the answer, participants were asked about the disclosure level:

Question:

Which level would you generally shop at?

Answer options:



○ Bronze ○ Silver ○ Gold ● I don't know

The results are shown in Figure 7.28 and summarised in Table 7.19.

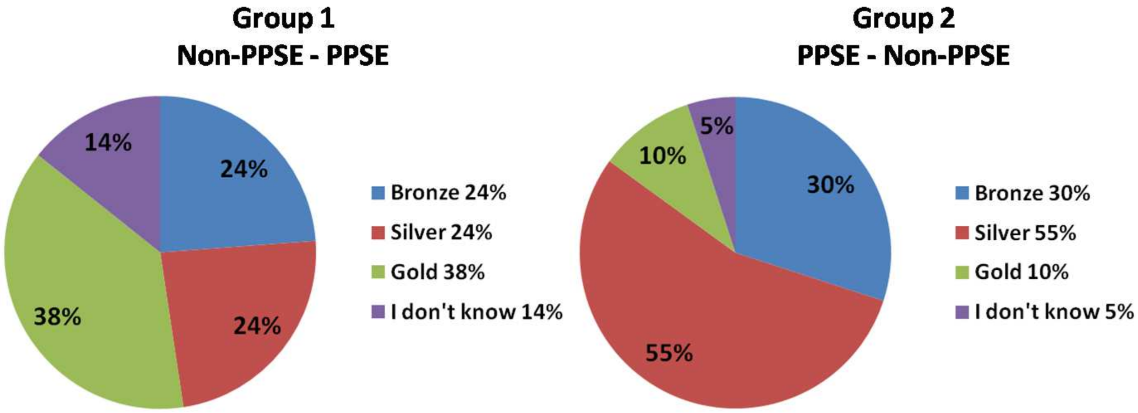


Figure 7.28: Final questionnaire - Participant's disclosure level

.

Group 1				
	Bronze	Silver	Gold	I don't know
Total	24 %	24 %	38 %	14 %
Age				
< 33	15 %	31 %	46 %	8 %
=> 33	12 %	25 %	38 %	25 %
Gender				
Male	36 %	27 %	37 %	0 %
Female	10 %	20 %	40 %	30 %
Internet Expertise				
Expert	25 %	25 %	38 %	12 %
Intermediate	20 %	30 %	30 %	20 %
Novice	33 %	0 %	67 %	0 %
Group 2				
	Bronze	Silver	Gold	I don't know
Total	30 %	55 %	10 %	5 %
Age				
< 33	13 %	67 %	13 %	7 %
=> 33	0 %	80 %	20 %	0 %
Gender				
Male	50 %	50 %	0 %	0 %
Female	17 %	58 %	17 %	8 %
Internet Expertise				
Expert	44 %	56 %	0 %	0 %
Intermediate	14 %	58 %	14 %	14 %
Novice	25 %	50 %	25 %	0 %

Table 7.19: Participant's disclosure level

The frequency of "*Bronze*, *Silver*, *Gold* and *Don't know*" from group 1 and group 2 is not statistically different (χ^2 = 9.38; P=0.403). Therefore the order of undertaking the scenario did not affect the participant's decision of using a particular disclosure level. It is important to notice that whether the options selected from group 1 were (in descendant order) gold, silver and bronze, for group 2 the options selected were (in descendant order) silver, bronze and gold. The only change is registered in male participants from group 1 where the options selected (in descendant order) were gold bronze and silver.

7.10 Discussion

7.10.1 Privacy perceptions - Based on privacy categories

With the aim of determining the participants' privacy needs, the results of the first questionnaire gathered the participant's perspective of privacy, related to the three parts of the privacy definition ("*control over disclosure*", "*control over body / personal information*" and "*the right to be left alone (set boundaries)*").

The results of the first question "*control over disclosure*", presented in Table 7.1, showed that, while the majority of participants decided not to face a risk situation, there was a considerable number of participants that decided to take the risky situation based on the scenario criteria (tiredness and price). This result suggests that, under circumstances that involve a calculated risk, participants would disclose their information. However, participants that considered themselves as Internet novice users did not follow the tendency, the majority of novices would buy the ticket instead of paying more with a reputable company. This results suggests that, for *control over disclosure*, novice users' privacy perception towards online shopping is more trusting than experts or intermediate's privacy perception.

In the case of "*control over body / personal information*", participants faced a situation where their information was already out of their control. The results shown in Table 7.2, suggest that, the loss of the control and misuse of their information has a major negative impact in their perception of privacy regardless of age. However, novice Internet users and females place less importance in the disclosure of information.

Finally, in this questionnaire's last question evaluating "*the right to be left alone (set boundaries)*", the high incidence of the "*Don't care*" option, shown in Table 7.3, suggest that, whether or not participants are concerned with establishing their personal boundaries, their reaction to a relative invasion of those boundaries does not represent a major negative effect.

Therefore, from this questionnaire it can be concluded that the participants privacy needs are, up to certain extent, flexible in the setting of their privacy boundaries. Under certain circumstances, some of them consider facing risk situations, but they do not tolerate the loss of control or misuse of their information. These results support hypothesis 1 (*People have privacy needs*).

7.10.2 Privacy perceptions - Based on privacy violations

In the *privacy perceptions according to privacy violations* questionnaire, participants identified the options that they considered violated their privacy. These options were linked to the three parts of the privacy definition (‘*control over disclosure*”, “*control over body / personal information*” and “*the right to be left alone (set boundaries)*”).

According to the association between the perceived privacy violations, and Westin’s privacy indices (*fundamentalist, pragmatic and unconcerned*), Section 7.7.4, the results for the category *control over disclosure*, presented in Table 7.4, showed that the majority of participants belonged to the pragmatic category followed by fundamentalist and unconcerned. In the category *control over body / personal information*, presented in Table 7.5, the majority of participants belonged to the pragmatic category, followed by fundamentalists and unconcerned. However, the category *the right to be left alone (set boundaries)*, presented in Table 7.6, the majority, pragmatic, was not followed by fundamentalists, but it was followed by the unconcerned group.

This shift in the distribution suggests that whether participants are conscious of their privacy needs and have a practical open-minded approach to privacy preserving mechanisms, they do not place the same importance when setting boundaries, and do not consider the interaction with others, and the delimitation of boundaries as vital as the disclosure of their information. These results prove that, where as participants have privacy needs, they place a different value in the different aspects of privacy. This finding also supports the hypothesis that people have privacy needs (1).

7.10.3 Privacy perceptions - Based on awareness

From the results obtain in the question presented to determine the participants’ privacy awareness, presented in Table 7.7, it can be noticed the lack of personal experience in privacy violations. This suggests that the participants’ privacy perception in the majority of the participants has not been created by firsthand experiences.

7.10.4 Privacy perceptions - Non-PPSE

Questionnaire A was presented to the participants at the end of their performing the bshop tasks in the non-PPSE evaluation environment. The responses from group 1 and group 2 in regard with “*control over disclosure*”, concurred on the general, with the majority of participants, as presented in Table 7.8, selecting the option *'Very angry'*. The results from group 1 and group 2 were not statistically different, however, group 1 presented a stronger negative reaction to the control over disclosure option after being presented with the privacy violation message.

In the first question (evaluating *control over disclosure*), only participants from group 2 selected the option *'Don't care'*, this suggests that all participants from group 1 did care about not being able to explain the context of the shopping (not for *Peter*, but for a Christmas party), after the violation message. This results suggest that participants react negatively when they become aware of a privacy violation out of their specific control, without being aware of the existence of other means of protecting their privacy.

In the case of the question directed towards the "*control over body / personal information*", the majority of participants selected *'Very angry'*, as presented in Table 7.9. In this case, the results from group 1 were also bigger than the results from group 2 without being statistically different. This suggests that participants that were faced first with a scenario where they were not able to control the disclosure of their body/personal information had a stronger negative perception than those who had used first a scenario with a privacy preserving mechanism.

In the results obtained for "*the right to be left alone (set boundaries)*", presented in Table 7.10, neither group selected the option *'Don't know'*. This results suggests that for both groups their privacy perspective was, up to certain extent, more flexible for "*control over disclosure*" and "*control over body / personal information*" than their privacy perspective when they realised that they agreed by default to the e-store's facility of disclosing the information to third parties. This suggests that participants had an initial expectancy of the information that the store would collect, use and disclose. The realisation that the e-store's had different objectives for the collected information was a cause for discontent. It is important to note that, whereas the prototype of the store had a *term and conditions* section, not a single participant read it before, during or after the test.

7.10.5 Privacy perceptions - PPSE

Questionnaire B was presented to the participants at the end of performing the bshop tasks in the PPSE evaluation environment. In the question to obtain the privacy perceptions regarding "*control over disclosure*", the majority of the participant's opinion, as presented in Table 7.11, was *'Very relieved'.* However, while the results were not statistically different, the response from group 2 was more positive than the response from group 1. This suggests that while group 2 gave a positive response towards the benefits of the PPSE (this was their first scenario), a certain level of awareness was created in group 1 and their answer was more conservative.

In the question related to "*control over body / personal information*", the results, presented in Table 7.12, showed whereas the majority of participants in both groups selected the option *'Happy'*, it was selected by more participants from group 2 than group 1. This suggests that while there is no statistically difference, participants from group 2 appreciated having their preferences and sensitive information in a repository where they could manage them.

In the question related to "*the right to be left alone (set boundaries)*", the results, presented in Table 7.13, showed that the participants' opinion was positive, although the difference between groups was not statistically significant. In this case, more participants from group 2 selected the option *'Happy'* than participants from group 1. This suggests that participants from group 2 took for granted that their information was not shared in an unauthorised way.

