Citation Details

Citation for the version of the work held in 'OpenAIR@RGU':

BIERIG, R., 2008. Event and map content personalisation in a mobile and context-aware environment. Available from *OpenAIR@RGU*. [online]. Available from: http://openair.rgu.ac.uk

EVENT AND MAP CONTENT PERSONALISATION IN A MOBILE AND CONTEXT-AWARE ENVIRONMENT

RALF BIERIG

A thesis submitted in partial fulfilment of the
requirements of
the Robert Gordon University
for the degree of Doctor of Philosophy

This research programme was carried out
in collaboration with SINTEF Telecom and Informatics

April 2008

It takes an awful long time to not write a book.

Douglas Adams

Abstract

Effective methods for information access are of the greatest importance for our modern lives – particularly with respect to handheld devices. Personalisation is one such method which models a user's characteristics to deliver content more focused to the user's needs. The emerging area of sophisticated mobile computing devices has started to inspire new forms of personalised systems that include aspects of the person's contextual environment. This thesis seeks to understand the role of personalisation and context, to evaluate the effectiveness of context for content personalisation and to investigate the event and map content domain for mobile usage. The work presented in this thesis has three parts:

The *first part* is a user experiment on context that investigated the contextual attributes of time, location and interest, with respect to participants' perception of their usefulness. Results show highly dynamic and interconnected effects of context on participants' usefulness ratings.

In the *second part*, these results were applied to create a predictive model of context that was related to attribution theory and then combined with an information retrieval score to create a weighted personalisation model.

In the *third part* of this work, the personalisation model was applied in a mobile experiment. Participants solved situational search tasks using a (i) non-personalized and a (ii) personalized mobile information system, and rating entertainment events based on

usefulness. Results showed that the personalised system delivered about 20% more useful content to the mobile user than the non-personalised system, with some indication for reduced search effort in terms of time and the amount of queries per task.

The work presented provides evidence for the promising potential of context to facilitate personalised information delivery to users of mobile devices. Overall, it serves as an example of an investigation into the effectiveness of context from multiple angles and provides a potential link to some of the aspects of psychology as a potential source for a deeper understanding of contextual processes in humans.

Acknowledgements

I thank my parents Richard and Edith Bierig and my sister Stefanie Bierig for their patience and support during the time in which I researched and wrote my doctoral thesis.

I also thank my two supervisors Dr. Ayşe Göker and Dr. Stuart Watt for their guidance and support during my doctoral studies that naturally had its highs and lows. I thank them for giving me the opportunity to work with them, for their patience, and for all the illuminating and fruitful discussions we had while I undertook my research. I thank my examiners, Prof. Pia Borlund and Prof. Dorothy Williams, for a searching viva and for the time and good advice that improved this document. I thank the international consortium of the EU-IST AmbieSense project and its project coordinator Hans I. Myrhaug for all the support (during the project and afterwards) and a very intense and dynamic work experience. Furthermore, I thank Dr. Alex Wilson, Dr. Roger McDermott and many anonymous experts from the L-STAT mailing list for their advice and endless patience in statistical matters.

I want to express my gratitude to all those who read parts of my work at various stages in progress; namely Dr. Diane Kelly, Prof. Ian Ruthven, Murat Yakici and Dr. Mark Baillie. I am grateful for all the resources and the support I received from the University of Strathclyde in Glasgow between 2005 and 2007; I thank Prof. Fabio Crestani and Fabio Simeoni for offering me a flexible work style that enabled continued progress on my thesis. Furthermore, I thank all those who proof-read my work over the years and made this thesis readable.

I am grateful for the advice I received from Prof. David Bawden and Dr. Tamara Eisenschitz from the Department of Information Science at City University in London; I am also thankful to the department in general for enabling the supervision to continue.

I thank my colleagues and fellow PhD students at Smartweb/CTC, who all made the time at Aberdeen an enjoyable and memorable one. I especially like to thank Zia, Fiona, Siddartha, Sudha, Ratiba, Stewart, Kefang, Ivan, Daniel, Dietrich, Rahman, Stella and Selpi for their friendship and company. I would like to express respect to the RGU security guards at St. Andrews Street building, whose rules I was able to bend exclusively on several occasions with great joy. I thank Prof. Susan Craw for her professionalism and her ability to deal with my sense of humour. I also thank Prof. David Harper for initiating Smartweb/CTC as one amazing workspace.

Then I would like to thank the office personnel of the School of Computing and their support and help during the years; Diane, Kathy, Ann and Marie for their constant and cheerful assistance.

A special thanks goes to Prof. Hans-Volker Niemeier and Prof. Mohsen Rezagholi from the University of Furtwangen for their early support during the application process for this PhD. I also thank Prof. Unruh for not supporting me, it made me a stronger person.

This study was fully funded by a research studentship from the School of Computing of the Robert Gordon University in Aberdeen and provided through the EU-IST AmbieSense project.

Declarations

This thesis has been compiled by myself and describes my own work. All mention of other work have been duly cited in the bibliography.

Ralf Bierig

Publications

The following publications have been produced during the time of this thesis that present work in connection with this thesis:

Göker, A., Myrhaug H. I., Bierig R. (2009). Context and Information Retrieval. *Information Retrieval: Searching in the 21st Century.* Goker A., and Davies J. (eds.), John Wiley and Sons, Ltd. ISBN-13: 978-0470027622.

Bierig R., Göker A. (2006). Time, Location, and Interest: An Empirical and User-Centred Study. *First Symposium on Information Interaction in Context (IIiX).* Copenhagen, Denmark. ACM-Press.

Bierig R.(2005). Personalisation and Adaptation of Content in an Ambient and Context-Aware Environment. Doctoral Forum of the *5th International Conference on Conceptions of Library and Information Sciences (CoLIS).* Glasgow, UK.

Göker A., Yakici M., Bierig R., and Myrhaug H. I. (2004). A Context-sensitive Information System for Mobile Users (Demo). *SIGIR Workshop on Information Retrieval in Context (IRiX).* Sheffied, UK.

Göker A., Watt S., Myrhaug H. I., Whitehead N., Yakici M., Bierig R., Nuti S.K., Cumming H. (2004). An Ambient, Personalised, and Context-Sensitive Information System

for Mobile Users. *Second European Symposium on Ambient Intelligence*. Eindhoven, Netherlands.

Contents

List of Figures

List of Tables

1 Introduction

> The White Rabbit put on his spectacles. ”Where shall I begin, please your Majesty?” he asked. ”Begin at the beginning,” the King said very gravely, ”and go on till you come to the end: then stop.”
>
> *Alice in Wonderland*
> Lewis Carroll

1.1 Motivation

This thesis presents an investigation into the effectiveness of context as a means to personalise content for mobile users. In particular, this work is centred around the personalised delivery of entertainment events with the use of geographic maps, two types of content that are relevant when mobile. For this, the thesis presents two studies in which a context model is established, evaluated in a simulated laboratory experiment, and used to define a personalisation model that is then evaluated in a simulated mobile experiment (see section 1.7 and 1.8).

Information and its retrieval is a crucial element of our daily lives. Similar to the steam engine that carried the world into the industrial age, information powers the information age. In only half a century, less than a human lifetime, an extraordinary development has taken place through the fast growth of a rich variety of computing equipment, intelligent software and an almost unlimited amount of information, mostly provided via the web. Today, the average person has access to more information than ever before. The growth of the web is still measured on an exponential scale. The NetCraft report from December 2007 [1] shows about 158 million hosts on the web, about 33 million more than in June 2007 and about 58 million more than at the end of 2006. This demonstrates how important information has become and how critical it is to manage information effectively. When Vannevar Bush wrote his key article that first stated the problem of "the massive task of making more accessible a bewildering store of knowledge" [Bush, 1945] – an expression that later coined the term 'Information Overload' – he may not have imagined how relevant his statement would become in the future. Different areas of information science are aimed at handling this challenge with a variety of methods and tools that allow people to effectively manage and use large and increasing amounts of information – content personalisation is one of these methods.

Generally speaking, personalisation is the umbrella term for tailoring products and services to personal needs; this is done for different reasons. As described in [Göker and Myrhaug, 2002], the automobile industry for example uses personalisation to allow customers to adapt their car to their individual needs by choosing between alternative basic features (e.g. the engine) and extending the basic model with additional features (e.g. a navigation system or an extra service package). The main purpose here is to create better targeted products which are more useful for individual customers which in turn improves the relationship between business and customer [Kobsa et al., 2001]. Similarly, a personalised information system adapts content information or system behaviour to the needs of an individual user (e.g. by including more detailed information or arranging information based on personal preferences). This is usually accomplished by considering additional information about this user like web logs, shopping cart history or user feedback and is normally facilitated through a user model. The aim is to improve the effectiveness of conventional information systems to fulfil users' information needs faster and more accurate.

[1] Available from `http://www.netcraft.com`, accessed April 14, 2008. NetCraft is an internet service that provides data and analysis of a range of different aspects and trends about the internet since 1995.

Although personalisation is appealing and promising it is also often over-rated [Nielsen, 1998] considering the effort required for providing quality content and presentation. It is difficult for a personalised service to improve poor content, inconsistent navigation and weak presentation. However, personalisation is a tool that can enhance a good website or an information system with extra value, as described in [Manber et al., 2000] for the MyYahoo! personalised portal. Generally, it is beneficial to carefully adjust a personalised solution to the domain for which it is made. This thesis suggests adapting a personalised service depending on *usage* (the way in which the system is applied by its user) and what kind of *content* it personalises. In particular, this thesis is concerned with semi-mobile and mobile usage (using a system while being away from the usual work or home environment or while moving) and focuses on personalising entertainment and map content for people. This will be further discussed in section 1.2 and section 1.3. After that, section 1.4 presents open challenges in personalisation research followed by section 1.5 that lists the research questions and defines the focus of this thesis. The contributions of this work are highlighted in section 1.6. The structure of this thesis is listed in section 1.7 and an overview to the studies of this thesis is presented in section 1.8.

1.2 Personalisation and Mobile Context

The shift to mobile computing is perhaps one of the most significant and rapid changes in recent years regarding how people use information systems. Users are mobile and want to use handheld devices for daily tasks while being away from their normal office and home environments or on the move between these environments. This new way of using information systems imposes new challenges on existing personalised solutions that so far have only operated in relatively static desktop environments. In the past, personalised information systems mainly operated based on information from user models; such models consider facts about the user and their past behaviour to recommend or adapt content in the future. Context, however, represents the situation around the user (e.g. weather, location and time) while allowing the inclusion of the user (e.g. interest, role and physical condition) as part of the situation. Context models represent a particular view on context for a specific application. Such models have been developed and used for information retrieval and seeking (IR&S), such as Ingwersen's cognitive IR&S model [Ingwersen, 1992] or Belkin's Episode Model [Belkin et al., 1995]. In the last decade, context modelling has also become a major focus of interest for context-aware computing as recent surveys show [Kaenampornpan and ONeill, 2004, Strang and Linnhoff-Popien, 2004, Baldauf et al., 2007]. Furthermore, context has been identified as a good source to improve personalisation [Myrhaug and Göker, 2003] in support of users' increasing mobility.

In the past decade, context-aware systems were often developed in the form of mobile guides. Subsequently, many different frameworks for context management were established that support the development of systems and the modelling of context from a technical perspective. Different context models were established and applied for different purposes in context-aware computing. However, few studies demonstrate empirical work and investigate these context models and their attributes in more detail. This thesis argues that further work is necessary to critically evaluate the impact of context. Individual context attributes need to be understood in much greater detail relative to how they affect people. Also, contextual effects that result from a number of perhaps interrelated attributes need to be understood. Naturally, attributes may not only occur in isolation but may be connected. Understanding their nature is therefore important for the delivery of personalised services that help users to obtain useful results and access and manage their information resources effectively. This thesis contributes to this need by investigating three common context attributes - time, location and interest - in chapter 4.

1.3 User Context for Event and Map Personalisation

As mobile computing differs from traditional desktop computing, mobile users are likely to prefer other forms of content than users that work at desktops. Even if the same types of content are used in both environments, it is likely that priorities on content are shifting based on the different usage situation. Map and event content are two forms of content which are interesting for mobile users as initial studies with the AmbieSense project have shown (see chapter 3). These results were obtained from two user studies in Seville/Spain and a large-scale market survey; the survey collected data from travellers and tourists on site in Seville and Oslo/Norway together with online questionnaires through travel websites. The results indicated that content about transportation, food/restaurants, sites/attractions, maps and events are among the most requested types, especially in relation to users' personal devices. In particular entertainment events and map content are regarded as highly important (85%) by mobile users [2]. This suggests that mobile applications that provide information about events combined with geographic maps can be especially useful in mobile settings. In this thesis, these two content types are investigated closely for their applicability in a personalised information system.

Map personalisation is a relatively new field of research and only been initially addressed. Map adaptation along device and bandwidth parameters has been explored [Chalmers, 2002], initial theoretical work on ideas for map personalisation has been published [Zipf, 2002] and initial investigations have started [Reichenbacher, 2003, Reichenbacher, 2007], however, the topic of mobile map personalisation is still young and has only started to emerge during the development of this thesis.

Furthermore, event content has not yet been addressed in the literature as a distinctive type of content; information on events has so far been treated as news content [Pazzani, 2002]. However, the results from the user studies in chapter 3 (see figure 3.7 on page 68) indicated that event and news content are considered to be very different for mobile users (85% relevance for events, but only 54% relevance for news). This suggests that events represent a distinct form of content. Initial evidence is provided that shows its particular relevance in mobile situations. This does not suggest a shift of event and map use to mobile devices, but rather an extension of use for these two content types via mobile applications.

1.4 Open Challenges for Personalised Information Systems

Numerous studies have been conducted on personalisation in the last decade. Many personalised systems demonstrated their benefit and increased value for users

[2] This was one of the studies that was conducted as part of AmbieSense EU-IST. These studies are reported and discussed comprehensively in chapter 3.

[Lieberman, 1995], [Mooney and Roy, 2000], [Pazzani, 2002]. A range of different personalised information systems and concepts have been presented in a special issue on recommender systems [Resnick and Varian, 1997] and in a later published special issue on personalisation [Riecken, 2000]. Personalisation and Recommender Systems have been explored for digital libraries [Smeaton and Callan, 2001] and for user modelling [Brusilovsky et al., 2005]. Nevertheless, there are still many challenges that remain open and unanswered:

- *Mobile personalisation:* Personalisation, that originally developed in desktop environments now faces a new situation where users become increasingly mobile. In the past few years, Personal Digital Assistants (PDA's) and mobile phones have developed into small and powerful minicomputers with a growing ability to handle secondary software and rich multimedia supported by constantly increasing performance in memory, processing power and communication abilities (such as Bluetooth and Wireless LAN). The users' access to and use of mobile devices provide new challenges for personalised information systems, which have only been addressed a minimum of attention.

- *Personalisation and context:* Context has been identified as being advantageous for a personalised service [Myrhaug and Göker, 2003]. Contextual information introduces a new point of view upon which personalisation can be performed and extended beyond the traditional method of personalisation based on user models.

- *Contextual effects:* Several context models exist and many different systems and frameworks have been developed to date. However, little is known about the precise effects of certain context attributes although being commonly applied in various systems. Detailed knowledge of contextual effects is important for the creation of effective context-aware applications; a challenging task when designing personalised information systems.

- *Evaluation of context:* The evaluation of context-aware systems is another area where little is known, as such systems are used in settings where conditions change continuously with respect to the user's task, attention, interest or physical location.

1.5 Research Questions and Focus

Based on the challenges expressed in the previous section, this thesis has three main research questions with several sub-questions:

1. *Role of personalisation and context:* What is the role of content personalisation and user context?
    (a) How does content personalisation relate to relevant research fields?
    (b) How do user context and personalisation relate to each other?
2. *Effectiveness of context:* How effective is user context for content personalisation in the mobile event and map content domain?
    (a) How do selected context attributes - time, location and interest - influence users' perception of usefulness?
    (b) How can user context be applied for a personalised information system?
    (c) How effective is user context in a personalised information system for providing useful content?
3. *Domain investigation:* What are the possibilities and limitations of the event and map content domain with respect to mobile use?
    (a) What are the specific characteristics of mobile use?
    (b) What are the specific characteristics of the event and map content domain?

This thesis is focused on content about entertainment events in relation to geographic maps that are delivered and applied in mobile environments by users who access and utilise information from handheld devices. This, however, does not necessarily restrict the findings in this work to this content and usage domain but might well expand into a much larger and more general application.

1.6 Contributions

This thesis contributes in four ways:

1. *Understanding personalisation from different angles:* Personalisation is a concept operationalised with a set of methods and tools (see 2.4 on page 35) that spans across many research fields some of which can be related to information science – like information retrieval [Saracevic, 1999], geographic information science [Goodchild, 1992], or adaptive hypermedia [Aroyo et al., 2004] – or simply relate to information science based on its shared goal of helping people to more effectively manage increasing amounts of information. Research is needed to further the understanding of personalisation and analyse its methods and tools from this multidisciplinary viewpoint. Chapter 2 reviews personalisation based on the fields of adaptive hypermedia, context-aware computing, information retrieval and seeking and geographic information systems. Furthermore, this thesis contributes to the ongoing discussion to extend the perspective of personalisation by considering not only a model of the user but additionally including the user's surrounding context. The related work of chapter 2 reviews current developments in personalisation research with respect to both user modelling and context modelling. This addresses the first research question on the role of both personalisation and context.

2. *Strategy for the development and evaluation of a context model:* An example is presented of the step-by-step development and evaluation of a context model that is used for content personalisation which addresses the first research question on how user context can be applied for a personalised information system. It first demonstrates how a content and usage domain is initially explored with AmbieSense studies in the field of travel and tourism based on the third research question about these two domains. It further shows how these results are used to specify a context model that is then analysed in a laboratory study, formalised into a personalisation model and then verified in a simulated mobile field experiment.

3. *Time, location and interest - evaluation of contextual relationships:* The experiment in chapter 4 investigates a context model for personalised information retrieval of entertainment events in a mobile application environment. To date, there is little evidence that such a context model has been established and evaluated in this depth for the entertainment event content domain. The context model consists of three context attributes (time, location and user's interest). These attributes

are frequently used for context-aware systems but not often subjected to closer investigation. This relates to the second research question; in particular how selected context attributes influence users.

4. *Connection between attribution theory and context:* This thesis reveals a connection between causal attribution theory [Hewstone, 1989], as developed in social psychology and the process of context modelling (see chapter 5). This contribution addresses the second research question on the effectiveness of context; in particular how selected context attributes influence people by providing a possible theory that explains this process. It was demonstrated that causal attribution as a theoretical framework allows context modelling to be viewed as a human process of finding explanations. The theory links to regression modelling based on the covariation principle and causal schema, two theoretical constructs presented by [Kelley, 1973] that explain how people relate effects to potential contextual causes.

1.7 Thesis Overview

The thesis is structured into 7 chapters (including this introduction chapter) in the following way:

Chapter 1 provides an overview of this thesis. The topic of mobile, context-aware personalisation is introduced and focused on event and map-content. Open challenges in current personalisation research are provided. The scoped objectives of this work are listed in connection with the main contributions that this thesis delivers.

Chapter 2 reviews related work and literature on personalisation. The review covers the various aspects of user and context modelling, the acquisition of modelling information and the creation of personalised output. In this respect, particular focus is put on personalised information retrieval and personalised maps – two relatively new areas of research.

Chapter 3 presents results from the AmbieSense project investigating three questions about travellers and tourists; one type of mobile user. Firstly, the importance of different kinds of digital content was investigated. Secondly, knowledge about travellers' information behaviour was gathered; mainly the ways with which they access information. Thirdly, their willingness to provide personal information was researched.

Chapter 4 introduces a context model focused on supporting event content personalisation for mobile usage. A laboratory-based user experiment is described that carefully investigates the effect of context on users perception of usefulness; the experiment targets three context attributes (time, location and user's interest) that are frequently used in context-aware systems. A detailed account of the experiment methodology is presented together with results and a discussion of effects.

Chapter 5 extends the results from the previous chapter and develops a predictive model of context. The model is connected to some of the theory that describes the human process of explanation finding, also called attribution theory. The theory is linked to factorial ANOVA and connected to a basic form of regression modelling. Regression is used to define a predictive context model where a score expresses the amount of usefulness (situational relevance) based on the attributes of the context model - time, location and interest. Within a personalisation model, this context score is combined with a traditional information retrieval score for personalised search.

Chapter 6 reports on the application of the personalisation model from chapter 5 during a mobile field experiment. Users engaged in a mobile environment where they completed simulated search tasks for entertainment events using a mobile application with personalised search. Results showed that context-aware personalised search was able to deliver about 20% more useful content to the mobile user than standard search. The study provided some indication that search tasks were solved faster and with fewer queries when using personalisation. This was achieved in spite of almost natural experiment conditions.

Chapter 7 concludes the studies of this thesis. Research contributions are summarised; limitations and potential future work is discussed.

Based on this more general overview to the structure of this thesis, the next section highlights the various studies from a more methodological viewpoint.

1.8 Overview to the Studies of this Thesis

The following diagram presents an overarching view of the different studies of this thesis and the studies that have inspired and influenced this research. This is a more focused methodological thesis overview in extension to the general overview that was given in the previous section.

Chapter 3 reports on results from a number of *related studies* that took place outside the focus of this thesis but nevertheless inspired and shaped this work. Most of the presented results are from studies conducted in the AmbieSense project and, to a smaller extend, one web study conducted by the Personalization Consortium[3]. The AmbieSense project functioned as a supporting project for this thesis in terms of funding, equipment and expertise but has also influenced this research to a certain extend. These studies explored the content types that mobile users expect, how they acquire, access and use content and how willing they are to provide personal information. Results from this work shaped the remainder of this thesis. The entertainment event content domain was identified as one type of digital content with relevance for mobile users. It was further identified that geographic maps were highly relevant for users in mobile situations. Therefore, it was decided to use these two content types as a focus for this study. Results confirmed the need particularly for context-aware and personalised services that provide and act based on the users' current situation. The studies identified that users are willing to support such systems with personal information. Priorities between different types of personal information emerged in these studies.

The results from chapter 3 significantly inspired and shaped the second part of the study. In chapter 4, a *laboratory user study* aimed to identify the effects of three selected contextual attributes on users' perception of the usefulness of entertainment event content. Besides the interest attribute, time and location were also part of the user study

[3]See section 3.3.3 on page 71 for more details

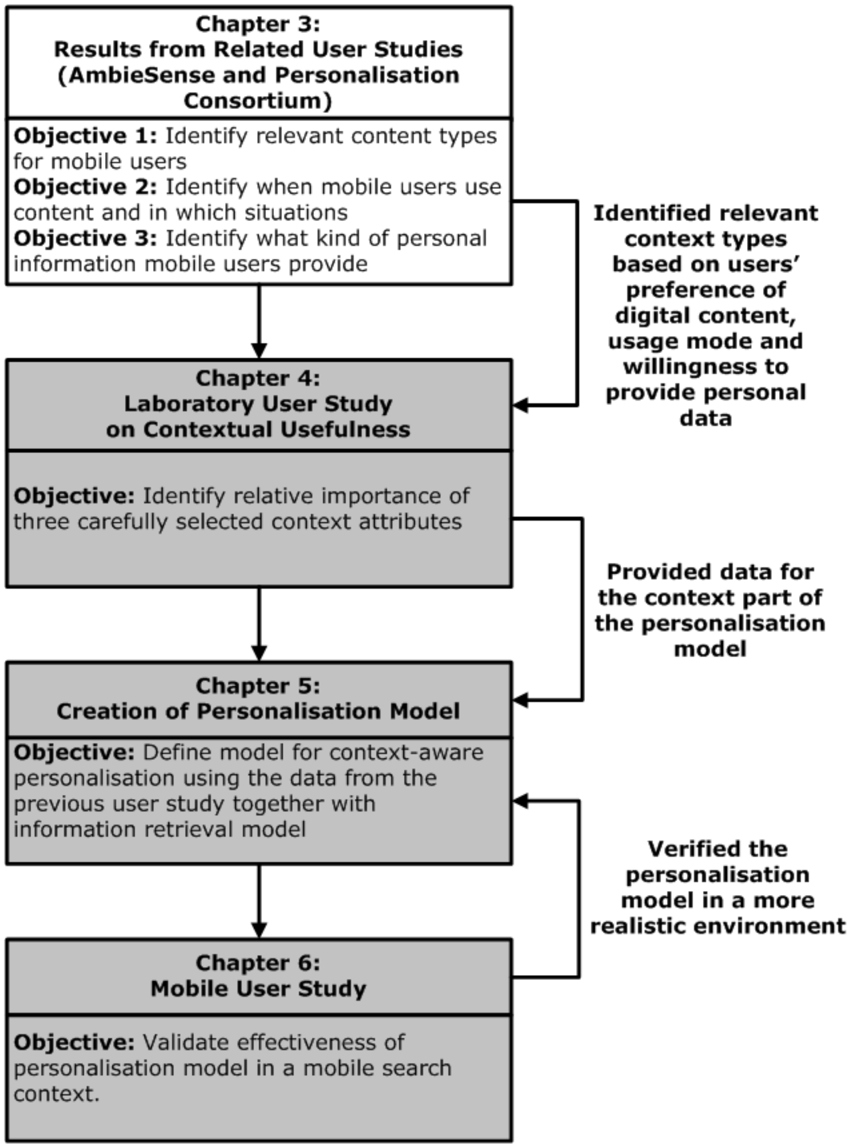


Figure 1.1: *Thesis Studies Design Overview. Grey boxes represent studies that were produced for this thesis. The white box represents all studies whose results inspired and influenced the studies of this thesis.*

that aimed to establish insight onto the nature of three common context attributes. More importantly, their interrelationships were also explored in greater detail.

The results from the laboratory user study reported in chapter 4 enabled the *creation of a personalisation model* that is reported in chapter 5. It merges two different kinds of relevance – the relevance predicted from the data of the laboratory user study and the relevance score from a state-of-the-art information retrieval algorithm. The personalisation model defined a ranking scheme for context-aware personalisation for event content that is connected to geographic maps.

In chapter 6, a *mobile user study* was conduced to measure the impact of the personalisation model on mobile information retrieval. The participants of this study were not travellers or tourists but general mobile users. This aimed to verify the personalisation model in a more realistic environment. As part of the experiment, a mobile event application was designed with a map based interface that delivered personalised event content to participants based on their contextual situation.

2

Related Work

> Thousands of geniuses live and die undiscovered – either by themselves or by others.
>
> Mark Twain

This chapter reviews research on personalisation throughout selected literature. The main focus of this thesis is to investigate the effectiveness of contextual information in the area of event and map-based content personalisation.

2.1 The Process of Personalisation

A personalised information system centres its services around an individual user. Traditionally, an adaptive system maintains a user model or profile that represents information about the user, as shown in in figure 2.1. The adaptive system processes data about the user (e.g. page viewing times) and collects inferred characteristics (e.g. level of topic knowledge and cognitive abilities) in a user model. The system then uses the information stored in the user model to process the adaptation effect (e.g. recommendations). The act of collecting inferred characteristics about a user is called user modelling. The processing of creating an adaptive effect from a user model is called adaptation.