The last question in this section explored the participants's perception over controlling their privacy with the PPSE. While the majority of both groups selected *'Some control'*, as presented in Table 7.14, a bigger number of participants from group 1 selected *'Absolute control'.* This suggests that the appreciation of the PPSE from participants that have not experienced a firsthand privacy violation is not as positive as participants that have experienced firsthand privacy violations.

7.10.6 Customers' satisfaction and loyalty

The final questionnaire aimed to determine the participants' satisfaction while using the PPSE and if they considered to use it again (loyalty). From the results obtained in the first question, the majority of participants in both groups recommended the use of the PPSE in case of customers with privacy needs. While the results, presented in Table 7.15, showed no statistically significative differences, more participants from group 2 indicated that they would recommend Peter the use of the PPSE. This more conservative opinion from group 1 suggests that the raise of awareness had influenced the change in the participants perception of privacy.

In the case of participants' satisfaction, when participants were asked about their perception of the PPSE preserving their privacy, the responses of majority of participants was positive, as shown in Figure 7.26, supporting the hypothesis 2 (the PPSE would satisfy the peoples' privacy needs).

Two questions were presented to obtain details of the disclosure level that they would suggest *Peter* to use, and that they would use. The responses, presented in Tables 7.16 and 7.19, showed that the majority of participants of group 1 would use and would suggest *Peter* to use Gold disclosure level while participants in group 2 would use and suggest *Peter* to use Silver disclosure level. This suggests that participants from group 1 selected to use the disclosure level designed for the group with pragmatic privacy concerns and group 2 selected to use the disclosure level designed for the unconcerned group. Whereas the differences of recommending one level or the other are not statistically significative in relation with the order of undertaking the scenarios, these results suggest that participants find differences in the uses of the disclosure levels and are satisfied with using them.

This willingness to use the disclosure levels, together with the participants positive response when asked if they would use the PPSE, presented in Table 7.18, support the hypothesis 3, users were satisfied with the PPSE.

7.11 Conclusions

This chapter presented the questionnaires used in the evaluation of customer' satisfaction and loyalty towards the PPSE to support the second part of the thesis statement.

The hypothesis 1 (participants have privacy needs) was supported based on the participants' perception over their privacy in relation with the privacy definition and privacy violations. Participants reported that, using the PPSE, they felt in control of their privacy, supporting hypothesis 2 (the PPSE can satisfy these privacy needs). Finally, hypothesis 3 (users are satisfied with the PPSE) was supported based on the positive responses about the participants' satisfaction in the use of the PPSE, their willingness of using the PPSE (customer loyalty) and recommending it to people with specific privacy needs.

The results obtained from groups 1 and 2 were not statistic different, indicating indicates that the order of undertaking the scenarios with a privacy preserved environment and with an environment that did not preserve their privacy had no effect in the participants privacy perceptions and their acceptance of the PPSE. However, the tendency suggested that the awareness created early in group 1 by presenting them with a privacy violation scenario, resulted in their more conservative acceptance of the PPSE.

It can be concluded then, that by presenting the results of the user test, the validation of the thesis statement is completed. Next section concludes this dissertation presenting final conclusions and suggesting future work.

Chapter 8

Conclusions

New violations and threats to privacy occur on an everyday basis. Organisations and media attempt to raise privacy awareness, but unfortunately, in some cases it is only when privacy is compromised and recovery attempted that the reality of the importance of *privacy matters* dawns.

The Web has provided improved and facilitated access mechanisms that were previously non-existent to various activities for instance, e-commerce provides improved access to making purchases. As a business, e-commerce, as explored in Chapter 2, has experienced a steady growth after the bubble burst effect in 2001. That growth and consequent fierce competition has encouraged e-tailers to adopt techniques, such as personalisation, to collect and analyse data in order to increase profits. However, the indiscriminate collection of information and potential misinterpretation of analysed data caused personalisation to be perceived as a privacy risk.

Preserving privacy has been the aim of a series of efforts analysed in Chapter 4. However, their use has been less than effective so far. One possible reason could be that the approach taken towards the preservation of privacy involves only one or, at best, two of the following approaches (presented in Section 4.4): raising *awareness*, *regulation* and the use of *technology* (ART).

The privacy preserving shopping environment (PPSE) was proposed here as a more holistic approach, since it combines all three ART aspects. A prototype of the PPSE was designed and implemented to support the evaluation of the first part of the thesis statement ("*It is possible to develop a privacy preserving shopping environment (PPSE), which respects the customers' privacy needs while allowing the company to gather and use sufficient reliable customer-specified data to achieve a level of personalisation*"). A user test was carried out to support the second part of the thesis statement (*"which can be used to encourage customer loyalty"*). The following hypotheses were tested during the user test:

- *hypothesis 1* - people have privacy needs,
- *hypothesis 2* - the PPSE can satisfy these privacy needs, and,
- *hypothesis 3* - users were satisfied with the PPSE.

In the user test, two scenarios involving the use of a privacy-preserved shopping environment (PPSE) and a non-privacy-preserved shopping environment (non-PPSE) were implemented. Although both scenarios involved the use of an e-grocery store, the non-PPSE scenario presented the participants with a message letting them know that the privacy of *Peter* (a "persona" with specific privacy issues) was violated. In the PPSE scenario, after using the e-grocery shop, participants were presented with a message informing them that their privacy was preserved. The findings of the user test can be summarised as follows:

Hypothesis 1 - People have privacy needs .

- The first three questionnaires explored the participants' perceptions of privacy with the following findings:
  - A small number of participants have experienced privacy violations.
  - The majority of participants exhibited a pragmatic approach towards the three aspects of the privacy definition ("*control over disclosure*", "*control over body / personal information*" and "*the right to be left alone (set boundaries)*")
  - Regarding *control over disclosure* and *control over body / personal information* fundamentalist group was bigger than unconcern group.
  - In the case of *the right to be left alone (set boundaries)*, unconcerned group was bigger than fundamentalist group.
  - From the scenarios used to directly assess the privacy definition, a number of participants who were willing to take risks under controlled circumstances (*control over disclosure*), reacted negatively when the information was disclosed and misused without their consent (*control over body / personal information*), and had a relatively flexible tolerance towards invasion of their space (*right to be left alone (set boundaries)*).

- The questionnaires presented during the PPSE evaluation delivered the following findings:
  - The participants' opinions of the use of a store that presented a message informing them about a misuse of their information was strongly negative. However, the response was stronger in the group that used the Non-PPSE scenario first.
  - The positive reception accorded the PPSE from participants of the group that used the PPSE first was bigger than from participants who used the non-PPSE first. This suggests acceptance of the PPSE, and expected heightened caution from the group that had faced a privacy violation.

Hypothesis 2 - The PPSE can satisfy these privacy needs .

- The majority of participants reported that, from their perspective, they were in control of their privacy (and their identified privacy needs based on the three elements of the privacy definition) when using the PPSE.

Hypothesis 3 - Users were satisfied with the PPSE .

- Participants indicated that they would recommend the use of the PPSE to people with privacy concerns.
- Based on those privacy concerns, participants suggested the use of *Silver* (designed for unconcerned users) or *Gold* (designed for pragmatist users) disclosure levels.
- Despite the fact that the PPSE introduced an extra step in the shopping process, the use thereof was perceived to be warranted.
- The majority of participants indicated that they would use it if it were available.

In conclusion, the user test supported the three hypotheses completing the validation of the thesis statement.

8.1 Future work

Due to the links that privacy has with multiple areas, such as consumer, e-tailers, and privacy organisations, any research carried out to preserve privacy have repercussions in a number of applications. This research leads into the following future work.

Raising of Awareness -

- A joint approach involving privacy organisations (such as Privacy International) could prove beneficial in researching better ways of raising awareness using the PPSE.

Regulation -

- A controlled environment allows a closer control over regulation, its use and evolution. The PPSE can be used as a practical case to explore different implementations of initiatives such as P3P.

Technology -

- Personalisation methods and recommendation lists can be explored in the PPSE stores.
- Since the Alter-Ego portal allows the entrance of new search preferences, the use of semantic dictionaries based on the terminology used by the stores in their stock classification, could limit and assist the universe of available entrances.
- Since the checkout process requires the disclosure of the client's personal information, alternative payment and delivery methods, such as PayPal, could be explored to guarantee full anonymity.

To industrialise the PPSE, the following stakeholders should be considered and consulted so as to satisfy all their needs:

- Privacy organisations (i.e. Privacy International)
- Customer groups (i.e. National Consumer Federation)
- Regulatory organisations (i.e. Better Regulation Commission[1])
- E-tailers (i.e. Electronic Retailing Association[2])
- Technology providers
- Researchers

[1] http://archive.cabinetoffice.gov.uk/brc/index.html

[2] http://www.retailing.org/

8.2 A final word

This research proposed an environment that gives customers the capability of carrying out their e-shopping in a privacy-preserved environment which:

- aims to raise their privacy awareness,
- facilitates the disclosure of their desired amount of information,
- provides a space to store their preferences and sensitive information so they can be used in any participant store to give them a personalised shopping experience without the need of spending time building a profile, and that
- abides to laws and regulations that protect privacy.

This environment, presented in the thesis statement, was designed, implemented and, finally, supported by the big majority of the participants who undertook a user satisfaction and customer loyalty test. These findings provide us with elements to conclude that personalised privacy preserved e-shopping is both feasible and desired.

Bibliography

[1] A. Abdelnur and S. Hepper. Java Portlet^{TM} Specification, Version 1.0. http://jcp.org/en/jsr/detail?id=168. Accessed 06 July 2008.