In this thesis, a personalised information system is defined as a system that processes not only user information but also contextual information; for that, Brusilovsky's original figure (shown in figure 2.1) is extended into figure 2.2. On one hand, a personalised

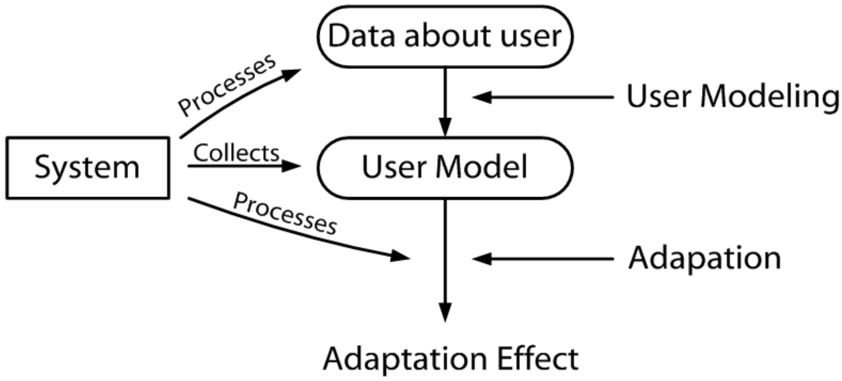


Figure 2.1: *The process of adaptation with an adaptive system (reproduced from [Brusilovsky, 1996]).*

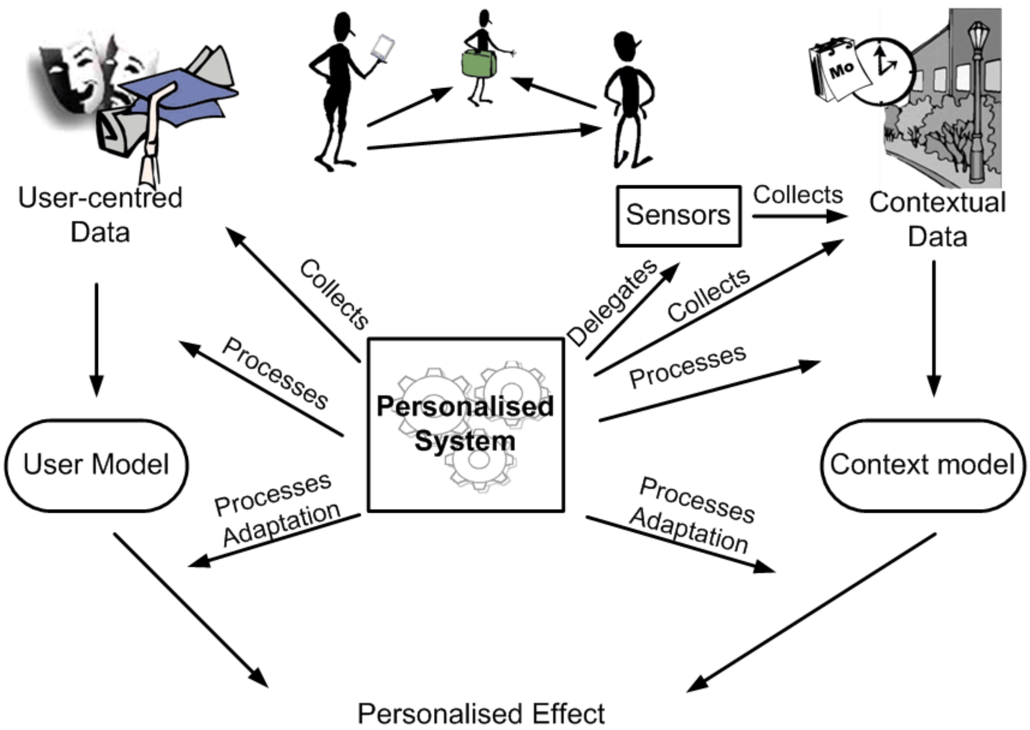


Figure 2.2: *The process of personalised adaptation from the perspective of user and context.*

system collects user-centred data (e.g. user's knowledge, mood) that are processed and stored in a user model; on the other hand, the system also collects contextual data (e.g. spatio-temporal or social context), a task that is sometimes delegated to sensors. This contextual data is then processed and stored in a context model. A personalised system then uses both models in some combination to produce the personalised effect. This thesis first investigates such a combined user context model in chapter 4. This was inspired by the background and related work reported from the Personalisation Consortium and the AmbieSense project in chapter 3. A model for personalised search is established in chapter 5 and evaluated in a mobile application environment in chapter 6. Based on the importance of user and context models for this thesis, the next section takes a closer look at these two types of models and how they relate to each other.

2.2 User and Context Models

A model is an incomplete representation of a part of the real world. Models are focused on particular problems and contain information about a number of entities that are relevant for these problems and their relationships. A model always represents one of possibly many different views to a problem; hence, many models may co-exist that present the same problem in very different ways. This thesis is focused on two types of models: user models (such as defined in [Brusilovsky et al., 2003, p. v]) and context models (such as defined in [Dey, 2001, p. 5]). Generally, a user model represents relevant information about a person, whereas a context model represents information about situations with an emphasis on the surrounding rather than the person. Both the user and the context model are designed based on the problem at hand, most commonly defined by an information system. Both types of models are relevant for the provision of personalised services as described in [Jameson, 2001]. Traditionally, the user model has been closely related with the topic of personalisation. The context model has only recently entered the scene with the appearance of the context-aware system, a new type of system that employs situational information to adapt system behaviour and content. In chapter 4 of this thesis, three attributes are investigated that relate to both the user and the context model. The following subsections describe both types of models with respect to personalisation. Commonalities and differences are then discussed.

2.2.1 User Models

One good definition of what constitutes a user model is provided by Peter Brusilovsky at a recent user modelling conference:

> ”A user model is an explicit representation of properties of individual users or user classes. It allows the system to adapt its performance to user needs and preferences.” [Brusilovsky et al., 2003, p. v]

Similarly, Kobsa defines user models as ”collections of information and assumptions about individual users (as well as user groups), which are needed in the adaptation process” [Kobsa, 1995].

User modelling can be used to understand humans (both as individuals and in groups) – the main objective of psychology. Every human is able to create implicit forms of such (user) models that are essential for daily tasks especially those involving communication [Rich, 1979]. A user model is a prerequisite for adaptive information systems in general and personalised information systems in particular. [Rich, 1979, p. 331] also identified the need for user models as a mean for personalisation.

Work on user modelling started around the beginning of the 1980s with the work of Allen, Cohen and Perrault, as described in [Kobsa, 2001a], and Rich [Rich, 1979]. The emergence of user models is an example of the continuing search for solutions that offer a potential aid for the classic ’information overload’ problem as originally described in [Bush, 1945]. Whereas early systems had no clear distinction between application and user model, systems developed after the mid 1980s started to separate the two. The focus was set to abstract the concept of user modelling and to introduce reusability for future applications and projects. It was not until the early 1990s before these new tools were coined as ’User Modelling Shell Systems’ [Kobsa, 1990]. Examples of such tools are UM [Kay, 1995], DOPPELGÄNGER [Orwant, 1995] and BGP-MP [Kobsa and Pohl, 1995] to mention a few. Such user modelling tools normally provide ways to define user stereotypes and contain a set of inferential techniques to ease the development of adaptive applications. These techniques are mostly applied for the automatic detection of users’ prepositional characteristics expressed as user properties (such as interests or pre-existing knowledge) based on users’ past interactions with the system.

According to [Kobsa, 2001a], the types of user properties and the way they are structured in a user model are largely based on intuition and experience. Kobsa, as part of a larger survey paper [Kobsa et al., 2001], provides a list of user properties as a potential guideline for the creation of a user model:

- *Demographic data* and other objective facts about the user such as name, address (city, area code), sex, age, education, profession and income.
- Information about users' *knowledge and abilities* such as the level of domain knowledge, amount of experience, familiarity with a fact or a concept and the ability to process a certain kind of information or perform a certain activity.
- *Goals and plans* such as users' short time information needs and long time intentions.
- *Interest* such as the strength of emotional involvement with certain product categories such as books, movies, travel destinations, etc. Interest has always been the most frequently applied user attribute in adaptive hypermedia systems [Brusilovsky, 2001] and is the key property for recommender systems. Interest is investigated closer in chapter 4 and applied as part of the personalisation model in chapter 5.

According to [Kobsa, 2001a, p. 53], most user modelling tools focus on determining values for user properties. Usage data is largely employed to determine properties but not as a separate entity that is modelled on its own.

2.2.2 Context Models

Numerous definitions of context exist. Schilit's paper on context-aware computing defines context as "where you are, who you are with, and what resources are nearby" [Schilit et al., 1994]. This suggests that context is focused on the user's surrounding as opposed to the user model that focuses on their inner states. In [Morse et al., 2000], context is described as "implicit situational information"; Schmidt goes beyond that and defines context as "interrelated conditions in which something occurs" [Schmidt et al., 1999b] pointing out possible links between context attributes. A more precise and complete definition is provided by Dey in a special issue on context where he defined it as "any information that can be used to characterise the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and application themselves" [Dey, 2001, p. 5]. It becomes apparent from these more or less informal definitions that context generally "suffers from the generality of the concept" [Schmidt et al., 1999b, p. 3]. Context as a term is used vaguely in everyday language and also has a range of different meanings in information and computer science where it is also used to describe aspects of human-computer interaction and elements in natural language processing [Schmidt et al., 1999b].

In parallel with the recognition of context in mobile and ubiquitous computing, context also emerged in other research areas. In a special journal issue, Cool and Spink [Cool and Spink, 2002] gave an overview to the various ways in which context relates to information retrieval. Two workshops on Context in Information Retrieval [Ingwersen et al., 2004, Ingwersen et al., 2005] provided a platform for the discussion of ideas and applications about context and information retrieval. A number of theoretical models exist for Information Retrieval and Seeking (IR&S). These models are generally focused on the concepts of interaction and context and are closely tied to the user of the retrieval system.

- *Ingwersen's Cognitive Model:* As described in [Saracevic, 1996], Ingwersen pioneered and promoted the cognitive IR&S model [Ingwersen, 1992, Ingwersen, 1996] that views IR interaction as a set of cognitive representations and processes. User's interact according to their cognitive space which is defined along a set of different factors (what Saracevic calls 'structured causality elements') embedded in the users personal context that closely resembles a user model. This cognitive space interacts with the social/organisational environment of that user that represents environment information as understood in user modelling or environment context as described later. As summarised in [Wilson, 1999], the model represents the different types of dynamic and interactive transformations that occur between the users experience of the problem and the search.

- *Belkin's Episode Model* is based on the paradigm of an "information seeking episode [that] consists of a series of kinds of interactions (slices of time) structured according to some plan associated with the person's overall goals, problem, experience, [...] goals..." [Belkin, 1996, p. 5]. The model also includes aspects from Belkin's earlier model on users' anomalous state of knowledge (ASK) [Belkin et al., 1982], a state where a user recognises a lack of knowledge about a particular area or topic but has difficulties expressing it in a precise way. ASK occurs in the users current contextual situation. The more specific episode model integrates the users' current contextual state as part of the information seeking process; furthermore, Belkin promotes the existence of cognitive scripts or plans that structure the information seeking procedure.

- *Saracevic's Stratified Interaction Model* [Saracevic, 1997], inspired by human-computer interaction, represents interactive information retrieval based on a user and a system side that are connected through an interface. Each side is divided into different levels or strata; the user side incorporates a strong contextual viewpoint that is divided into three different levels. The *cognitive* level deals with the ways users organise and structure information mentally such as their state of knowledge and how they infer relevance from it. The *affective* level handles users' intentionality such as motivation, feelings and desires. The *situational* level represents users' surrounding situation that triggers their information needs that are put forward to the retrieval system.

Besides information retrieval and seeking, the importance of context and its relevance in relation to user modelling has also been identified for ubiquitous computing with a recent special issue by Jameson and Krueger [Jameson and Krueger, 2005].

A few years after the first context-aware systems emerged, context management systems followed. The *Context Toolkit* [Salber et al., 1999] was one of the first system architectures that supported the development of context-aware applications. It provided a reusable framework of context widgets – small software components that encapsulated context and its acquisition for an easier integration in existing applications. An overall context model was not provided by the toolkit and had to be defined by the application. Others followed quickly such as the server based Context Managing Framework [Korpipää et al., 2003], the SOCAM system [Gu et al., 2004], CASS [Fahy and Clarke, 2004], the distributed

CoBrA (Context Broker Architecture) [Chen and Finin, 2003] and the system developed for the Hydrogen project [Hofer et al., 2002] – all comprehensively described in the most recent survey on context-aware systems [Baldauf et al., 2007].

There has recently been increasing interest in context modelling [Indulska and Roure, 2004]. The information seeking community in CoLIS 2005 [Crestani and Ruthven, 2005] investigated in particular theoretical approaches for better understanding and modelling of context and has been followed by the First and Second Symposium on Interaction in Context [Ruthven et al., 2006, Borlund et al., 2008]. Context modelling is motivated by a general need for theory about context and its structures that will consequently help in building better frameworks and more effective systems. A number of surveys explored existing context models [Kaenampornpan and ONeill, 2004, Strang and Linnhoff-Popien, 2004, Baldauf et al., 2007]. In [Baldauf et al., 2007] and [Strang and Linnhoff-Popien, 2004] context models of various systems were investigated, mainly from their technical aspect such as architectures, data formats, communication protocols and the use of standards. However, the type of contextual information that the model contains (e.g. attributes and their relationships) is perhaps more important than the technical structure of the model. The choice of the kind of attributes and their values directly effects the performance of the system using it. The survey provided by [Kaenampornpan and ONeill, 2004] reviews context models based on the information they model and relate them to each other. In the following list, a selective number of models is described:

- In [Schmidt et al., 1999b] a context model is defined by two categories – human factors and the physical environment. The human factors are categorised into user information, the user's social environment and the user's task. Attributes describing the physical environment are divided into user's physical location, the available infrastructure and the physical conditions around the user.

- In [Chalmers and Sloman, 1999], a context model is defined along attributes of location (in the sense of positioning and proximity), device characteristics, environment and the user activity. The emphasis on environment and device is due to the focus on Quality of Service solutions for mobile applications.

- The User Context model proposed by Myrhaug and Göker [Myrhaug and Göker, 2003] divides the contextual spectrum into the five categories of environment context, personal context, task context, social context and spatio/temporal context. The environment context captures the entities surrounding the user. The personal context models the attributes of the user and is further categorised into the physiological context and the mental context. The task context represents all attributes that describe what the user is doing. The social context models a user's social relationships. The spatio-temporal context describes user's location and time related attributes including the potential movement in relation to other entities of the user's environment (i.e. buildings or vehicles).
- For the purpose of map personalisation, Reichenbacher defines context into the six categories of situation, user, user activity, physical environment, information and system [Reichenbacher, 2007].
- Zipf's earlier work on map personalisation [Zipf, 2002] identified relevant context attributes for map personalisation consisting of attributes about the user's physical condition, the weather, the user's task, user's cultural background and others.

There are a number of papers that do not propose a context model, but contribute ideas and suggestions for context modelling. The context attributes that are considered relevant in [Hull et al., 1997] are user attributes (i.e. health, identity) and attributes about the users' environment (i.e. location, time, computing resources, physical environment). Lieberman and Selker distinguish context into system, user and task [Lieberman and Selker, 2000]. Whereas the system context is defined by the system implementation, the user context consists of the user state, history of past activities and preferences. The task context is defined by goals and actions. In [Lucas, 2001], context is categorised into physical context, device context and information context. The paper has an overall strong focus on spatial aspects where the physical context refers to the location of environment features, the device context refers to device attributes and the device location and the information context also focuses on the location of information objects and their proximity to each other.

In the following, the context categorisation provided by [Myrhaug and Göker, 2003] is used as one example of a comprehensive context model. Its categories are used as a guideline to link context categories and attributes of other models:

- *Spatial context*[1]*:* The most important attributes of a context model are those that capture the spatial aspects of a situation. *Location* is the most common aspect of a spatial context and used in almost every context-aware system. In particular, mobile guides such as Cyberguide [Abowd et al., 1997], GUIDE [Cheverst et al., 2000] and the CRUMPET system [Zipf, 2002] belong to a special class of applications that are commonly referred to as location-based services. Based on its relevance, location modelling has emerged as one research branch in location-based services; a comprehensive overview is provided in [Jiang and Yao, 2007]. Many location-based services employ location for the personalisation of geographic maps[2]. Location can be represented either geographically or semantically [Beigl, 2002]. The geographic representation exhibits locations by its position (i.e. coordinates provided from the Global Positioning System (GPS)). On the other hand, the semantic representation describes locations in a more descriptive and humanly understandable way, yet still able to be processed by a computer. The comMotion system [Marmasse and Schmandt, 2000] for example learned meaningful locations semantically by analysing users' GPS logs over time. Besides location, spatial context also models attributes such as the direction of movement, the viewing direction and the speed of movement.