[2] G. Adomavicius and A. Tuzhilin. Personalization technologies: a process-oriented perspective. *Commun. ACM*, 48(10):83–90, 2005.

[3] J. Agar. Identity cards in britain: past experience and policy implications. http://www.historyandpolicy.org/archive/pol-paper-33.html, 2005. Accessed 12 Sept 2006.

[4] J. Alexander. Confidentiality and privacy: what's the difference? http://www.library.cmu.edu/ethics2.html, 2004. Accessed 10 Sept 2006.

[5] APACS. Number of people banking online increases more than 500% in past seven years. http://www.apacs.org.uk/08_07_24.htm. Accessed 07 August 2008.

[6] AT&T Corp. Privacy bird ®. http://www.privacybird.org/. Accessed 28 July 2007.

[7] B2B Marketing online. B2B Marketing online. http://www.b2bm.biz/. Accessed 01 August 2008.

[8] M. Balabanović and Y. Shoham. Fab: content-based, collaborative recommendation. *Commun. ACM*, 40(3):66–72, 1997.

[9] R. Baraglia and F. Silvestri. Dynamic personalization of web sites without user intervention. *Commun. ACM*, 50(2):63–67, 2007.

[10] M. Barbaro and T. Zeller. A face is exposed for aol searcher no. 4417749. http://www.nytimes.com/2006/08/09/technology/09aol.html?ex=1312776000&en=f6f61949c6da4d38&ei=5090, 2006. Accessed 29 August 2006.

[11] R. Barrett, P. P. Maglio, and D. C. Kellem. A confederation of agents that personalize the web. http://www.almaden.ibm.com/cs/wbi/papers/agents97.pdf, 1997. Accessed 29 June 2005.

[12] R. Barrett, P. P. Maglio, and D. C. Kellem. WBI: a confederation of agents that personalize the Web. In *AGENTS '97: Proceedings of the first international conference on Autonomous agents*, pages 496–499, New York, NY, USA, 1997. ACM Press.

[13] BBC News. Q&a: Identity card plans. http://news.bbc.co.uk/1/hi/uk_politics/3127696.stm, 2008. Accessed 5 March 2008.

[14] BBC News. Q&a: Loss of freedoms? http://news.bbc.co.uk/1/hi/uk/7451552.stm, 2008. Accessed 28 October 2008.

[15] BBC News. Record increase in DNA database. http://news.bbc.co.uk/1/hi/uk/7662683.stm, 2008. Accessed 28 October 2008.

[16] B. Berendt, O. Günther, and S. Spiekermann. Privacy in e-commerce: stated preferences vs. actual behavior. *Commun. ACM*, 48(4):101–106, 2005.

[17] N. Bevan. Measuring usability as quality of use. *Software Quality Journal*, 4(2):115–130, 1995.

[18] P. BeynonDavies. *E-business.* Pilgrave MacMillan, 2004.

[19] C. Buyer. Japan's marubeni to test smart tags in retail stores. html://crmbuyer.com/story/46031.html, 2005. Accessed 14 September 2005.

[20] J. Cassidy. *Dot.con : the greatest story ever sold.* London : Allen Lane, 2002.

[21] S. Chang, S. Changchien, and R. Huang. Assessing users' product-specific knowledge for personalization in electronic commerce. *Expert Systems With Applications*, 30(4):682–693, 2006.

[22] S. C. Chen and G. S. Dhillon. Interpreting Dimensions of Consumer Trust in E-Commerce. *Information Technology and Management*, 4(2):303–318, 2003.

[23] H.-C. Chu, M.-Y. Chen, and Y.-M. Chen. A semantic-based approach to content abstraction and annotation for content management. *Expert Systems With Applications*, 2008.

[24] CIFAS. What is CIFAS? http://www.cifas.org.uk/. Accessed 07 August 2008.

[25] R. Clarke. Human Identification in Information Systems: Management Challenges and Public Policy Issues. *Information Technology & People*, 7(4):6–37, 1994.

[26] C. Clifton and J. Vaidya. Privacy-preserving data mining: Why, how, and when. *IEEE Security and Privacy*, 2(6):19–27, 2004.

[27] S. Collett. Outsourcing: Losing control. `http://www.computerworld.com/printthis/2004/0,4814,91085,00.html`, 15 March 2004. Accessed 27 May 2005.

[28] R. M. Conlan and P. Tarasewich. Improving interface designs to help users choose better passwords. *Conference on Human Factors in Computing Systems*, pages 652–657, 2006.

[29] ContactCenterWorld.com. Global personalization. `http://www.contactcenterworld.com/research.asp?\#Global\%20Personalization\%20Technologies\%20-\%20Datamonitor`, 2006. Accessed 02 April 2006.

[30] G. Conti and E. Sobiesk. An honest man has nothing to fear: user perceptions on web-based information disclosure. *Proceedings of the 3rd symposium on Usable privacy and security*, pages 112–121, 2007.

[31] L. Cranor. 'I didn't buy it for myself' privacy and ecommerce personalization. *Proceedings of the 2003 ACM workshop on Privacy in the electronic society*, 2003.

[32] L. Cranor, M. Arjula, and P. Guduru. Use of a p3p user agent by early adopters. In *WPES 0́2: Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society*, pages 1–10, New York, NY, USA, 2002. ACM.

[33] L. Cranor, B. Dobbs, S. Egelman, G. Hogben, J. Humphrey, M. Langheinrich, M. Marchiori, M. Presler-Marshall, J. Reagle, M. Schunter, D. A. Stampley, and R. Wenning. The platform for privacy preferences 1.1 (p3p1.1) specification. `http://www.w3.org/TR/2006/WD-P3P11-20060210/Overview.html`, February 2006. Accessed 11July2006.

[34] Critical Infrastructure Assurance OfficeCIAO and et al. British standard 7799. `http://www.knowledgeleader.com/KnowledgeLeader/Content.nsf/content/SecurityBritishStandard7799ISO17799?OpenDocument`, 2006. Accessed 11 Sept 2006.

[35] S. Devaraj, M. Fan, and R. Kohli. E-loyalty: elusive ideal or competitive edge? *Communications of the ACM*, 46, 2003.

[36] R. Dilley. When the british fought off id cards. `http://news.bbc.co.uk/1/hi/magazine/3129302.stm`, 2003. Accessed 12 Sept 2006.

[37] S. Eads. The Web's Still-Unfulfilled Personalization Promise. `http://www.businessweek.com/bwdaily/dnflash/aug2000/nf2000084\_506.htm`. Accessed 27 May 2005.

[38] M. Eirinaki and M. Vazirgiannis. Web mining for web personalization. *ACM Transactions on Internet Technology (TOIT)*, 3(1):1–27, 2003.

[39] T. R. Eisenmann and S. Pothen. Online Portals, Case Number 9-801-305. Harvard Business School, 2000.

[40] Electronic Privacy Information Center and Privacy International. Privacy and Human Rights 2003: Overview. `http://www.privacyinternational.org/survey/phr2003/overview.htm`, 2003. Accessed 23 Oct 2007.

[41] V. Elliot. 'intelligent' shopping trolley is new front in battle against obesity. Newspaper, The Times, 2007. October 8, Page 27, Section News.

[42] EMarketer. 2005 will be the year of personalization. `http://www.marketwire.com/mw/release/_html/_bl?release/_id=78581`, 2005. Accessed 29 June 2005.

[43] A. Field. *Discovering Statistics Using SPSS.* Sage Publications Inc, 2005.

[44] A. Field and G. Hole. *How to Design and Report Experiments.* Sage Publications, 2003.

[45] K. Foord. Defining privacy. Victorian Law Reform Commission, 2000.

[46] D. Galvez-Cruz and K. Renaud. Privacy by agreement. *IADIS E-commerce*, pages 338–341, 2006.

[47] D. Galvez-Cruz and K. Renaud. You Know My'Alter-Ego'–but You Don't Know Me! In *Databases, 2007. BNCOD'07. 24th British National Conference on*, pages 101–109, 2007.

[48] D. Galvez-Cruz and K. V. Renaud. What e-grocery customers really want: Personalised personalisation. In *LA-WEB '06: Proceedings of the Fourth Latin American Web Congress*, pages 109–112, Washington, DC, USA, 2006. IEEE Computer Society.

[49] H. Garstka. Europe's 'privacy cops': The U.S. isn't our beat (int'l edition). `http://www.businessweek.com/1998/49/b3607178.htm`, 1998. Accessed 29 June 2005.

[50] J. M. Germain. Online consumers window shop more than impulse buy. `http://www.ecommercetimes.com/story/42761.html`, 2005. Accessed 29 June 2005.

[51] Gerti Kappel (Editor) and Birgit Pröll (Editor) and Siegfried Reich (Editor) and Werner Retschitzegger (Editor). *Web engineering : the discipline of systematic development of web applications.* Chichester ; Hoboken, NJ : John Wiley & Sons, 2006.

[52] M. Gifkins and D. Hitchcock. *The EDI handbook : trading in the 1990s.* Blenheim Online, 1988.

[53] C. Goodwin. Privacy: Recognition of a consumer right. *Journal of Public Policy & Marketing*, 10(1):149–166, 18p, Spring91.

[54] Google. About us. `http://www.ebay.co.uk/`. Accessed 01 August 2008.

[55] D. Hajewski. Kohl's, northwestern mutual top customer satisfaction survey. `http://www.jsonline.com/bym/news/feb05/301759.asp`, 2005. Accessed 29 June 2005.

[56] C. E. Heckman and J. O. Wobbrock. Put your best face forward: anthropomorphic agents, e-commerce consumers, and the law. In *AGENTS '00: Proceedings of the fourth international conference on Autonomous agents*, pages 435–442, New York, NY, USA, 2000. ACM Press.

[57] R. D. Hof and L. Himelstein. Now it's your web. `http://www.businessweek.com/1998/40/b3598023.htm`, 1998. Accessed 29 June 2005.

[58] L. J. Hoffman. Computers and privacy: A survey. *ACM Computing Surveys (CSUR)*, 1, 1969.

[59] A. Holzinger. Usability engineering methods for software developers. *Communications of the ACM*, 48(1):71–74, 2005.

[60] International Biometric Group, LLC. IBG BioPrivacy Initiative $^{\text{TM}}$. `http://www.bioprivacy.org/`. Accessed 26 August 2008.

[61] International Telecommunication Union. X.800 : Security architecture for open systems interconnection for ccitt applications. `http://www.itu.int/rec/T-REC-X.800/en`, 1991. Accessed 11 March 2008.

[62] International Telecommunication Union. X.805 : Security architecture for systems providing end-to-end communications. `http://www.itu.int/rec/T-REC-X.800/en`, 2003. Accessed 11 March 2008.

[63] Internet Retailer. Web personalization will come back and change retailing, experts say. `http://internetretailer.com/dailyNews.asp?id=13879`, 2005. Accessed 29 June 2005.

[64] A. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):4–20, 2004.

[65] Y.-J. Joung, C. Yen, C.-T. Huang, and Y.-J. Huang. On personal data license design and negotiation. In *Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International*, volume 1, pages 281–286 Vol 2, 2005.

[66] T. Kamba and Y. Koike. Presentation of personalized information using anthropomorphous agents. In *CHI '99: CHI '99 extended abstracts on Human factors in computing systems*, pages 246–247, New York, NY, USA, 1999. ACM Press.

[67] B. Kasanoff. *Make it Personal how to profit from personalization without invading privacy.* Perseus Publishing, 1 edition, 2001.

[68] M. Kempiak and M. A. Fox. Online Grocery Shopping: Consumer Motives, Concerns, and Business Models. *First Monday*, 7(9), 2002.

[69] A. Kobsa, R. K. Chellappa, and S. Spiekermann. Privacy-enhanced personalization. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 1631–1634, New York, NY, USA, 2006. ACM.

[70] S. Korper and J. Ellis. *The E-Commerce Book: Building the E-Empire.* Morgan Kaufmann, 2001.

[71] V. Kotelnikov. Business Model. `http://www.1000ventures.com/business\_guide/business\_model.html#Xerox`. Accessed 14 July 2008.

[72] V. Kotelnikov. Case Study:Amazon.com New Business Model and Venture Financing Chronology. `http://www.1000ventures.com/business\_guide/cs\_biz\_model\_amazon.html`. Accessed 13 July 2008.

[73] B. Krishnamurthy, D. Malandrino, and C. E. Wills. Measuring privacy loss and the impact of privacy protection in web browsing. In *SOUPS '07: Proceedings of the 3rd symposium on Usable privacy and security*, pages 52–63, New York, NY, USA, 2007. ACM.

[74] D. M. Kristol. HTTP Cookies: Standards, privacy, and politics. *ACM Transactions on Internet Technology (TOIT)*, 1(2):151–198, 2001.

[75] P. Kumaraguru and L. Cranor. Privacy Indexes: A Survey of Westins Studies. *Institute for Software Research International*, 2005.

[76] B. Lavoie and H. F. Nielsen. Web characterization terminology & definitions sheet. `http://www.w3.org/1999/05/WCA-terms/\#WEBSCOPE`, May 1999. Accessed 06 August 2008.

[77] J. LeClaire. Study Shows Where Multi-Channel E-Tailers Must Improve. `http://www.ecommercetimes.com/story/39929.html`. Accessed 29 June 2005.

[78] J. LeClaire. Sun Microsystems' Mike Green Analyzes E-Commerce in 2005. `http://www.ecommercetimes.com/story/40100.html`, 2005. Accessed 29 June 2005.

[79] P. Legris, J. Ingham, and P. Collerette. Why do people use information technology. *A critical review of the technology acceptance model. Information & Management*, 40(3):191–204, 2003.

[80] B. M. Leiner, V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, and S. Wolff. A brief history of the internet. `http://www.isoc.org/internet/history/brief.shtml`. Accessed 30 April 2007.

[81] J. Lewis. All air passengers to give their fingerprints ... but is the reason security or simply to raise profits for the duty-free shops? `http://www.dailymail.co.uk/news/article-1038879/All-air-passengers-fingerprints-reason.-security-simply-raise-profits-duty-free-shops.html`. Accessed 26 August 2008.

[82] C. Lin, S. Wu, and R. J. Tsai. Integrating perceived playfulness into expectation-confirmation model for web portal context. *Information & Management*, 42(5):683–693, 2005.

[83] J. Ling and P. van Schaik. The effect of text and background colour on visual search of Web pages. *Displays*, 23(5):223–230, 2002.

[84] Lorrie Cranor (Chair) and Rigo Wenning (W3C). Platform for Privacy Preferences (P3P) Project. `http://www.w3.org/P3P/`. Accessed 28 July 2008.

[85] S. E. Lunce, L. M. Lunce, Y. Kawai, and B. Maniam. Success and failure of pure-play organizations: Webvan versus Peapod, a comparative analysis. *Industrial Management & Data Systems*, 106(9):1344–1358, 2006.

[86] W. Ma, J. Campbell, D. Tran, and D. Kleeman. A Conceptual Framework for Assessing Password Quality. *International Journal of Computer Science and Network Security (IJCSNS)*, 7(1):179, 2007.

[87] P. Markillie. A perfect market: A survey of e-commerce. *The Economist*, 15:3–18, 2004.

[88] R. Martinez-Pelaez, J. Rico-Novella, V. Morales-Rocha, and M. Huerta. Digital pseudonym identity card to create digital identities. *IADIS E-commerce*, pages 313–318, 2006.

[89] M. McDonald and I. Dunbar. *Market Segmentation: How to Do It, how to Profit from it.* Butterworth-Heinemann, 2004.

[90] D. H. McKnight and L. C. Norman. What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce*, 6(2):35–59, 1/02.

[91] D. Miller. *A Theory of Shopping.* Cornell University Press, 1998.

[92] B. Mobasher, R. C. D. Paul, and J. Srivastava. Automatic personalization based on web usage mining. *Communications of the ACM*, 43, 2000.

[93] M. Molina, H. Bettiol, M. Barbieri, A. Silva, S. Conceição, and J. Dos-Santos. Food consumption by young adults living in Ribeirão Preto, SP, 2002/2004. *Braz J Med Biol Res*, 40(9):1257–1266, 2007.

[94] A. Morris, G. Jones, and J. Rubinsztein. Entry-level information systems personnel: a comparative study of ethical attitudes. *Proceedings of the 1993 conference on Computer personnel research*, 1993.

[95] E. Morris. Online customer experience: Will we get it right one day? `http://www.ecommercetimes.com/story/42274.html`, 2005. Accessed 29 June 2005.

[96] R. D. Newbold. *Newbold's Biometric Dictionary* . AuthorHouse, 2007.

[97] B. news. Airport's photo policy reviewed. `http://news.bbc.co.uk/1/hi/england/manchester/4415941.stm`. Accessed 26 August 2008.

[98] B. News. Anger as nhs patient records lost. `http://news.bbc.co.uk/1/hi/uk/7158498.stm`, 2007. Accessed 07 January 2008.

[99] J. Nielsen. F-shaped pattern for reading web content. `http://www.useit.com/alertbox/reading\_pattern.html`, April 2006. Accessed 18 June 2008.

[100] P. NORBERG, D. HORNE, and D. HORNE. The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007.

[101] OED. Oed. `http://dictionary.oed.com/`, 2006. Accessed 05 Sept 2006.

[102] T. U. H. Office. Identity crime definitions. `http://www.identity-theft.org.uk/definition.html`. Accessed 08 August 2008.

[103] Office of Public Sector Information (OPSI). Regulation of investigatory powers act 2000. `http://www.opsi.gov.uk/acts/acts2000/ukpga_20000023_en_1`. Accessed 21 July 2008.

[104] Office of Public Sector Information (OPSI). The Electronic Commerce (EC Directive) Regulations 2002. `http://www.opsi.gov.uk/si/si2002/20022013.htm#note11`. Accessed 21 July 2008.

[105] Office of Public Sector Information (OPSI). Data protection act 1998. `http://www.opsi.gov.uk/acts/acts1998/ukpga_19980029_en_2#pt1-l1g2`, 1998. Accessed 09 June 2008.

[106] N. Olivero and P. Lunt. Privacy versus willingness to disclose in e-commerce exchanges: The effect of risk awareness on the relative role of trust and control. *Journal of Economic Psychology*, 25(2):243 – 262, April 2004.

[107] Organisation for Economic Co-Operation and Development. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. `http://www.oecd.org/document/0,2340,en\_2649\_34255\_1815186\_1\_1\_1\_1,00.html`. Accessed 27 February 2008.

[108] Organisation For Economic Co-Operation And Development. The economic and social impact of electronic commerce preliminary findings and research agenda, 1999.

[109] M. Oxley and J. Miller. What Makes A Web Site "Sticky"? A Critical Element in Online Relationship Marketing Strategies. In B. Inc., editor, *ESOMAR Congress 2000 Proceedings*, September 2000.

[110] J. Park and S. Ram. Information systems interoperability: What lies beneath? *ACM Trans. Inf. Syst.*, 22(4):595–632, 2004.