- *Temporal context:* Temporal context refers to time and is generally identified as at least as important as location [Reichenbacher, 2007]. Temporal context may be represented absolutely as a measurement or less precisely in a more semantic way (e.g. 'in the evening' or 'before a meeting'). In [Hull et al., 1997], time is part of the users' environment whereas [Göker and Myrhaug, 2002] view it as a separate context type that is tightly bound to spatial conditions due to the focus on mobility. Similarly, [Reichenbacher, 2007] views it together with location as part of a situation. In [Schmidt et al., 1999b], temporal context is related to all

[1] The user context model described in [Myrhaug and Göker, 2003] originally categorised spatial context and temporal context combined into a single category as spatio-temporal context. For reasons of presentation, it has been separated into two distinct categories.

[2] A relevant selection of LBS'es will be described in section 2.5.3 on page 50 in more detail.

other context attributes representing the contextual change of those attributes over time. In [Bradley and Dunlop, 2004, p. 3], temporal context is described as being "embedded within everything, and is what gives a current situation meaning...".

- *Personal context* is equal to the user information that is stored in a typical user model as discussed in section 2.2.1. It can be distinguished into user's physiological context (e.g. age or body weight) and user's mental context (e.g. interest). This is similarly described in [Schmidt et al., 1999b, p. 3], that states that "information on the user" comprises for instance "(knowledge of habits, emotional state, biophysiological conditions, ...)". Attributes that are represented by the mental context can generally be found in user models. Examples include user's identity, preferences, knowledge and skills as listed in Reichenbacher's context model [Reichenbacher, 2007]. User's interest is one of the most important attributes that is commonly modelled in most user models and has been widely applied in personalised information systems [Brusilovsky, 2001].

- *Environment context* captures the entities around the user. It includes objects of the surrounding environment (e.g. buildings, outdoor facilities, infrastructure) and their state (i.e. temperature, light, humidity, noise). It also describes information – what Lucas calls information context [Lucas, 2001] – as part of the environment. Furthermore, devices that reside in users' vicinity are also accounted for as part of the environment. Attributes of these devices define the extent with which personalisation can be performed [Chalmers and Sloman, 1999]. Examples includes the processing power, screen size/resolution, colour support, sound capabilities and the kind and number of input devices. Similar to [Göker and Myrhaug, 2002], Schmidt also categorises contextual information about device(s) as part of the physical environment [Schmidt et al., 1999b].

- *Task context:* The task context contains information about what the user is doing or aiming for. It describes "the functional relationship of the user with other people and objects" [Bradley and Dunlop, 2004], including the benefits and constraints of this relationship. It can be modelled as explicit goals, actions and activities. User's activity is a common type of context in a number of context models [Reichenbacher, 2007, Chalmers and Sloman, 1999].

- *Social context* represents the social environment of a user. It may describe a user's relationship to like-minded people that are connected to the user. Social filtering systems[3] implement one special kind of social context modelling [Griffith and O'Riordan, 2000]. The recommendation output is solely based on a user being similar to other users who rated items previously. This social connection is then exploited to recommend more items. Social filtering systems have demonstrated good results for a range of relevant topics such as music [Shardanand and Maes, 1995] and news [Resnick et al., 1994]. Social context may model a user's list of friends or colleagues, perhaps explicitly expressed by that user or implicitly acquired through an email address book or buddy list on a website.

In the past decade, many context-aware systems have been developed, often in the form of mobile guides and focusing primarily on location sensing. Different frameworks for context management were established and provided the technical support for the design and the development of context-aware systems. Several context models were established, as the previous pages have shown; however the literature shows a lack of empirical work that investigates these context models and their attributes. More work is required to more closely evaluate the influence of context attributes on users of context-aware systems with respect to users' information seeking behaviour. Contextual effects that result from a number of possibly interrelated attributes also need to be understood. Attributes may not only occur in isolation but may be connected. Understanding their nature would allow more effective systems to be built. This thesis contributes with empirical investigative studies in an attempt to deepen the understanding about the dynamics of context. In chapter 4, three context attributes are analysed with respect to how they affect peoples' perception of the usefulness of information: *Location* as a central attribute of spatial context; *time* as the key attribute of temporal context; and *interest* as the main attribute of personal context that has been widely used for personalisation. The three attributes were inspired by results obtained from the AmbieSense project (see chapter 3). The User Context model proposed in [Myrhaug and Göker, 2003] was used to structure these attributes; it is particularly suitable since it provides a wide range of general context categories and gives enough room for specialisations and refinements. In chapter 5, results

[3] Also referred to as collaborative recommender systems or more general as collaborative filtering.

from this study are further connected with theory and used to establish a personalisation model for personalised information retrieval that is then evaluated in a mobile user study in chapter 6.

2.2.3 User and Context Models: Commonalities and Differences

The previous two sections reviewed various user and context models in relevant literature. This section highlights their commonalities and points out their differences. Two main aspects have been identified in this respect:

- *History:* User modeling and context modeling developed in different research areas. User modelling originated from ideas in psychology and cognition at the end of the 1970s / early 1980s. Later, user modeling found wide application in adaptive hypermedia that resulted in a large number of adaptive and personalised systems. On the other hand, context modelling mainly originated from mobile and ubiquitous computing in the 1990s mainly initiated by the technological boom in personal digital assistants (PDAs). Based on this development, first applications were mainly location-based systems and services.

- *Focus and Aim:* The most dominant and obvious difference between the two kinds of models is the data they use to build the model. User models are created solely from collected user data (i.e. explicit and implicit user feedback), whereas context models are generally focused on data sources that describe aspects of the surrounding (e.g. the geographic location). The strong initial focus of context models on location was later enriched by other contextual attributes. However, the main focus largely remained on surrounding information gathered by sensors. According to [Byun and Cheverst, 2001], this makes context models predominantly data models whereas user models can be data models as well as behavioural models.

Despite the fact that user and context models maintain their focus on different kinds of information, there is a current trend to integrate elements from each type of model. On one hand, many user modelling systems have started to model aspects from the surroundings of the user and therefore widen their initial approach. On the other hand, literature also shows that context models have expanded from their technical, measurement and sensor-based viewpoint (i.e. device and positioning information) to a viewpoint that also includes the user as part of the surroundings. Some context models are labelled 'User Context' in at least two recent papers [Göker and Myrhaug, 2002, Tazari et al., 2004]. This thesis also contributes to furthering the understanding of a combination of both models for the purpose of personalisation and presents an example of an investigation into the effectiveness of user context based on a number of selected attributes in a series of studies.

2.3 Data Acquisition for Modelling

This section reviews different strategies and technologies applied to gather data for the population of user and context models. In user modelling and adaptive hypermedia, data acquisition mainly focuses on the users and how they use information systems [Kobsa, 2001b]; however, in context-aware computing the emphasis is usually on the acquisition of information that describes the surrounding environment. In this section, both areas are addressed. Note that in this thesis, the data that were used for modelling were not collected with a specific information system but instead with an experiment where participants provided explicit ratings of usefulness for content based on simulated situations. This method of data collection relates closely to the explicit information acquisition in section 2.3.1 below. A specific application however could very well also employ implicit methods as described below and/or use one or more of the methods and techniques described in section 2.3.2 to obtain contextual information about the user's environment.

2.3.1 Acquiring User and Usage Information

From the perspective of user modelling, there are two common ways of acquiring information from the user: explicit and implicit.

Explicit Acquisition

Explicit techniques prompt the user for information; for example, the user completes an initial questionnaire when signing up for a personalised service as described in

[Rich, 1979]. In a wider sense, *explicit personalisation*[4] refers to a personalisation style where users enjoy high levels of control over the entire personalisation process. Explicit styles therefore not only affect the model acquisition process but also the production of the personalisation output (as described later in section 2.5). There are numerous examples of explicit personalisation practices. On the MyYahoo! portal [Manber et al., 2000], users manually select content categories thus explicitly indicating interest. This information is used by the personalised portal to select and present similar news stories. Personalised Google[5] as well as many other sites, provide localised weather information based on postcodes. Users of LIBRA [Mooney and Roy, 2000], a content-based book recommender system, manually rate books and receive book recommendations as personalised output. In general it is very intuitive to acquire user information in explicit form. Many early systems have used the explicit style, such as Tapestry [Goldberg et al., 1992], Ringo / Firefly [Shardanand and Maes, 1995] or Fab [Balabanovic and Shoham, 1997].

The main *advantage* of explicit model acquisition is system transparency. Explicit acquisition of information for a personalised service from the user allows them to relate the cause (e.g. input of interest categories) with the effect (e.g. output of personalised news stories) and to become more confident and comfortable with the system. A user study on relevance feedback [Koenemann and Belkin, 1996] and the effect of different levels of system transparency showed that users obtain more relevant results from a system that offers users more control. The study also demonstrated that users generally prefer systems that offer more information and control about the feedback process. These aspects have room for further investigation[6].

As a *disadvantage*, the user has the total responsibility and the full workload of providing accurate feedback. Humans tend to base their feedback on subjective measures such as the presentation rather than the content; they may also be influenced by emotions and other personal factors. More generally, relevance is a widely debated concept that is understood as a multidimensional and dynamic concept that follows systematic and measurable patterns as summarised in [Schamber et al., 1990, Borlund, 2003a, Ruthven, 2005]. The

[4] Sometimes also called 'customisation' or 'checkbox personalisation'.
[5] `http://www.google.com/ig/`, accessed April 14, 2008
[6] Based on suggestions by Diane Kelly [personal communication].

multidimensional aspect points out that relevance is in fact not one, but a multitude of different relevances that coexist and play their part in the judgment process. The dynamic notion of relevance refers to the fact that users judge information at a certain point in time which continually changes. However, relevance is still a systematic and therefore a measurable concept, although complex in general. In practise, relevance is usually recorded on a single scale which limits the amount of feedback that can be acquired from the user. Self evaluation may be difficult for users when, for example, being asked about existing knowledge. Another important issue is that most user-centred information becomes eventually outdated (i.e. interests) or even invalid (i.e. place of residence) and needs to be refreshed. Most systems deal with these issues by requesting users to regularly confirm their data. A step forward in this direction is the use of implicit acquisition techniques.

Implicit Acquisition

Implicit techniques gather knowledge about the user in an indirect and unobtrusive way. Since users may not be fully aware of the full extent of the data collection process, implicit data acquisition almost always sparks a debate on privacy and trust; this is addressed in more detail in appendix A.1.

In a wider sense, *implicit personalisation* refers to a personalisation style where users experience the personalisation features of a system while only being indirectly confronted with the acquisition process of user data. Therefore, implicit techniques minimises the interface between user and system for easier data collection. Liebermann's Letizia system [Lieberman, 1995] built a user profile based on the sites the user visited earlier. New webpages were recommended based on their similarity to the user model. The Amazon website provides personalised product recommendations based on users' purchase history.

Implicit feedback and its application for personalised information systems was reviewed in [Nichols, 1997] and [Oard and Kim, 1998] providing a categorisation for different kinds of implicit feedback. [Oard and Kim, 1998] differentiates observable user feedback, earlier surveyed and identified by [Nichols, 1997], in the categories of "Examination", "Retention" and "Reference" later further refined in [Oard and Kim, 2001]. In [Kelly and Teevan, 2003] 30 papers were reviewed based on how well they addressed these categories.

Implicit acquisition techniques have a range of *advantages*. Firstly, the effort of data collection is minimised and shifted to the system. The user applies the information system normally and has no additional obligations. This also means that implicit methods can be applied in situations where users are generally unwilling to provide information (i.e. in a busy work environment) or not able to provide them (i.e. on mobile devices with limited ways of input). Secondly, it is easy to obtain large amounts of usage data that potentially provide many clues for the identification of user properties. Thirdly, it is easy to use implicit acquisition techniques in a client-sided personalised system as it provides direct access not only to users' actions but also to other system parameters.

As a *disadvantage*, implicit acquisition techniques are generally regraded as less effective than their explicit counterparts [Nichols, 1997] (cited by [Kelly and Teevan, 2003]). For example, results from the ANATAGONOMY system, a personalised newspaper, demonstrated implicit user feedback as being less effective for personalised page generation in comparison to an explicit method [Sakagami and Kamba, 1997]. Furthermore, data that is collected from users' behaviour may be ambiguous or even contradictory; even so, the fact that it can be easily collected and additionally enriched with explicit feedback turns implicit acquisition into a very promising method.

2.3.2 Acquiring Contextual Information

Like a user model, a context model needs to be populated with accurate and up-to-date information. There are two common methods with which context can be acquired. Firstly, context information can be gathered from hardware that is located on the device or in the users environment. Secondly, data can also be acquired directly from the user – as described previously in section 2.3.1. The following subsections analyse and discuss the acquisition of different types of context guided by the context categories that have been described in section 2.2.2.

Acquisition of Spatial Context

Location is the most applicable context attribute and widely used for Location-Based Services – one type of context-aware system. For that reason, the detection of spatial context is one of the most common types of context acquisition. For the low level determination of the users' current location, a wide range of different positioning technologies are used. These technologies depend on wether the application operates indoors or outdoors.

For *outdoor positioning*, the Global Positioning System (GPS) is widely used. GPS is a globally available satellite navigation system that allows GPS enabled devices to determine their position, speed and the direction of movement. Its use is free of charge and a large range of available hardware products enables its use with mobile devices. The Cyberguide system [Abowd et al., 1997], a tourist guide for visitors of a research centre, used GPS in outdoor environments. The location-based system comMotion [Marmasse and Schmandt, 2000] employed the technology to detect users' movement patterns and determines meaningful locations in order to associate personal information (i.e. todo lists) with users' current location. However, in indoor environments GPS lacks reliability and accuracy. Its low signal strength is easily blocked by most buildings and additionally disturbed by reflections.

For that reason, different technologies emerged that cope with the problem of *indoor positioning*:

- *Infrared* technology (IrDA) provides a simple and cost effective method for point-to-point indoor positioning. The basic operation of an IrDA system is based on a sender that transmits light impulses to a receiver within a short range of a few meters. The signal also identifies the sending device, which allows the receiver to map the sender to a location. Due to the optical connection, the signal stays within the limits of a room but is also easily interfered with small obstacles. Its low power consumption makes it particularly attractive for mobile devices. The Active Badge system [Want et al., 1992] was a system that routed telephone calls based on users current locations. It consisted of a small mobile device that propagated an infrared signal into the environment that was then picked up by sensors to determine the position of the device and therefore the user. Infrared also found application for ParcTab [Schilit et al., 1994], a prototypical system that investigated the abilities of mobile computers in an office environment. The Cyberguide system [Abowd et al., 1997] employed infrared in combination with GPS to cover both indoor and outdoor situations.