[111] R. Pastor-Satorras and A. Vespignani. *Evolution and structure of the Internet : a statistical physics approach*. Cambridge Univeristy Press, 2004.

[112] Pay Pal. About us. `https://www.paypal-media.com/aboutus.cfm`. Accessed 07 August 2008.

[113] PayPal. About Us. `https://www.paypal.com/uk/cgi-bin/webscr?cmd=p/gen/about-outside`. Accessed on 01 July 2007.

[114] J. Preece. *Online Communities: Designing Usability and Supporting Sociability*. John Wiley and Sons, 2000.

[115] Princeton University. "privacy." WordNet 3.0. Princeton University. `http://wordnet.princeton.edu/perl/webwn?s=privacy&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h=`. Accessed 05 Sept 2006.

[116] Privacy and human rights 2003. Privacy and human rights 2003: Overview. `http://www.privacyinternational.org/survey/phr2003/overview.htm`, 2003. Accessed 23 October 2007.

[117] C. W. Proskauer Rose LLP. *Proskauer on Privacy: A Guide to Privacy and Data Security Law in the Information Age.* The Practising Law Institute PLI, 2006.

[118] A. A. Razaq, W. T. Luk, K. M. Shum, L. M. Cheng, and K. N. Yung. Second-Generation RFID. *IEEE Security and Privacy*, pages 21–27, 2008.

[119] P. Resnick and R. Zeckhauser. Trust Among Strangers in Internet Transactions: Empirical Analysis of eBays Reputation System. *The Economics of the Internet and E-Commerce*, 11:127–157, 2002.

[120] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on eBay: A controlled experiment. *Experimental Economics*, 9(2):79–101, 2006.

[121] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on eBay: A controlled experiment. *Experimental Economics*, 9(2):79–101, 2006.

[122] Retailer Internet. Retailers moving toward personalization, despite its challenges. `http://internetretailer.com/dailyNews.asp?id=13915`, 2005. Accessed 29 June 2005.

[123] L. Rosencrance. Amazon charging different prices on some dvds. `http://www.computerworld.com/industrytopics/retail/story/0,10801,49569,00.html`, 2000. Accessed 03 March 2007.

[124] G. Roussos and T. Moussouri. Consumer perceptions of privacy, security and trust in ubiquitous commerce. *Personal Ubiquitous Comput.*, 8(6):416–429, 2004.

[125] J. Rykowski and W. Cellary. Virtual web services: application of software agents to personalization of web services. In A. I. C. P. Series, editor, *6th international conference on Electronic commerce*, Delft, The Netherlands, 2004. ACM Press New York, NY, USA.

[126] E. Samhaber. *Merchants make history : how trade has influenced the course of history throughout the world.* London : Harrap, 1963.

[127] A. Sarner. Prepare to reinvest in e-commerce for growth. `http://www.gartner.com/research/spotlight/asset\_114586\_895.jsp`, 2005. Accessed 29 June 2005.

[128] J. B. Schafer, J. Konstan, and J. Riedi. Recommender systems in e-commerce. In *EC '99: Proceedings of the 1st ACM conference on Electronic commerce*, pages 158–166, New York, NY, USA, 1999. ACM.

[129] B. Schneier. Can new passport rfid technology be trusted? `http://www.crmbuyer,com/rsstory/53176.html`, 2006. Accessed 24 September 2006.

[130] J. E. Scott and C. H. Scott. Online Grocery Order Fulfillment Tradeoffs. *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 90–90, 2008.

[131] R.-A. Shang, Y.-C. Chen, and L. Shen. Extrinsic versus intrinsic motivations for consumers to shop on-line. *Information & Management*, 42(3):401–413, 2005.

[132] U. Shankar and C. Karlof. Doppelganger: Better browser privacy without the bother. *Proceedings of the 13th ACM conference on Computer and communications security*, pages 154–167, 2006.

[133] V. Shankar, A. Smith, and A. Rangaswamy. Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, 20(2):153–175, 2003.

[134] H. Sharp, Y. Rogers, and J. Preece. *Interaction Design: Beyond Human Computer Interaction.* John Wiley & Sons, 2007.

[135] P. Smith and D. Chaffey. *Emarketing Excellence: The Heart of Ebusiness.* Butterworth-Heinemann, 2002.

[136] R. M. Smith. Privacy and the dot-com bubble. `http://www.ftc.gov/bcp/workshops/infomktplace/comments/smith-richard.pdf`, 2000. Accessed 13 July 2008.

[137] J. Snell, D. Tidwell, and P. Kulchenko. *Programming Web Services with SOAP.* O'Reilly Media, Inc., 2002.

[138] P. K. Sokol. *From EDI to electronic commerce : a business initiative.* McGraw-Hill Inc, 1995.

[139] S. Spanbauer. Internet tips: Take charge of what web sites know about you. `http://www.pcworld.com/article/id,124583/article.html\#`, 2006. Accessed 02 April 2006.

[140] S. Spiekermann, J. Grossklags, and B. Berendt. E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavior. In *Electronic Commerce.* ACM, 2001.

[141] S. Sproule and N. Archer. Defining Identity Theft. *Management of eBusiness, 2007. WCMeB 2007. Eighth World Congress on the*, pages 20–20, 2007.

[142] J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan. Web usage mining: discovery and applications of usage patterns from web data. *SIGKDD Explor. Newsl.*, 1(2):12–23, 2000.

[143] F. Stajano. Will your digital butlers betray you? In *WPES '04: Proceedings of the 2004 ACM workshop on Privacy in the electronic society*, pages 37–38, New York, NY, USA, 2004. ACM.

[144] Sun microsystems. Java Web Start. `http://java.sun.com/javase/6/docs/technotes/guides/javaws/developersguide/faq.html\#101`. Accessed 19 September 2008.

[145] P. Syverson. The paradoxical value of privacy. In *2nd Annual Workshop economics and Information Security.* Robert H. Smith School of Business, 2003.

[146] T. Zeller Jr. Personal data for the taking. http://query.nytimes.com/gst/abstract.html?res=F30B1EF83D5D0C7B8DDDAC0894DD404482, 2005. Accessed 29 June 2005.

[147] K. Y. Tam and S. Y. Ho. Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Info. Sys. Research*, 16(3):271–291, 2005.

[148] A. S. Tanenbaum and M. van Steen. *Distributed Systems: Principles and Paradigms.* Upper Saddle River, N.J. : Prentice-Hall, 2002.

[149] C. R. Taylor. Private Demands and Demands For Privacy: Dynamic Pricing and the Market for Customer Information. *SSRN eLibrary*, 2002.

[150] The Associated Press. CIA employees easily identified by Internet searches. http://www.usatoday.co/tech/news/internetprivacy/2006-03-12-cia-net_x.htm, 2006. Accessed 23 March 2006.

[151] The European Interactive Advertising Association (EIAA). Mediascope europe study. http://www.eiaa.net/research/media-consumption.asp?lang=6. Accessed 08 July 2008.

[152] The Hindu Business Line. Shopping for experience. http://www.thehindubusinessline.com/catalyst/2005/03/10/stories/2005031000130200.htm, 2005. Accessed 29 June 2005.

[153] P. Underhill. *Why We Buy: The Science Of Shopping.* Simon & Schuster, 2000.

[154] U.S. Department of Commerce. QUARTERLY RETAIL E-COMMERCE SALES 46^{th} QUARTER 2007. http://www.census.gov/mrts/www/data/pdf/07Q4.pdf. Accessed 27 May 2008.

[155] J. Vogel. Getting to know all about you. http://archive.salon.com/21st/feature/1998/10/14featureb.html, 1998. Accessed 20/02/2007.

[156] V. Wagner. Your reputation online, part 1: How damage is done. http://www.technewsworld.com/story/63740.html. Accessed 23 July 2008.

[157] S. D. Warren and L. D. Brandeis. The right to privacy. *Harvard Law Review.*, 4(5), 1890.

[158] What's News at JUNKBUSTERS. Shoppers cards used against shoppers? http://www.junkbusters.com/new.html. Accessed 26 August 2008.

[159] D. Whiteley. *E-commerce: strategy, technologies and applications.* McGraw-Hill, London, 2000.

[160] Yahoo. Lenox increases email response rate 32% and sales 41% using scene7 to personalize images. http://www.forrelease.com/D20050202/sfw008.P2.02012005020033.11106.html, February 2005. Accessed 29 June 2005.

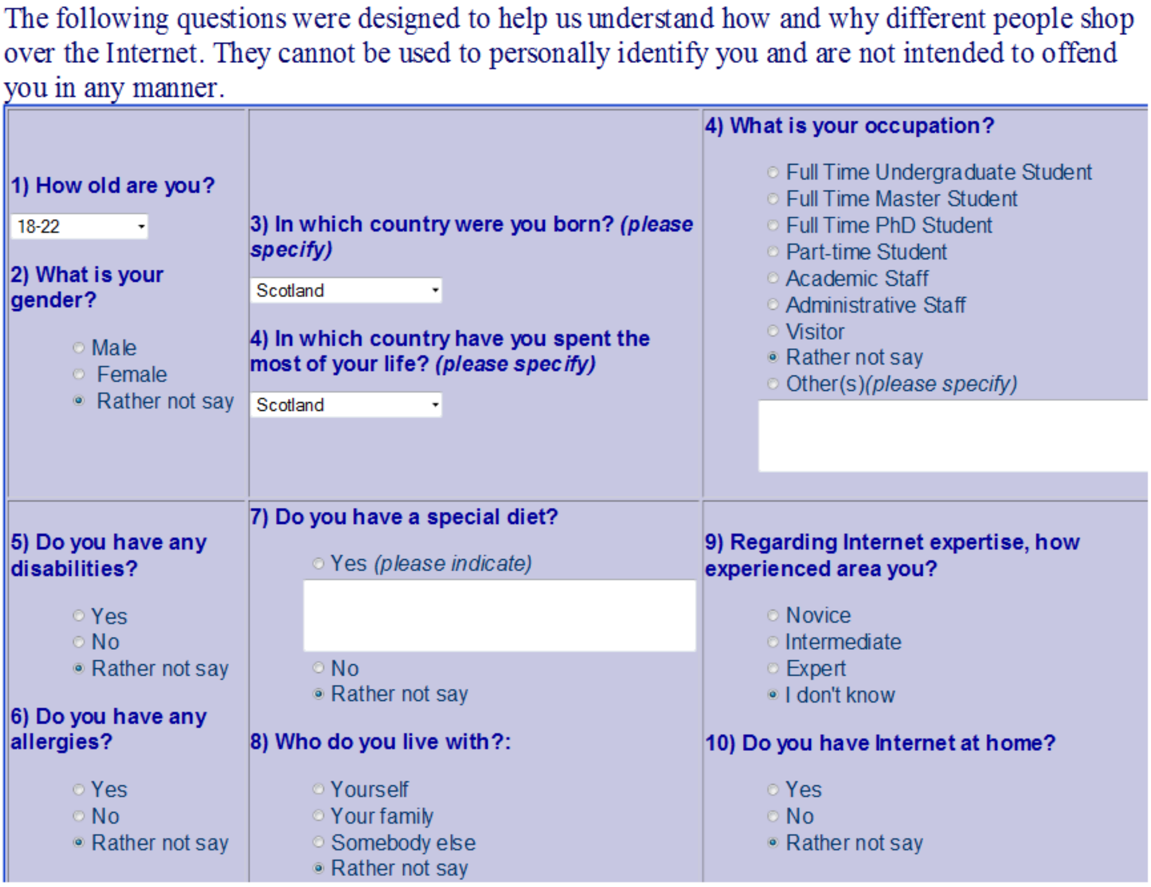
[161] Z. Yang, S. Cai, Z. Zhou, and N. Zhou. Development and validation of an instrument to measure user perceived service quality of information presenting Web portals. *Information & Management*, 42(4):575–589, 2005.

[162] B. Yuille. Stolen innocence: Child identity theft. http://articles.moneycentral.msn.com/Banking/FinancialPrivacy/StolenInnocenceChildIdentityTheft.aspx, 2008. Accessed 04 March 2008.

Appendix A

Questionnaires

A.1 Demographics



The following questions were designed to help us understand how and why different people shop over the Internet. They cannot be used to personally identify you and are not intended to offend you in any manner.

1) How old are you?

18-22 ▾

2) What is your gender?

○ Male
○ Female
● Rather not say

3) In which country were you born? *(please specify)*

Scotland ▾

4) In which country have you spent the most of your life? *(please specify)*

Scotland ▾

4) What is your occupation?

○ Full Time Undergraduate Student
○ Full Time Master Student
○ Full Time PhD Student
○ Part-time Student
○ Academic Staff
○ Administrative Staff
○ Visitor
● Rather not say
○ Other(s)*(please specify)*

5) Do you have any disabilities?

○ Yes
○ No
● Rather not say

6) Do you have any allergies?

○ Yes
○ No
● Rather not say

7) Do you have a special diet?

○ Yes *(please indicate)*
○ No
● Rather not say

8) Who do you live with?:

○ Yourself
○ Your family
○ Somebody else
● Rather not say

9) Regarding Internet expertise, how experienced area you?

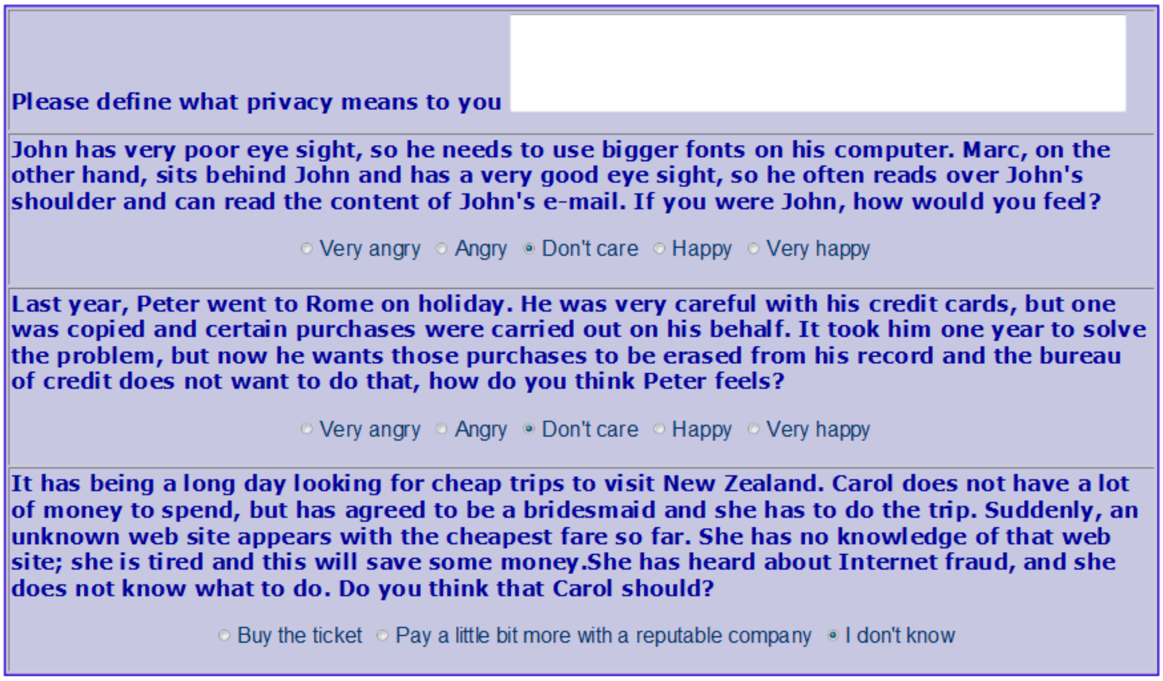
○ Novice
○ Intermediate
○ Expert
● I don't know

10) Do you have Internet at home?

○ Yes
○ No
● Rather not say

Figure A.1: Questionnaire - Demographics.

A.2 Privacy perceptions



Please define what privacy means to you

John has very poor eye sight, so he needs to use bigger fonts on his computer. Marc, on the other hand, sits behind John and has a very good eye sight, so he often reads over John's shoulder and can read the content of John's e-mail. If you were John, how would you feel?

○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

Last year, Peter went to Rome on holiday. He was very careful with his credit cards, but one was copied and certain purchases were carried out on his behalf. It took him one year to solve the problem, but now he wants those purchases to be erased from his record and the bureau of credit does not want to do that, how do you think Peter feels?

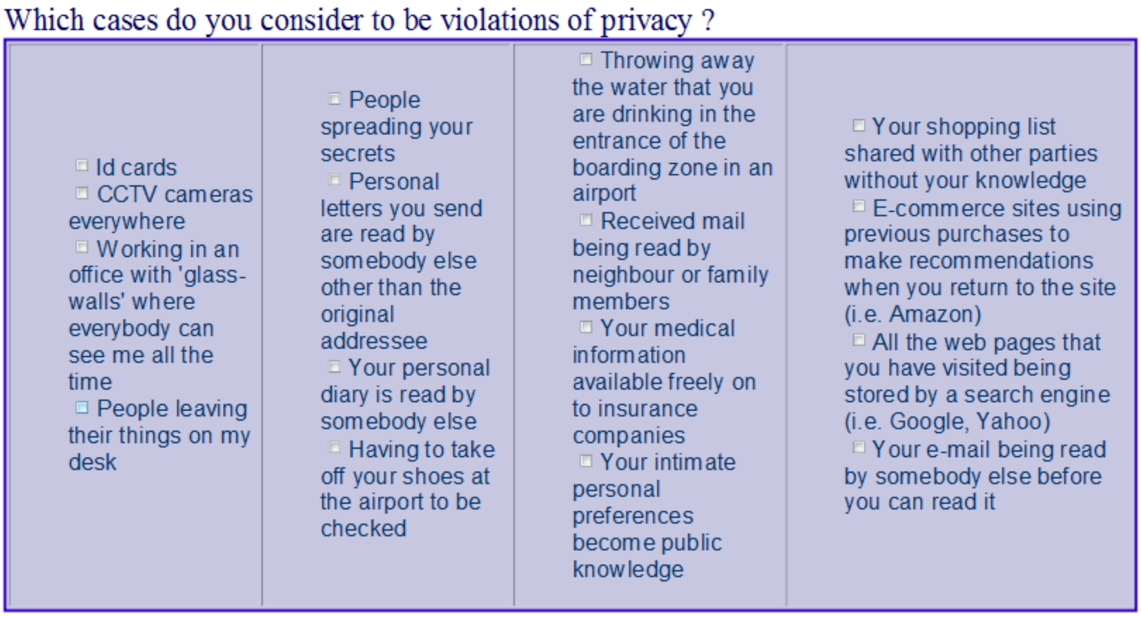
○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

It has being a long day looking for cheap trips to visit New Zealand. Carol does not have a lot of money to spend, but has agreed to be a bridesmaid and she has to do the trip. Suddenly, an unknown web site appears with the cheapest fare so far. She has no knowledge of that web site; she is tired and this will save some money.She has heard about Internet fraud, and she does not know what to do. Do you think that Carol should?