- *Radio Frequency Identification (RFID)* is an emerging technology that is mainly used for security access and product identification/tracking, but may also be used for positioning. RFID positioning is facilitated via hardware identifiers (RFID tags) that are mounted in the environment and a radio signal that is emitted by the sending mobile device and reflected by the RFID tag. Depending on the type of tag, RFID operates in the range between centimeters up to a few meters and offers a reliable signal that stays in the vicinity of the location. In [Mantyjarvi et al., 2006], RFID has been used for a museum guide to recognise a users proximity to an artwork. The LANDMARC positioning system [Ni et al., 2004] uses arrays of RFID tags and reports positioning accuracies of 1 meter (50% precision).

- *Ultrasound* technology provides an alternative to systems that operate on IrDA and RFID. Its signal is relatively slow (1,200 km/h), does not penetrate walls and does not require an optical connection between sending and receiving device. This makes it particularly useful for indoor positioning with high accuracy in the range of a few centimeters. On the other hand, it is not appropriate for large spaces as the signal is very restricted in range. In the ActiveBat system [Ward et al., 1997] users carried a device that periodically emitted ultrasonic impulses when triggered by a central controller. The signal time was determined by a grid of receivers mounted on the ceiling that allowed the 3D position of the user to be determined with very high precision. The Cricket system [Priyantha et al., 2000] used small beacons that emitted environment information plus ultrasonic impulses. Mobile devices received these signals and calculated their position based on environment information and the signal travel time.

- A *Wireless LAN* may also be used for positioning. It establishes a cell around the wireless access point that offers mobile devices connections with minimum setup time. Devices may be operated stationary or at walking speed. Location accuracy is generally less than infrared and RFID as the signal penetrates walls and therefore extend over several rooms, floors or even buildings. The GUIDE system [Cheverst et al., 2000] used Wireless LAN for positioning based on the user's current (and presumably closest) connection to a Wireless LAN access point to provide city visitors of Lancaster with contextualised content. The RADAR system [Bahl and Padmanabhan, 2000], developed by Microsoft Research, achieved a positioning accuracy in the range of 2-3 meters. In [Xiang et al., 2004], a Wireless LAN based indoor positioning technology is presented and tests reported a similar positioning accuracy between 2 and 5 meters depending on the degree of movement with 90% probability.

- *Bluetooth* is aimed for wireless communication to substitute cable connections between devices at short distance. It is meant for stationary use with no or little movement. The connection time is longer than with Wireless LAN and more sensitive to the number of devices in the surrounding[7]. Bluetooth operates more reliably in a particular room or part of a building than Wireless LAN as the signal does not easily penetrate walls. Bluetooth was used for the detection of users' location for the mobile tourist guide described in [Myrhaug and Göker, 2003] for both indoor and outdoor situations. The commercial Ubisense system[8] enables people to be accurately positioned within the range of centimeters. A Bluetooth triangulation positioning system was developed for the Alipes project [Hallberg et al., 2002]. The BIPS system [Mantyjarvi et al., 2006] employed Bluetooth in combination with Wireless LAN for positioning in order to pinpoint and guide mobile users through a building.

Acquisition of Other Context Types

Unlike spatial context, temporal context is easy to obtain by using the built-in system clock that is usually available in every mobile device and provides the precise time. The precise time measure further allows a system to infer more semantic levels of temporal context such as the time of the day ('early morning' or 'late evening'), seasons ('summer'

[7]A recent IEEE article [Ferro and Potorti, 2005] reports average network setup times of 5 + n * 1,28 seconds where n represents the number of devices in the environment.

[8]`http://www.ubisense.com`, accessed April 14, 2008

or 'winter') or special days of the year (i.e. user's birthday or bank holidays). Often temporal context is used in relation to spatial context. The GPS signal, for example, is combined with a high precision time signal that allows enabled systems to determine users' speed of movement.

A wide range of sensors exist to acquire context about the environment, the person, the social surrounding and the task. In [Schmidt, 2002, chapter 3], a comprehensive list of sensors is provided that are suitable for the detection of these context types for mobile computing. Among others, sensors are listed for the detection of light, audio, temperature, humidity, air pressure, movement and acceleration, magnetic fields and orientation and touch. 'Smart-Its' [Gellersen et al., 2004] provides a sensor board with an integrated collection of sensors to measure and communicate motion, audio, light, pressure and temperature.

Light and *audio* sensors are small, cheap and reliable; they provide an easy way for the automatic acquisition of contextual information about the near environment of the user. The TEA project [Schmidt et al., 1999a], an EU-IST project that contextually investigated user activities, demonstrated that audio can distinguish human voices from other sounds such as music or noise. The measurement of *temperature* can be easily performed by low cost sensors providing reliable clues about the state of the environment. In combination with *humidity* and *air pressure* sensors, this can provide information about the current weather conditions in outdoor situations. *Movement and acceleration* detectors may provide coarse-grained clues about users' physical movement, however, details about what the user is currently doing are best obtained by implicit indicators as discussed previously. For example, electronic calendars, often provided on PDA's and mobile phones, offer a rich source of context information about the user's current task. An electronic compass helps determine the *orientation* of the device and the user; the mobile guide described in [Rantakokko and Plomp, 2003] applied such a sensor to adapt the orientation of a map display. Furthermore, a number of technologies for spatial context acquisition can also be used to determine the device orientation. A *touch* sensor that measures the conductance of the skin, pulse or body temperature allows the recognition if a device is used; a potential high-level indicator for the user's task. In a case study, descried in [Schmidt and Laerhoven, 2001], a mobile phone equipped with touch sensors recognises if it is held in the user's hand. This sensor can also be used to collect information about the physiological state of the user; furthermore, it may hint at information about the mental state of the user. For example, a high pulse may indicate anxiety. The sensing of information about personal emotions is particularly targeted by the Affective Computing Laboratory at MIT [Picard, 1997].

2.4 Techniques for Model Acquisition and the Creation of Personalised Content

In this section, a short overview is provided of the most common techniques used for model acquisition and for the creation of personalised output. The section starts with a brief overview of the most common and most general statistical techniques, one being used in chapter 5 as the underlying technique for representing context as part of a personalisation model. Literature is then reviewed based on two common styles of how personalised information systems acquire models and how they produce personalised output: content-based and collaborative. Information retrieval and filtering techniques are common as content-based methods that acquire a model from the content a user consumes. In this thesis, information retrieval is used as the main technique for one part of the personalisation model (see section 5.4). The personalisation model is then used for the mobile user experiment in chapter 6 to evaluate its performance for personalised information retrieval.

Although collaborative techniques were not used as part of this work, they provide an interesting alternative for future work. Collaborative methods are specifically aimed for the acquisition of modelling information from the feedback provided by user groups.

2.4.1 General Statistical Learning Techniques

According to the overview provided by [Kobsa et al., 2001], statistical techniques became increasingly popular for user modelling and have been commonly applied to user modeling problems, as [Zukerman and Albrecht, 2001] shows:

- *Linear models* use an additive, linear combination of variables in combination with weights for the prediction of an outcome, i.e. the value of a user property. Regression represents one variation of a linear model and was used for the prediction of usefulness based on context as described in chapter 5.

- *Markov models* present sequences of events where the prediction of a future event is based only on a limited number of previous events. The DOPPELGÄNGER system [Orwant, 1995], for example, applied Markov models to model user behaviour.

- *Classification* partitions information objects into classes based on the information they contain. Classes are not defined beforehand but are based on the closeness of information objects to each other in a multidimensional hyperspace.

- *Rule-based methods* learn rule sets that are able to classify observations from its attributes. Such a rule-based system then expresses an explicit representation of knowledge – normally represented as decision trees or probabilities.

- *Bayesian networks* are acyclic, directed graphs where each node represents a variable and each edge represents a causal link between them. Each node has a conditional probability distribution that represents the probability of this node for each based on each value of the parent nodes. Bayesian networks are therefore extensible and may contain many variables.

2.4.2 Content-based Filtering versus Collaborative Filtering

Techniques for the acquisition of data that is used to populate a user or context model are usually distinguished into two groups – content-based filtering and collaborative filtering – as highlighted more generally in [Belkin and Croft, 1992] as well as in [O'Riordan and Sorensen, 1999] from a more technical perspective.

Content-based filtering

Content-based filtering techniques employ the content a user is reading, working with or otherwise consuming in order to learn a model. The underlying idea is that the content information represents a model of the users' future information need. The literature shows that content-based techniques are usually from the two areas of information filtering and information retrieval:

- *Information filtering* (see figure 2.3) works on the basis of incoming, dynamic streams of mostly textual information (i.e. a news feed or a mailing list). In such a scenario, it is more important to remove non-relevant content from the stream rather than selecting relevant content for the user. On the other hand, the user has a long term information need or interest that is formalised in a profile. This profile is built from implicit or explicit feedback and used for the process of filtering. The filtering itself is a continuing process over a longer period of time in which the profile might be adapted to the changing long term information needs of the user. For information filtering systems, timeliness of content delivery is often of particular importance as content is updated. Users of information filtering systems are usually more passive and acting as information consumers. Issues of privacy and trust are repeatedly debated mainly because filtering systems tend to collect and maintain private data in profiles. These issues are discussed in a wider sense in appendix A.1.

- *Information retrieval* (see figure 2.4), unlike information filtering, is based on more static sets of information (i.e. journal articles, but also web pages). It is therefore historically more concerned with issues of information management and storage (see [Belkin and Croft, 1992, p. 32]) and has a much stronger connection with the field of library and information science as described in [Saracevic, 1999, p.1]. Since information is more static, information retrieval is more concerned with the extraction of relevant information rather than the removal of the irrelevant. For that, the user employs a query to express an information need that has generally a more limited scope and validity than a profile (normally not exceeding a session) and is more add-hoc than a profile. An information retrieval system involves the user more strongly as an active partner and therefore demands higher knowledge, but also provides more control over the retrieval process. Due to the more static nature of the content, content is generally less time critical. Traditionally, users of IR systems are generally less concerned about questions of privacy and trust.

Information filtering and information retrieval, although different in how they apply and treat content and users, share their techniques. The reason for that is the long tradition

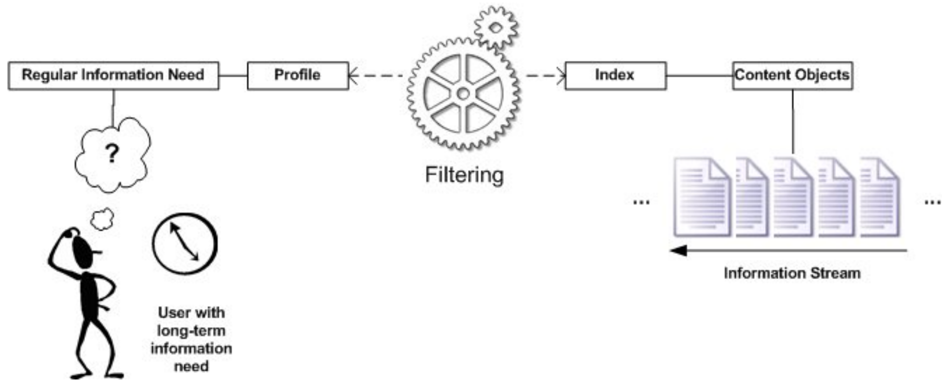


Figure 2.3: *Information Filtering Process (adapted from [Belkin and Croft, 1992])*

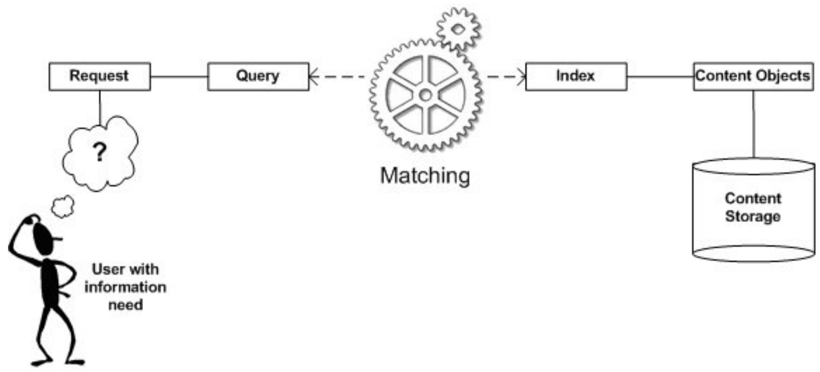


Figure 2.4: *Information Retrieval Process (adapted from [Belkin and Croft, 1992])*

of information retrieval in comparison to the young discipline of information filtering. In a way, information filtering is a specialisation of information retrieval and mainly supported by new application areas [O'Riordan and Sorensen, 1999] – such as adaptive hypermedia [Aroyo et al., 2004]. The literature often connects them because of their similar goal [Faloutsos and Oard, 1995] or labels systems interchangeably as information filtering or information retrieval systems based on the context of application[9]. This means that many content-based techniques are both used for information filtering and information retrieval, whereas historically they may originate from information retrieval.

Most modern information retrieval systems are based on either of the most common models – the *Vector Space Model* [Salton, 1971] or the *Probabilistic Model* [Maron and Kuhns, 1960]. It is still debated which of the models performs better

[9]In [O'Riordan and Sorensen, 1999], the SMART information retrieval system is listed as a information filtering system.

and a clear conclusion is still missing [Grossman and Frieder, 2004]. It is argued that comparisons are usually performed on the system level that incorporates not only a particular model but also numerous adaptations and tools. This makes it very difficult to compare the fundamental, underlying principles of both basic models.

The information system that was used in the mobile user study described later in chapter 6 applied an information retrieval system in combination with context to enable personalised search. For the purpose of investigating the effectiveness of context as a means to personalise content for mobile users, the Lucene search engine API[10] was applied as one implementation of the vector space model. The API was used for the development of the mobile information system that was applied in the experiment described later in chapter 6. The decision to use Lucene is motivated by the technical environment in which this search engine was applied. The search engine was executed on a Sharp Zaurus 5500 PDA. Although the PDA was up-to-date at the time of the experiment, for today's standard it is of relatively low performance. No other search system was available that ran in such restricted conditions, neither on the Sharp Zaurus nor on any other handheld device. Nevertheless, the vector space model is a well established information retrieval model that has been tested comprehensively with SMART [Salton, 1971]. Lucene represents a reasonable implementation of a vector space model enriched by a number of optimisations as shown below. In the following, the vector space model and Lucene's specific implementation of the model are described in more detail.

In the vector space model, relevance is determined by projecting content (d)[11]) and queries (q) into a multidimensional vector space and scoring the content (in our case, information about entertainment events) based on its distance to the query. This distance is determined with a scoring function expressed by a conventional vector product

$$score(q, d) = \sum_{k=1}^{n} w_{qk} * w_{dk} \tag{2.1}$$

[10]In the following, it will be referred to as Lucene.

[11]In the long tradition of information retrieval, it is common to refer to information objects as documents (d). In the focus of this study, it is preferred to call them more generally 'content'. However, the abbreviation for document (d) is not changed in the formulas of this thesis for reasons of consistency with information retrieval literature.

Each term has one weight for its occurrence in the query (w_{qk}) and one weight for its occurrence in the text of the content (w_{dk}). In the simplest case, this is binary where the term occurrence is represented by 1 and its absence by 0. The weight is then simply the number of terms that occur both in the query and the content. In practice, however, this does not offer enough variation in the score and limits the ability to distinguish relevant content from irrelevant. Good weighting functions are important as the retrieval effectiveness of an IR system depends highly on it [Buckley, 1993]. It is possible to create a weighting function from any set of parameters, however, statistical information is largely used at least for initial weights. In [Salton and Buckley, 1988], three basic kinds of weighting were proposed and tested with various content collections. This was further summarized and evaluated in [Chisholm and Kolda, 1999]:

1. A *local weight* L_{ij} represents the function of how many times the term t_i appears in the content d_j.

2. A *global weight* G_i is the function of how many times the term t_i appears in the entire content collection.

3. A *normalisation factor* N_j corrects the advantage of content with long texts over content with short texts.

Since the score is computed from both query and content terms, the scoring function contains these three components once for the query and once for the content. Many different strategies have been proposed and evaluated over the years – some of which are discussed and compared in [Buckley, 1993] and [Chisholm and Kolda, 1999]. The Lucene search library that was incorporated in this study expresses the *local weight* as the square root of the term frequency for both the query terms (q) and the terms of the content (d).

$$Lij = \begin{cases} \sqrt{|t\epsilon q|} \\ \sqrt{|t\epsilon d|} \end{cases} \tag{2.2}$$

The square root has a normalisation effect on high term frequencies that provides more weight to infrequent terms than frequent terms. An alternative to the square root is the logarithm as this was done in the SMART system. In [Chisholm and Kolda, 1999] however, square root transformation was found to produce better results than traditional log transformation. The *global weight* in Lucene is a standard inverse document frequency (IDF) as originally introduced in [Spärck Jones, 1972].

$$Gi = log(\frac{|d|}{|df + 1|}) + 1 \tag{2.3}$$

The IDF relates the total number of content items ($|d|$) to the total number of documents that contain term t (df) and expresses in this way the specificity of the term; as a result, terms that only appear in a few content objects are scored more highly than terms that occur regularly across the entire content collection.

For the normalisation factor, Lucene applies two different strategies:

$$Ni = \begin{cases} \frac{1}{\sqrt{\sum_{t\epsilon q}(\sqrt{|t\epsilon q|} * (log(\frac{|d|}{df+1}) + 1))^2}} \\ \frac{1}{\sqrt{|t\epsilon d|}} \end{cases} \tag{2.4}$$

For query terms ($t\epsilon q$), a standard *cosine normalisation* is applied as originally proposed in [Salton and Buckley, 1988]; for content terms ($t\epsilon d$), a different strategy is introduced that normalises over the number of terms in the content. This method is referred to as the *approximated normalisation factor* and has been comprehensively evaluated in [Lee et al., 1997]. Approximated normalisation basically offers the same advantages as standard cosine normalisation while being computationally much more efficient.

Besides these standard weighting elements, Lucene also offers *coordination level matching*

$$\frac{|t\epsilon d \wedge t\epsilon q|}{|t\epsilon q|} \tag{2.5}$$

that additionally boosts terms based on their frequency of co-occurrence in query and content ($|t\epsilon d \wedge t\epsilon q|$) normalised by their frequency in the query ($|t\epsilon q|$). In combination with the TFxIDF it provides an extra level of content distinction which is regarded as an advantageous strategy [Spärck Jones, 1972]. Under consideration of the previous, the Lucene search engine API presents one standard vector space model implementation with a small number of modifications for improved efficiency.

Content-based techniques have a number of disadvantages as described in [Griffith and O'Riordan, 2000]:

1. *Content Focus:* Content-based techniques focus on the content information the user is consuming, but omit the knowledge other people might have about it. Furthermore, these techniques are generally not able to look beyond the machine readable representation of the content. This means that features such as the representation style, the popularity of the information or other subjective features are not part of the content and can therefore not be included for the personalised service.

2. *Content Bias:* Besides the focus on the content representation, there is also a bias against content that the user has already seen. Since models are trained based on the user's past content, the personalised service will provide similar content in the future. Although this can be helpful for users in narrowing down potentially very large amounts of information, a model can become overfitted so that it introduces a bias that prevents people from retrieving different kinds of information for potential use.

Collaborative filtering

Collaborative filtering is based on the assumption that a group of like-minded people do also have a common taste for certain things. These things can be entirely virtual (e.g. certain kinds of information such as interesting websites) or non-virtual (e.g. books, movies or other goods). Systems that employ collaborative filtering therefore record and represent information from an entire user population in one large model. For this reason, it is also called social (information) filtering, a method that "automates the process of 'word-of-mouth' recommendations" [Shardanand and Maes, 1995, p. 2][12].

A system that provides personalised services though collaborative filtering records users' ratings for items that they have seen, used, bought or otherwise experienced. This information is used to populate the collaborative model that is represented as a matrix

[12] Although no collaborative techniques were used in this thesis, this section presents this overview as an interesting alternative with a potential use for future work.

containing a list of users on one axis and a list of rateable items (e.g. book titles or websites) on the other. The model is populated by user ratings for items thus gradually fill the matrix. Later, the model is processed by an algorithm that calculates the correlation between all users in order to find all those that are like-minded and therefore close in terms of interest. There are a number of different algorithms which can be used – the most common are Pearson correlation, Spearman rank correlation, mean-square difference or vector similarity as described and overviewed in [Griffith and O'Riordan, 2000]. This similarity measure can then be used to select the closest neighbourhood for a user. The output of collaborative systems are recommendations. The recommendation for a particular user is produced by selecting the closest neighbours for that user and recommending rated items from this 'neighbourhood' of like-minded people. More details about common, as well as more advanced techniques for collaborative filtering are provided in [Griffith and O'Riordan, 2000] and [Griffith and O'Riordan, 2002].

Since collaborative filtering focuses on people, the approach provides an alternative to content based techniques. The disconnection from the actual content representation allows this technique to be applied for content that is difficult to analyse, such as image or video content. It is also possible to address subjective content that cannot normally be analysed, such as the presentation style of a website. Furthermore, recommendations can be made for real life objects or entities of our thought (e.g. ideas and beliefs) which do either not have any processable representation or only in a very limited form [Lueg, 1997]. Besides this representational aspect, the collaborative approach also has the benefit to open up new avenues as this method is able to provide recommendations of entirely different items the user has not considered; however, this mostly depends on the user population in the community, their ratings and the diversity of available items to rate.

Collaborative filtering also has a number of significant drawbacks:

- *Sparsity:* Users do usually not rate many items. Nevertheless, the system is expected to provide quality recommendations. Few ratings lead to sparsely populated rating matrices. Based on the statistical nature of the approach, this might lead to the point where no recommendations can be provided or only few of poor quality. In [Shardanand and Maes, 1995, p. 70], it is reported that the system only started working properly after reaching the "critical mass" of 250 users. This is commonly referred to as the 'sparsity problem' and a serious scalability issue; for this reason, it is a big challenge for collaborative filtering applications to produce high quality recommendations despite few existing ratings. This is perhaps the reason why collaborative systems work well for common items (like books [Mooney and Roy, 2000] or music [Shardanand and Maes, 1995]) and commonly used information (like newsnet news [Konstan et al., 1997]) as this provides the best premise for frequent ratings.

- *Performance:* On the other side, a fully functioning collaborative filtering system with many items (e.g. books) and an equally large number of users (e.g. book enthusiasts) will need to compute similarity measures for every user with respect to every other user considering all ratings of these users. This leads to high demands on performance as the model needs to refresh its representation as the basis for up-to-date recommendations.
- *Bias:* Since collaborative systems are based on social opinion in a community, there is a tendency of such systems to recommend popular items as they attract more ratings; this creates a bias for such items as they achieve even more ratings after being recommended more often. Such a bias can originate from seasonal effects (e.g. websites about flower shops on Valentine's Day) or dramatic events.

2.5 Personalisation Output

This chapter has so far reviewed different kinds of user and context models and various methods and techniques for information acquisition to populate such models and perform personalised services. In this section, different kinds of personalised output are reviewed as some of these are applied in chapter 6 with the mobile information system that is used to evaluate the personalisation model.

A personalisation output is understood as a piece of information that has been transformed from a general 'one-size-fits-all' representation into an individualised representation. This section reviews different adaptation types; all common to adaptive hypermedia but with a special focus on mobile applications and personalisation. Based on this overview, two particular kinds of personalised adaptations are reviewed in more detail – personalised information retrieval and map personalisation – as they have been applied for the mobile information system in this thesis and more generally appear promising for mobile computing.

2.5.1 Types of Personalised Output

This section describes the various types of output that a personalised information system might produce for its users. This relates specifically to the 'Personalised Effect' as shown in figure 2.2 on page 15. Personalised output is differentiated into the two categories of content information and structure (e.g. alternative webpages with different details and alternative webpage linking) as well as content presentation (e.g. alternative media such as text descriptions versus video tutorials) based on a number of prominent reviews in adaptive hypermedia [Brusilovsky, 1996, Brusilovsky, 2001, Kobsa et al., 2001].

Personalising Content Information and Structure

This refers to personalised adaptations either to modify the core content or the way it is linked.

- *Content variation:* A system can maintain alternative versions of the same content and present it based on, for example, who is accessing or in what situation it is accessed. Variations may be provided coarse-grained (e.g. entire webpages) or more fine-grained (i.e. text fragments/paragraphs). Alternative versions of content are used for personalised web portals like MyYahoo! [Marmasse and Schmandt, 2000], where content is selected from categories based on user preferences. The Personalized Information Description Language (PIDL)[Koike et al., 1999] is an XML-based standardisation effort that defines the management and application of alternative content variations across different personalisation media and methods.

- *Hyperlink sorting/ranking* transforms lists of links into a ranked form. The rank position indicates the degree with which a hyperlink is recommended by the system and serves as a personalised recommendation. Letizia [Lieberman, 1995] provided link recommendations in a separate browser window to assist the user during a web session. Hyperlink ranking is provided by every common web search engine that generates a list of ranked results based on their relevance to a user query. Chapter 5 presents a personalisation model that implements one way of the personalised ranking of search results. More examples of personalised information retrieval systems are provided in section 2.5.2 below.

- *Hyperlink annotation* enhances an existing content hyperlink structure with additional information. Syskill&Webert [Pazzani et al., 1998], WebWatcher [Joachims et al., 1997] and Personal WebWatcher [Mladenic, 1996] applied this technique to recommend potentially relevant websites to the user.

- *Optional linking* personalises the information space to the user by adding or removing parts of the linking structure thus enabling or disabling access to certain content. HIPS [Oppermann and Specht, 2000], a mobile museum guide, for example included links to paintings based on proximity and interest.

Personalising Content Presentation

This refers to all personalised adaptations that alter the presentation of content whereas content information and its structuring stays.

- *Text presentation and colouring:* The text representation can be personalised by changing the font type, further emphasising the font (e.g. bold, italics) and its size. Furthermore, the text representation my be coloured in parts – a technique also called 'fragment colouring' [Kobsa et al., 2001]. This allows a system to shift users' attention to relevant parts of the content by emphasising one part and withdrawing importance from other parts. For that, colours may linked to meaning based on the cultural background of the user. This still allows all users to access all content while conveying different levels of relevance to different users.

- *Modality:* A change in modality allows users to access information in different media types based on their preferences and abilities. The AVANTI system [Fink et al., 1998] presented content either as a map or text based on users' physical abilities. This, however, may not only be determined by the user but also by the user's current situation. For example, a route planner application may provide a map to a walking user but an audio description to a driver. This is especially relevant in mobile scenarios where such situational changes occur more spontaneously.

This thesis is especially focused on two forms of personalisation - personalised information retrieval and map personalisation. Both types of personalised output have been applied as part of the mobile information system in chapter 6 which is more comprehensively described in appendix H. Both forms of personalisation are early in their development but have a promising future. The next two subsections describe these two types in more detail.

2.5.2 Personalised Information Retrieval

As shown in section 2.4 on page 35, information retrieval and filtering is an important technique that offers tools for people to manage large amounts of information and helps them to find relevant content. The success of search engines in recent years is an example of the necessity of such tools. Information retrieval has reached a position of widespread public attention and becomes an important part of everybody's daily information life.

Despite the popularity of search engines, only a few provide personalisation features. A recent study of 60 search engines [Khopkar et al., 2003] revealed that most personalisation features were minimalistic and generally difficult to access and use. User models were hardly employed and there was no mention of the use of any model that employs context. The personalised portal of Yahoo was described as having the most effective and integrated personalised solution, but still lacking anything in depth that goes beyond basic customisation. The following list describes a selective number of relevant research systems that provide different forms of personalised information retrieval.

- The *OBIWAN system* [Pretschner and Gauch, 1999] was a personalised web interface for web information retrieval that was developed at the University of Kansas. It employed ontology-based user profiles structured as concept hierarchies. Results from an internet search engine were personalised by re-ranking and filtering and moderate empirical results were achieved.

- In *PowerScout* [Lieberman et al., 2001], a software agent provided recommendations to websites of potential interest. The agent retrieved and personalised search results from a conventional search engine (i.e. Altavista). PowerScout constantly constructed and submitted queries to this search engine based on the user's browsing history. Search results were categorised into concepts and presented to the user.

- The *Outride system* [Pitkow et al., 2002] was a personalised web search component. It originated at the Xerox Palo Alto Research Center (PARC) which later led to the company Outride Inc. which was acquired by Google in 2001. The system operated as a browser sidebar delivering personalised output by modifying user queries based on a profile and then filtering and re-ranking returning search results. The system demonstrated that personalised information retrieval can help users (experts as well as novices) to reduce their search effort on the web by more than 50%.

- *PResTo!* [Keenoy and Levene, 2005] was similar to Outride and used as a web browser plug-in acting as a client-sided application between user and search engine. The system personalised search results by re-ranking them based on a user model that modelled information of the user's previous interactions with the search engine.

- Billsus and Pazzani's PDA edition of the *Daily Learner* [Billsus and Pazzani, 2000] provided news from 9 different categories (e.g. politics and entertainment) to its users. Besides the recommendation of news stories, the mobile application also offered a search feature. Search results were personalised by re-ranking depending on how users accessed news stories in the past.

- A mobile information system developed for the *WebPark* EU-IST project [Mountain and MacFarlane, 2007] offered information search and map-based functionalities. The application was a solution for visitors of an outdoor recreational park area. The search provided access to geographically indexed documents that were ranked depending on a number of geographic criteria.