○ Buy the ticket ○ Pay a little bit more with a reputable company ● I don't know

Figure A.2: Questionnaire - Privacy perceptions.

A.3 Privacy perceptions according to privacy violations



Which cases do you consider to be violations of privacy ?

□ Id cards
□ CCTV cameras everywhere
□ Working in an office with 'glass-walls' where everybody can see me all the time
□ People leaving their things on my desk

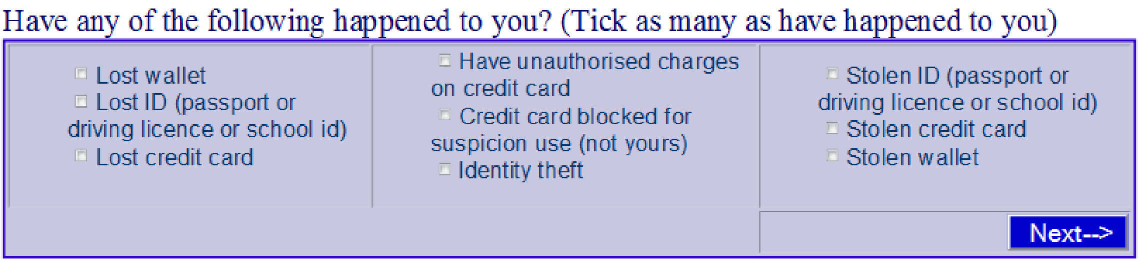
□ People spreading your secrets
□ Personal letters you send are read by somebody else other than the original addressee
□ Your personal diary is read by somebody else
□ Having to take off your shoes at the airport to be checked

□ Throwing away the water that you are drinking in the entrance of the boarding zone in an airport
□ Received mail being read by neighbour or family members
□ Your medical information available freely on to insurance companies
□ Your intimate personal preferences become public knowledge

□ Your shopping list shared with other parties without your knowledge
□ E-commerce sites using previous purchases to make recommendations when you return to the site (i.e. Amazon)
□ All the web pages that you have visited being stored by a search engine (i.e. Google, Yahoo)
□ Your e-mail being read by somebody else before you can read it

Figure A.3: Questionnaire - Privacy perceptions according to privacy violations.

A.4 Privacy perceptions according to privacy awareness



Have any of the following happened to you? (Tick as many as have happened to you)

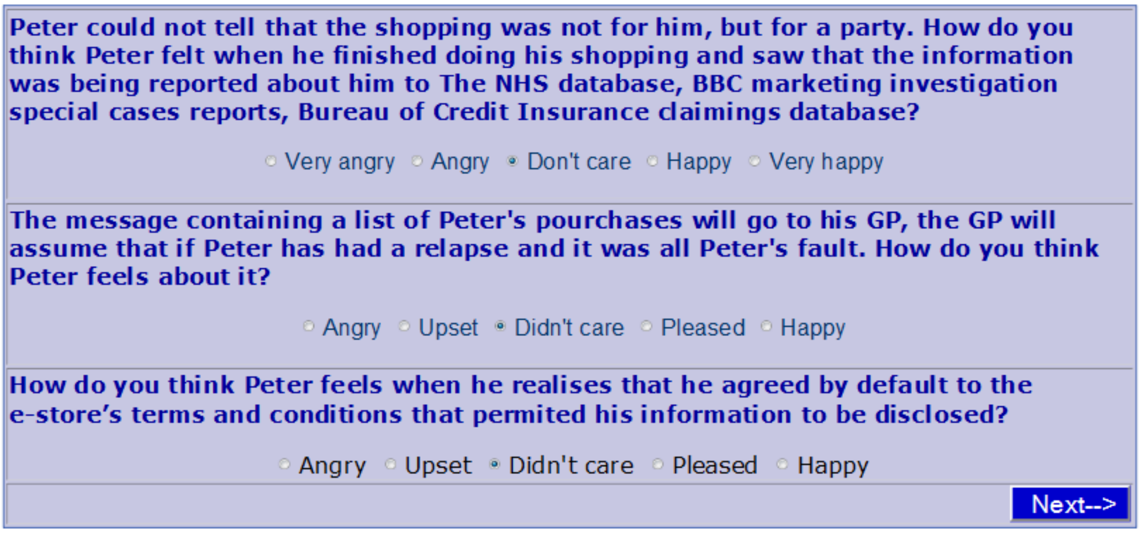
□ Lost wallet
□ Lost ID (passport or driving licence or school id)
□ Lost credit card

□ Have unauthorised charges on credit card
□ Credit card blocked for suspicion use (not yours)
□ Identity theft

□ Stolen ID (passport or driving licence or school id)
□ Stolen credit card
□ Stolen wallet

Figure A.4: Questionnaire - Privacy perceptions according to privacy awareness.

A.5 Questionnaire A



Peter could not tell that the shopping was not for him, but for a party. How do you think Peter felt when he finished doing his shopping and saw that the information was being reported about him to The NHS database, BBC marketing investigation special cases reports, Bureau of Credit Insurance claimings database?

○ Very angry ○ Angry ● Don't care ○ Happy ○ Very happy

The message containing a list of Peter's pourchases will go to his GP, the GP will assume that if Peter has had a relapse and it was all Peter's fault. How do you think Peter feels about it?

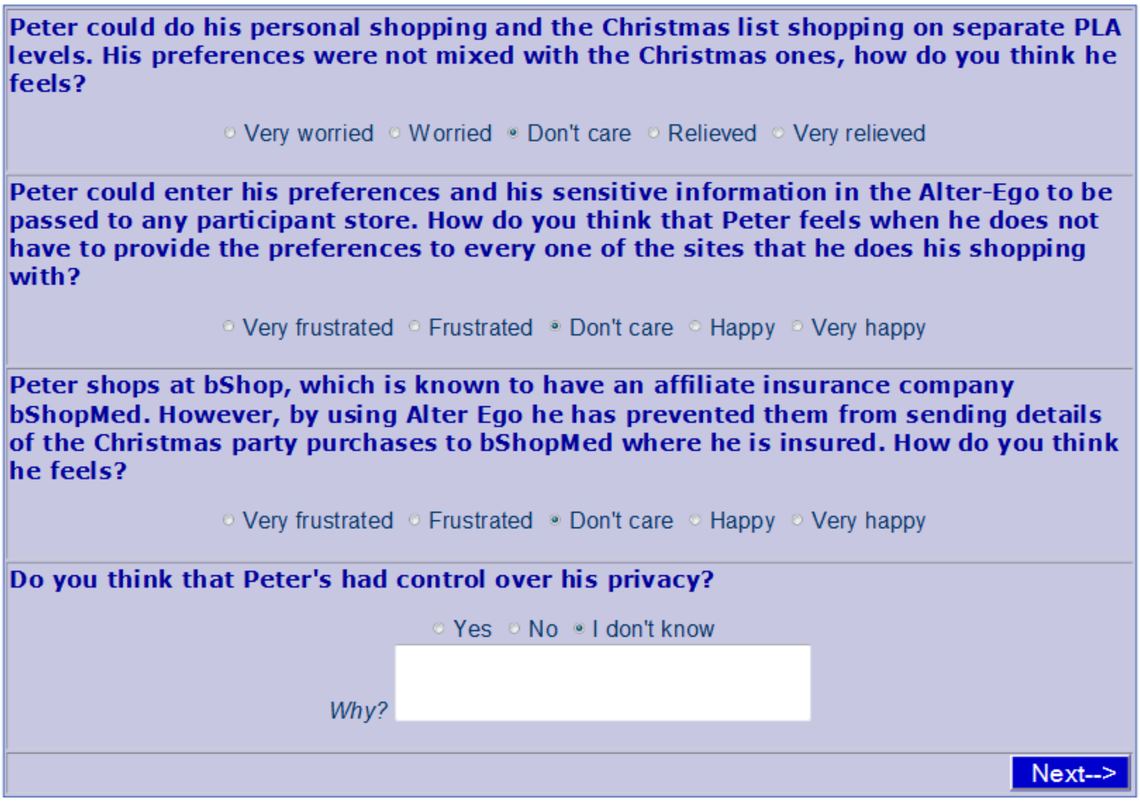
○ Angry ○ Upset ● Didn't care ○ Pleased ○ Happy

How do you think Peter feels when he realises that he agreed by default to the e-store's terms and conditions that permited his information to be disclosed?

○ Angry ○ Upset ● Didn't care ○ Pleased ○ Happy

Figure A.5: After using the bshop in a non-PPSE environment.

A.6 Questionnaire B



Peter could do his personal shopping and the Christmas list shopping on separate PLA levels. His preferences were not mixed with the Christmas ones, how do you think he feels?

○ Very worried ○ Worried ● Don't care ○ Relieved ○ Very relieved

Peter could enter his preferences and his sensitive information in the Alter-Ego to be passed to any participant store. How do you think that Peter feels when he does not have to provide the preferences to every one of the sites that he does his shopping with?

○ Very frustrated ○ Frustrated ● Don't care ○ Happy ○ Very happy

Peter shops at bShop, which is known to have an affiliate insurance company bShopMed. However, by using Alter Ego he has prevented them from sending details of the Christmas party purchases to bShopMed where he is insured. How do you think he feels?