Personalised information retrieval systems usually adapt at two different stages in the search process – at query time and/or at result time.

At query time, some personalised information retrieval systems employ *query* *modification*, an information retrieval technique used for more than 30 years [Spink and Losee, 1996, Kelly and Teevan, 2003]. This is either done by the addition of extra query terms to an existing user query (query extension) or the modification of the weights of existing query terms (query re-weighting). The WebMate browsing and searching agent [Chen and Sycara, 1998] assisted users' search by expanding search queries with learned keywords based on a model of correlated word pairs. The Outride system did that in similar fashion but additionally used implicit feedback; as users also browsed for content using an ontology, Outride used its category information to augment subsequent queries.

At result time, most personalised information retrieval systems adapt the result list that is returned from the search engine. This may be done in two different ways:

- *Re-ranking* is the reordering of an existing list of search results depending on the information provided in a user and/or context model. The OBIWAN system [Pretschner and Gauch, 1999] used a publically available search engine and modified its generic search result rank into a personalised search rank based on the user's interest in a number of topical categories. The strength of these categories was constantly adjusted along page content and the user's viewing time. Re-ranking in the Outride system [Pitkow et al., 2002] was performed more simplistically based on the correlation between the content (titles and metadata) and the user profiles using the vector space model. In PResTo! [Keenoy and Levene, 2005], re-ranking was also performed along a vector space model taking into account the URL structure as well as temporal information. Temporal information included information about the hit frequency (amount of times this URL has been found before), lookup frequency (amount of times this URL has been accessed) and the age of the URL in the user model.

- *Filtering* here refers to the removal of non-relevant content from the result list. In [Pitkow et al., 2002], it suggests removing search results that the user has seen before, to support focusing on new results. This argument is strongly challenged in [Bradley and Smyth, 2002] in support for re-ranking as being generally advantageous over filtering in situations that require a system to deliver a good coverage of available information. Filtering is indeed much more challenging as it actively suppresses content that might have been useful for the user; however, this is justified in situations of constant and comprehensive information streams that does not allow users to fully examine their content.

The mobile information system that is applied in chapter 6 for the mobile user study makes use of personalised information retrieval at result time. It uses the personalisation model described in chapter 5 to re-rank search results based on the user's query and the user's current contextual situation. Additionally, the system provides a map visualisation for personalised search results thus an example of producing a personalised map. Other projects, such as WebPark, also demonstrate that personalised information retrieval strongly relates with geographic concepts as soon as applications become mobile. The integration of information retrieval with map content appears promising and valuable for the purpose of personalised information provision in mobile settings, therefore, the next section focuses further on another form of personalised output – the personalisation of maps and geographic information.

2.5.3 Map personalisation

The creation and use of maps has an equally long tradition as the written word and the institution of the library. In the past as much as now, accurate maps are an important form of knowledge. Half a century ago, maps started becoming digital giving birth to Geographic Information Systems (GIS) and, more recently, Geographic Information Science as its emerging discipline that can be related to information science [Goodchild, 1992]. Such systems are centred around the generation, management and extraction of geographic information. Such systems ran exclusively on powerful computers, were controlled by experts and were mainly used for special applications (e.g. environmental monitoring). Now, the electronic map has started to move into the private sector. The rapid growth of the world wide web boosted this new development and allowed for widespread access to maps. According to Wikipedia [13], MapQuest[14] was the first popular routing service introduced in 1996 followed by MultiMap[15] in the same

[13] `http://www.wikipedia.com`, accessed April 14, 2008
[14] `http://www.mapquest.com`, accessed April 14, 2008
[15] `http://www.multimap.com`, accessed April 14, 2008

year that became one of the most popular UK routing websites. In 2000, the web was already the major medium for the dissemination of map content [van Elzakker, 2000]. The wave of wide-spread dissemination of geospatial information has continued rapidly in recent years with websites such as Google Maps [16] or products such as Google Earth [17] or NASA World Wind [18].

At the same time, the rising trend for Personal Digital Assistants (PDA's) also boosted the need for mobile map applications and services. As a consequence, Location-Based Services (LBS) emerged as a new class of applications able to incorporate the current location of a user for its service provision. Some of them employ complex user and context modelling features although context is mainly restricted to spatial attributes such as location and proximity. The following list presents a number of selective mobile guides that provide (personalised) map adaptations to mobile users:

- *Cyberguide* [Abowd et al., 1997] was an early example of an LBS that worked either indoors based on infrared beacons or outdoors based on GPS. The application offered tourist guide functionality that provided visitors of a research centre with a lab map and information about ongoing projects. The map application employed automatic scrolling and displayed the user's current position.

- The location-based system *comMotion* [Marmasse and Schmandt, 2000] associated personal information (i.e. todo lists) with users' current location based on GPS technology. The system provided maps with position information as well as other web information. It personalised the map by learning new locations over time and asking the user to annotate these locations with meaningful labels; this gradually transformed the map into the user's personal map.

- The *GUIDE* project [Cheverst et al., 2000] provided city visitors of Lancaster with up-to-date and context-aware hypermedia information using both a user model and an environment model. Users' location was approximated though the use of wireless access points and maps were employed and integrated on a web interface.

[16] `http://maps.google.co.uk`, accessed April 14, 2008
[17] `http://earth.google.com`, accessed April 14, 2008
[18] `http://worldwind.arc.nasa.gov`, accessed April 14, 2008

- The EU-IST project *CRUMPET* [Zipf, 2002] was a mobile guide for tourists in outdoor scenarios. The system offered interactive maps that highlighted users' current position and sights. Based on users' interest and location, the system recommended tourists attractions and provides personalised tours on the map display.

- In *[Rantakokko and Plomp, 2003]*, mobile maps were adapted to the user's location, the orientation of the device and the user's physical interaction with the map. A variety of sensors provided the software with necessary information although no additional effort was made to personalise the map beyond this basic adaptation.

- The EU-IST project *AmbieSense* [Myrhaug and Göker, 2003] developed a mobile guide for tourists in indoor and outdoor scenarios. The system allowed users to contextually search and to browse and also provided georeferenced content proactively based on users' proximity to hardware mounted in the environment.

- The *Taeneb City Guide* project [Dunlop et al., 2004] was a more recent mobile guide for tourists. The system had a map interface and provided information about tourist attractions and other places of interest that was personalised by dynamic content filters (e.g. for restaurant selection based on food type and price range).

Most mobile guides, including the information system used in this thesis, use map layering – a concept that is beneficial for the adaptation of map content, structure and presentation. A map layer is a data set that describes a single aspect of geographic data. This can be a set of lines that represent the streets of an area, or a set of data points that describe users or event locations in a region. The GIS application stacks these layers to a single map as shown in figure 2.5. Even though the user of the application might view the map only as one single unit, it consists of different layers each contributing its own unique kind of information. Map layering is an important GIS concept with two main benefits:

- Data can be reused for different purposes. While some layers may be static (e.g. city streets) others may be created dynamically and periodically over time (e.g. user's current position). Whereas static layers may remain as a fundamental part of the map, dynamic layers may be refreshed regularly allowing the GIS to target its use of resources; a relevant issue for mobile devices.

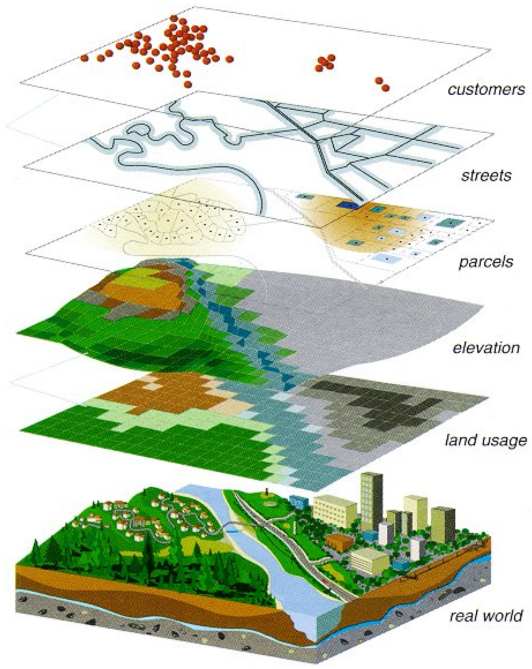


Figure 2.5: *Composition of a map by layers. Illustration obtained and adapted from http://ssnds.uwo.ca/sscnetworkupdate/2006winter/gissupport.html, accessed April 14, 2008.*

- Collections of different geographic data sets about one map area can be managed and used in parallel. As data is separated on different layers, a personalised system may choose data provided by layers only when appropriate instead of plotting all data at once thus overpopulating the map with details.

Map Content Adaptation

In this thesis, map content are all features that a map may contain. This refers to the map layers as structural entities as well as everything that is contained within a layer. Map content is distinguished from its presentation which is described in the next section. The following types of map content adaptations are identified:

- *Layer Variation:* One of the most coarse grained forms of map content adaptation is through the selection of pre-existing layers. Since layers contain certain kinds of information about one geographic region (e.g. hotels, restaurants, shops or sightseeing locations), it is possible to generate an adaptive map simply by selecting the most relevant layers. The map is then a composition of layers and a personalised map application would compose a map through the variation of map layers for different users based on information about the user and the user's contextual situation. For example, a tourist who just arrived in a new city would receive a map with a layer containing hotel locations. As soon as the user has reached the hotel, the map would exchange this layer with other layers providing content about sightseeing, restaurants, shops etc.

- *Content Selection:* On a more fine grained level, content selection determines which geographic features enter a map. Content may be selected through the addition of new features onto an existing map or layer as well as through the removal of features. Content selection may be performed using information retrieval using an algorithm that queries an information repository adding a relevant ranked list of results to the map. This approach was used for the mobile information system applied in chapter 6. Alternatively, an information filtering algorithm may constantly process an incoming stream of information based on a user or context model and adding all relevant features to the map or removing all non-relevant features. The Taeneb City Guide [Dunlop et al., 2004] used query filters to control the amount of content that enters the map. Users were able to restrict features through the explicit statement of interest through the user interface. A personalised map-based application might also use implicit methods to gather and employ user knowledge or directly use and exploit contextual information in order to add or remove map content.

- *Encoding and Quality of Service:* Geographic features can be stored, transmitted and presented in different encodings; either as vector graphics or as raster graphics [Harmon and Anderson, 2003, p. 73-78]. Vector graphics are high quality, usually very precise, but may consume large amounts of memory if many features are present. Raster graphics (i.e. images) offer only a predefined amount of detail and quality but are more efficient with memory and bandwidth. The quality of a map can be adapted by encoding information in various formats and quality levels using these encodings, thus adjusting to the personal requirements of quality of service. This is important when maps are stored remotely and loaded on demand. Depending on the complexity of the map layer, it can be advantageous either to use an image or a vector layer in order to provide a minimum service quality. These service classes can be determined by the user's situation. For example, a user who demands routing information between two locations in a town will be more interested in the precise path and information on directions rather than a high quality image layer. On the other hand, the same user randomly walking and exploring the city will be much more interested in exactly those high quality images that show necessary details about streets and other highlights.

Map Presentation Adaptation

Map presentation can be adapted in the following ways:

- *Modality:* Personalised maps may provide different forms of visualisation based on users' preferences, abilities and their current situation. The AVANTI system [Fink et al., 1998] provided content personalisation based on users' physical abilities by switching between map and text representation. The BMW Personal Navigator (BPN) [Krüger et al., 2004] is an example of a multi modal map navigation system that spans over different platforms (PC and PDA) and usage (travel planning, driving navigation, pedestrian navigation) supported by a multi modal map interface.

- *Scaling:* Whereas traditional paper maps have predefined scales, electronic maps allow for its adaptation. This can be based on various contextual conditions (e.g. user travelling speed) or based on the user (e.g. user preference or behaviour). For example, the mobile guide presented in [Rantakokko and Plomp, 2003] re-scales the map visualisation via zoom that is based on the closeness of the users face to the screen.

- *Orientation:* Traditionally, a map assumes north to be at the top for a standardised presentation. This can be irritating when moving with changing directions. Personalised electronic maps are able to adapt to this by re-orienting the map to the user's direction of movement. The mobile guide presented in [Rantakokko and Plomp, 2003] provided map orientation based on an internal electronic compass.

- *Detail level:* Even though the map carries a rich set of information, perhaps divided over many layers, it is not necessary to present every detail at all times. Different people prefer different levels of detail. An expert might only need the most essential map features whereas a new user might want the full level of detail. Zipf's 'Focus Maps' [Zipf, 2002] graphically highlight relevant parts of a map with more details while removing information from other, less relevant parts of the map. On the other hand, contextual information can help selecting the right level of detail based on the user's activity. For example, a fast moving car would not need a high quality visualisation but instead prompt and accurate routing descriptions.

- *Symbol and text presentation:* Points of interest are usually presented by symbols which can be adapted for personalised map representation. This includes the kind of symbol, its styling, the colouring (e.g. considering users' cultural background), the size and opacity (i.e. based on actuality and relevance of the geographic content). The mobile information system of chapter 6 used coloured dots for the visualisation of events in the map and adapted these colours based on relevance. The map service provided in CRUMPET [Zipf, 2002] adapted the map visualisation based on the cultural characteristics of users (e.g. the cultural association of colours based on the user's country of origin) along a number of different graphical properties (e.g. the colouring of map features). Since maps might contain textual information (e.g. in the form of labels), the presentation of this text can also be personalised by front type, size, style and colour based on user and situation.

2.6 Summary

In this chapter, a wide range of literature on personalisation has been reviewed along user and context modelling – the overall focus of this thesis. The review has been structured along the process of personalisation covering the user model and the context model as conceptional entities, the acquisition of user and context information, the review of relevant techniques and the production of personalised output. Particular emphasis has been tributed to personalised information retrieval and map personalisation, two very young yet very promising areas of personalisation. In the review, numerous personalised information systems have been described. Links have been provided from the literature to relevant latter sections of this thesis.

3

Information Needs and Behaviours of Mobile Users

Winter comes, you wish it were summer. Summer comes, you live in dread of winter. That's why we never tire of travel.

La Leggenda del Pianista sull'Oceano
(Legend of 1900)
GIUSEPPE TORNATORE

3.1 Introduction

3.1.1 Overview

An effective personalised information system should be adapted to its application domain. In chapter 1 we suggested adapting a personalised information service based on the way it is used and based on the content it personalises. This thesis targets semi-mobile and mobile usage (applying a system while being away from the usual work or home environment or while being on the move) and focuses on personalising entertainment and map content. In order to do that appropriately, it is essential to know more about mobile users.

This key requirement was shared with the AmbieSense EU-IST project that focused on travellers and tourists as one type of mobile users and their information needs.