○ Very frustrated ○ Frustrated ● Don't care ○ Happy ○ Very happy

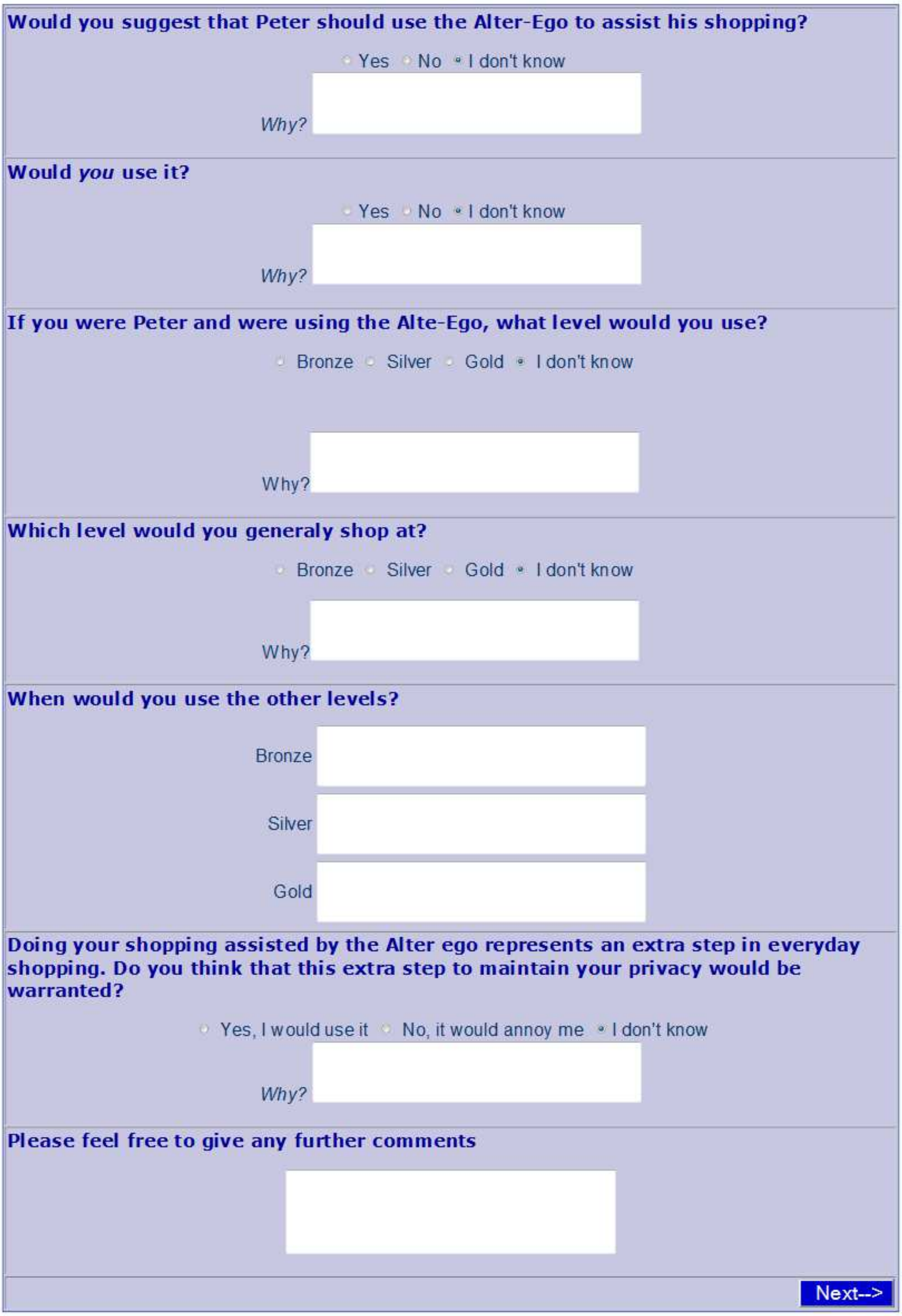
Do you think that Peter's had control over his privacy?

○ Yes ○ No ● I don't know

Why?

Figure A.6: After using the bshop in a PPSE environment.

A.7 Final questionnaire



Would you suggest that Peter should use the Alter-Ego to assist his shopping?

○ Yes ○ No ● I don't know

Why?

Would *you* use it?

○ Yes ○ No ● I don't know

Why?

If you were Peter and were using the Alte-Ego, what level would you use?

○ Bronze ○ Silver ○ Gold ● I don't know

Why?

Which level would you generaly shop at?

○ Bronze ○ Silver ○ Gold ● I don't know

Why?

When would you use the other levels?

Bronze

Silver

Gold

Doing your shopping assisted by the Alter ego represents an extra step in everyday shopping. Do you think that this extra step to maintain your privacy would be warranted?

○ Yes, I would use it ○ No, it would annoy me ● I don't know

Why?

Please feel free to give any further comments

Figure A.7: After using PPSE and non-PPSE environments.

Appendix B

Scenarios

B.1 Scenario 1

Peter went to the doctor and he was very happy when he found that his hypertension and cholesterol are controlled now. He only has to keep up a low consumption of salt and red meat if he wants to avoid hospitalisation. Because things are better, he has been allowed to drink a little red wine; no beer and no other alcohol can be consumed. Peter's doctor was very clear: no chances are to be taken; he will be under close observation.

The Christmas Party:
The Christmas party is approaching and Peter has to buy all the things for the party, here is the list:

- The accountant girls love wine, so he has to buy at least 5 bottles of red wine and 5 bottles of white wine
- Susan volunteered to prepare some food, so she needs 10 packages of beef escalope for the kebabs
- Finally, he has to buy some crisps, enough for 20 people.

Please use the bShop to do Peter's shopping.

In the checkout section use the following values

- **Name: Peter**
- **Address: a**
- **City: a**

- **Postcode: a**

B.2 Scenario 2

Peter has being having severe problems with his health, and because it is not the first time that he has fallen ill, he does not want to risk any problems with his health anymore. The doctor has advised him to control his consumption of salt and red meat. Because Peter is a very busy person, he does not want to spend a lot of time buying his groceries that have some limitations now. Therefore, he decides to use the Alter-Ego web site to set his preferences and use it to assist his buying.

> **Please use the same user name and password that you used during the training to enter Alter ego web site**

The preferences that Peter wants to set are (feel free to fill the others options with the values you desire):

In Silver level

- Beef 'No (but show them to me anyway)'
- Shellfish 'No (but show them to me anyway)'

In Gold level

- Salt 'Never'
- Alcohol 'No (but show them to me anyway)'
- Fat 'Never'

Shopping with "Silver level"

Peter has his preferences set in the Alter ego web site, so he goes to the left side navigator bar to "Select & Shop" and selects the bShop he wants to use and clicks the add icon on the right of the name of the participating sites.

Because this shopping is for the Christmas party he does not want it to be constrained by his normal shopping preferences, so he decides to do his shopping using silver level preferences only.

Please select 'Silver' by clicking the medals icons in the left hand side images. Then click in the store name and do the shopping.

After setting the preferences, Peter can face his Christmas shopping task.

The Christmas Party:
The Christmas party is approaching and Peter has to buy all the things for the party, here is the list:

- The accountant girls love wine, so he has to buy at least 5 bottles of red wine and 5 bottles of white wine
- Susan volunteered to prepare some food, so she needs 10 packages of beef escalope for the kebabs
- Finally, he has to buy some crisps, enough for 20 people.

Please use the bShop to do Peter's shopping. In the checkout section use the following values

- **Name: Peter**
- **Address: a**
- **City: a**
- **Postcode: a**

Acronyms

ART: Raising *Awareness*, *Regulation* and use of *Technology.* Three approaches identified in this research, used by organisations and initiatives towards the preservation of privacy.

B2C: Business-to-Consumer.

C2C: Consumer-to-Consumer.

B2B: Business-to-Business.

PLA: Personal Level Agreement, an agreement that has the objective of formalise the exchange of non-identifiable sensitive and/or belief-based information between customer and e-tailer.

PPSE: Privacy Preserving Shopping Environment, an approach to the e-shopping where the customer's privacy needs are respected while allowing the company to gather and use sufficient customer-specified data to achieve a level of personalisation which can be used to encourage customer loyalty.

W3C: World Wide Web consortium, international consortium where member organisations, full-time staff, and the public work together to develop Web standards.

P3P: Platform for Privacy Preferences, project created with the main aim of expressing privacy practices in a machine readable way.

Appendix C

Glossary

Alter-Ego: A third party Web portal which objective is facilitate and mediate the customer's disclosure of information and the e-tailer's user-specified data requirements.

Customer: In the context of this research, customer is the participant that uses the PPSE approach to assist his or her shopping.

Data: "“data” means information which (a)is being processed by means of equipment operating automatically in response to instructions given for that purpose, (b)is recorded with the intention that it should be processed by means of such equipment, (c)is recorded as part of a relevant filing system or with the intention that it should form part of a relevant filing system, or (d)does not fall within paragraph (a), (b) or (c) but forms part of an accessible record as defined by Section 68" [105].

Market: Can be defined as "a customer need that can be satisfied by the products or services seen as alternatives" [89].

Participant store(s): The store(s) that comply with the privacy level agreement (PLA), the PPSE privacy policy and are considered part of the privacy preserving shopping environment (PPSE).

Privacy invasive system: Facilitates or enables the usage of personal data in a fashion inconsistent with generally accepted privacy principles [96][p133].

Privacy neutral system: Privacy is not an issue or in which the potential privacy impact is slight. Privacy-neutral systems are difficult to misuse from a privacy perspective, but do not have the capability to protect personal privacy [96][p133].

Privacy protective system: Used to protect or limit access to personal information or which provide a means for an individual to establish a trusted identity [96][p133].

Privacy sympathetic system: Limits access to and usage of personal data and in which decisions regarding design issues such as storage and transmission of biometric data are informed, if not driven, by privacy concerns [96][p134].