AmbieSense functioned as a source of support [1] for this thesis in terms of funding, equipment and other resources. Furthermore, AmbieSense studies influenced the studies conducted in this thesis including some of its methodology. The studies presented in the later chapters of this thesis may be seen as a more specialised extension of the work produced by AmbieSense and the various AmbieSense studies. For this reason, the present chapter focuses on a set of relevant AmbieSense results. These were gathered in a number of AmbieSense user studies in 2004 that influenced and shaped this research to some extent. These user studies were planned and conducted during the time of the AmbieSense project in a collaborative effort by all members of the international project team including myself. Therefore, the data presented in this chapter are not a direct result of this thesis. Nevertheless, the data provide a basis for the arguments of this work. This chapter differs from the more general review of related work presented in the previous chapter. It includes direct personal involvement and therefore a much stronger connection between the research conducted in AmbieSense and the research conducted for this thesis. More details about this connection are described in the following subsection.

The AmbieSense data are presented in terms of three questions that are particularly relevant to this thesis. Firstly, it is important to know what kind of digital content mobile users prefer and expect. Secondly, it is essential to gather knowledge about mobile information behaviour. Thirdly, it is important to know about the types of personal information people are prepared to provide in order to receive personalised services. Section 3.3 will discuss these points in more detail. The next subsection provides more information about the AmbieSense project and its connection with this research.

3.1.2 The AmbieSense EU-IST Project

The AmbieSense EU-IST project[2] was centred around the problem of delivering situationally relevant, digital content to travellers and tourists, one type of mobile users, who use handheld devices to access information services in their vicinity. Special

[1] More about the precise relationship between this thesis and the AmbieSense EU-IST project is described in the next subsection. From this point on, the AmbieSense EU-IST project will be called 'AmbieSense'.

[2] Contract number IST 2001-34244. Project website available at `http://www.ambiesense.net/`, accessed April 14, 2008

embedded hardware was located in key locations to distribute content to information services on personal digital assistants and mobile phones, thus forming an infrastructure for personalised and context-aware computing. Travellers and tourists from Seville (Spain) and from Oslo Airport (Norway) were recruited as participants for mobile user studies and questionnaires some of which are presented in section 3.3.

AmbieSense combined research and development efforts from a range of academic institutions and companies[3]. AmbieSense also funded several PhD research projects including my own. In particular, it provided the following support for the research reported in this thesis:

- *Funding* based on active participation in the project; working on questionnaire design, conducting a wide range of user studies, playing an active role in system development and assisting in project management throughout the entire project.
- *Opportunity to shape research* conducted in AmbieSense based on own research interests. For example, the questionnaires for two of the user studies (Seville June 2004 and Seville September 2004) were extended by one of my own questions asking users what types of information they are willing to provide for personalisation (see section 3.3.3 below); other questions were shaped collaboratively through participation in questionnaire design as part of the team.
- *Access* to project technology, infrastructure and expertise. For example, the mobile device used for the experiment reported in chapter 6 was provided by AmbieSense and *practical experience* for conducting mobile studies was gained while planning and conducting mobile studies during AmbieSense.

Based on that, there is a certain degree of connection between AmbieSense and this thesis which is typical for every larger research project that also hosts individual research efforts.

The data presented in this chapter was obtained by AmbieSense in a collaborative effort by all project partners. The results reported in section 3.3 are therefore not a

[3]SINTEF ICT (Norway) as the project coordinator, The Robert Gordon University (UK), YellowMap AG (Germany), Oslo Airport (Norway), Lonely Planet (UK), CognIT AS (Norway), The Norwegian University of Science and Technology (Norway), Sevilla Global SA (Spain), and Siemens (Austria).

direct output of this thesis. Nevertheless they are presented here to provide a basis for the arguments of this thesis as a continuation of the work that was conducted in AmbieSense. This chapter considers selected AmbieSense results that relate to three questions relevant for this thesis – the content preference and expectations of mobile users, mobile information behaviour, and users' willingness to provide personal information in exchange for personalised services.

The next section provides more details about the travel and tourist domain – the user domain that was targeted by AmbieSense – and further highlights why this domain serves as a good example for the application of mobile computing. In section 3.3, selected results from AmbieSense are presented. These results have been obtained from two user studies in Seville (Spain) and a large-scale market survey. The implications of these results with respect to this thesis are then discussed in section 3.4.

3.2 Travel Domain as an Example for Mobile Computing

International tourism has increased over the last two decades. According to the UK Office for National Statistics[4], the number of visits made to the UK by overseas' residents doubled between 1985 and 2005 reaching a total of 30 million visitors in 2005. A record sum of £14,2 billion was spend in 2005, 8% more than in 1985 and 6 % more than in 2004. On the other hand, UK residents' visits abroad has more than tripled since 1985, reaching a total record of 66,4 million in 2005. Two thirds of these were for holidays – most of them to Europe (80%) with Spain as the top destination (about 20%). The total spending of UK residents abroad quadrupled between 1985 and 2005 to a record sum of about £32 billion in 2005. The rising popularity of travel and tourism has continued in the meantime. Between January and April 2007, international tourist arrivals worldwide increased by over 6% in comparison with the same months in 2006 based on the June 2007 edition of the World Tourism Barometer, a triannual report that is published by the World Tourism Organization [WTO, 2007].

This demonstrates the high and increasing significance of the travel and tourism

[4]Available from `http://www.statistics.gov.uk`, accessed April 14, 2008

domain that was chosen as the domain in AmbieSense. As mentioned earlier, the main aim of AmbieSense was to provide relevant digital content to travellers and tourists in the right situation. Travellers and tourists are naturally on the move and therefore represent an appropriate case for a mobile user population. Travellers and tourists potentially benefit from mobile information services that provide digital travel information to their personal devices. The project provided an opportunity to investigate and better understand travellers and tourists. This is valuable to this thesis since travel in itself has a rich set of situations that include the inquiry of content such as the ones covered in these studies.

The next section presents and discusses the mentioned results that have been obtained from questionnaire data collected in two AmbieSense user studies that took place in Seville/Spain. Additionally, a large-scale AmbieSense market survey gathered data about information needs and information behaviour in Seville as well as Oslo/Norway. The survey was additionally conducted on the web where travellers and tourists accessed the questionnaire through the Oslo Airport and Lonely Planet websites.

3.3 Relevant Results from AmbieSense

This section presents relevant results from AmbieSense and some background about how this data has been collected.

Results have been obtained from two AmbieSense user studies conducted in Seville/Spain (13 participants in June, 76 participants in September)[5] and a large-scale market survey (438 participants) that gathered questionnaire data from travellers and tourists on site and web users that accessed the questionnaire online through the Oslo Airport [6] and Lonely Planet [7] websites.

[5] Both studies consisted of questionnaires and mobile search tasks. This chapter however only revisits some of the questionnaire data that is relevant to the aim of this thesis. Results from the mobile information search tasks have not been used in this chapter since they reflect more specifically the goal of AmbieSense. Data from the mobile search tasks are therefore not described here.

[6] `http://www.osl.no`, accessed April 14, 2008

[7] Lonely Planet is a well known publisher for travel guides (`http://www.lonelyplanet.com`, accessed April 14, 2008).

Questionnaires were selected for both Seville studies as a convenient and time-efficient instrument for collecting background information. It did not demand too much cognitive effort since participants also had to complete a series of mobile search tasks. Questionnaires were also used for the market survey as a time-efficient data collection instrument. This was crucial since people are only willing to spend a few minutes. They were also selected since the consortium decided to replicate the questionnaire online.

Seville was selected as a test location based on its inner city with its many attractions (e.g. restaurants, sites, and shops) that made it particularly valuable for conducting walkable mobile user studies with tourists. Oslo was chosen for its airport as a manageable example for a test site that allowed to collect data from people in travel situations (e.g. arrival, departure, waiting at the gate, and airport shopping).

For the market survey, tourists and travellers were directly approached in the streets of Seville or at Oslo airport with the questionnaire as shown in appendix B.2. Recruiting for the two mobile user studies was organised by our project partner Sevilla Global – an urban agency for the economic development of the city of Seville – actively supported by other project partners including myself. Recruiting of participants was performed from information stands located at various points in the inner city. All personnel at these information stands was instructed to target tourists and therefore sample participants for the relevant AmbieSense user group. Tourists were approached and directed to the lobby of a nearby hotel where the study was initiated. Questionnaires were completed in the hotel lobby before and after the mobile tasks. Appendix B presents all questionnaires that were used to collect the AmbieSense related data presented in this chapter. More details about the AmbieSense studies (including the mobile search tasks) can be found in the various technical reports [Myrhaug and Göker, 2004, Myrhaug et al., 2004b] that have been published in part [Göker et al., 2004, Göker and Myrhaug, 2007].

Figure 3.1 shows both the age and gender distribution for the mobile study conducted in June 2004. All 13 users were in the 19-39 age group with a larger proportion of male participants.

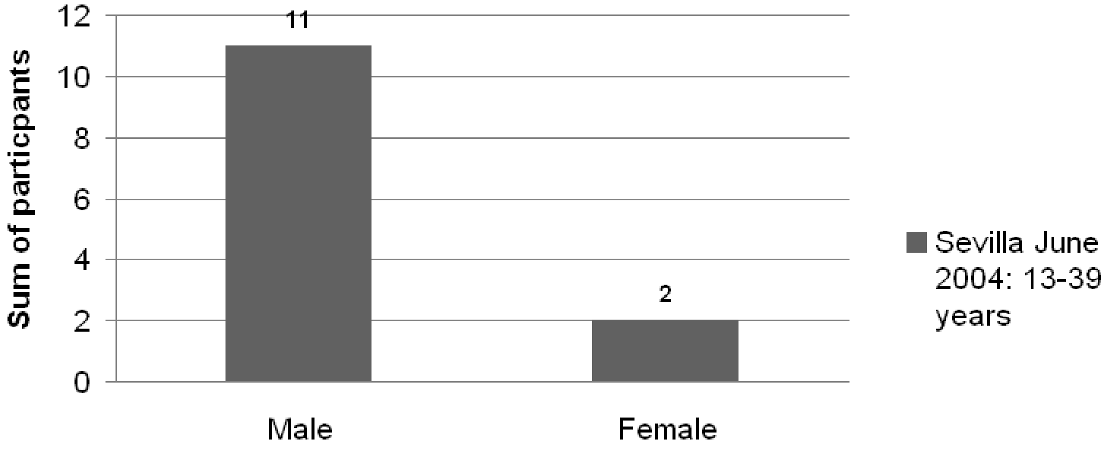


Figure 3.1: *Age and gender distribution of the 13 participants of the mobile study in Seville (June 2004)*

Due to time restrictions and operational constraints during the experiments, pre-questionnaires – and therefore age and gender data – from the mobile Seville September study were only partially collected for 24 participants out of 76 (see figure 3.2). This

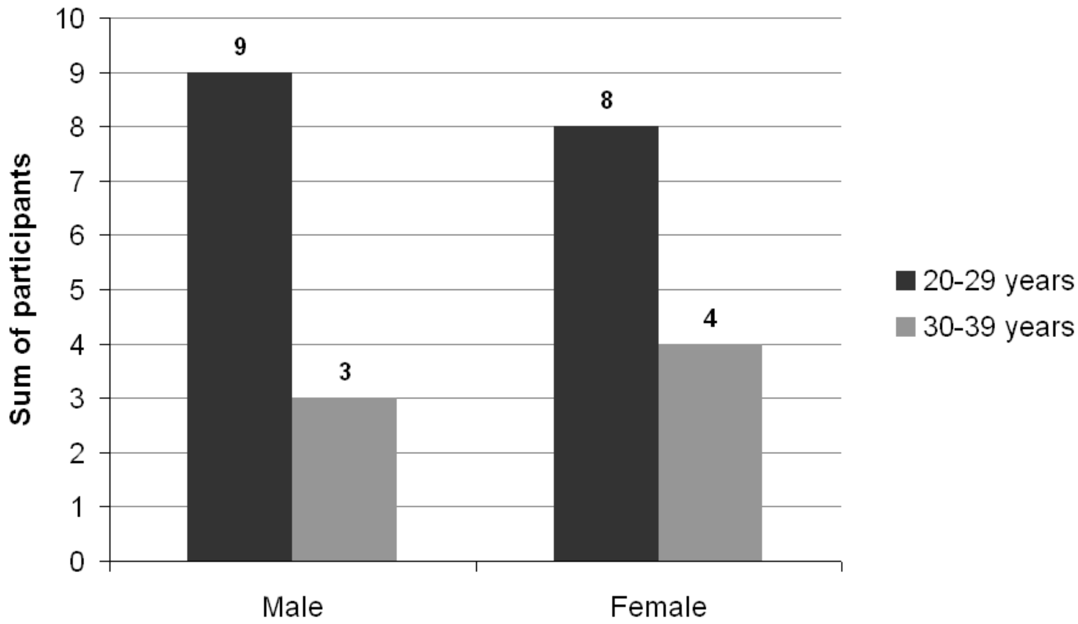


Figure 3.2: *Age and gender distribution for 24 (out of 76) participants of the mobile study in Seville (September 2004)*

partial distribution shows an equally balanced user population with most participants being between 20 and 29 years old.

The market survey collected opinions from a generally very wide range of age groups with most participants being between 20 and 29 years old (see figure 3.3). The age

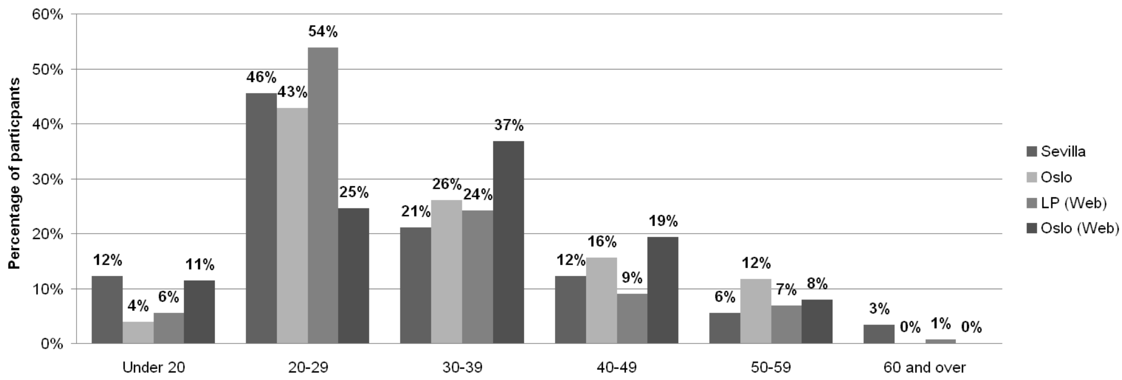


Figure 3.3: *Age distribution of the 438 participants of the AmbieSense market survey*

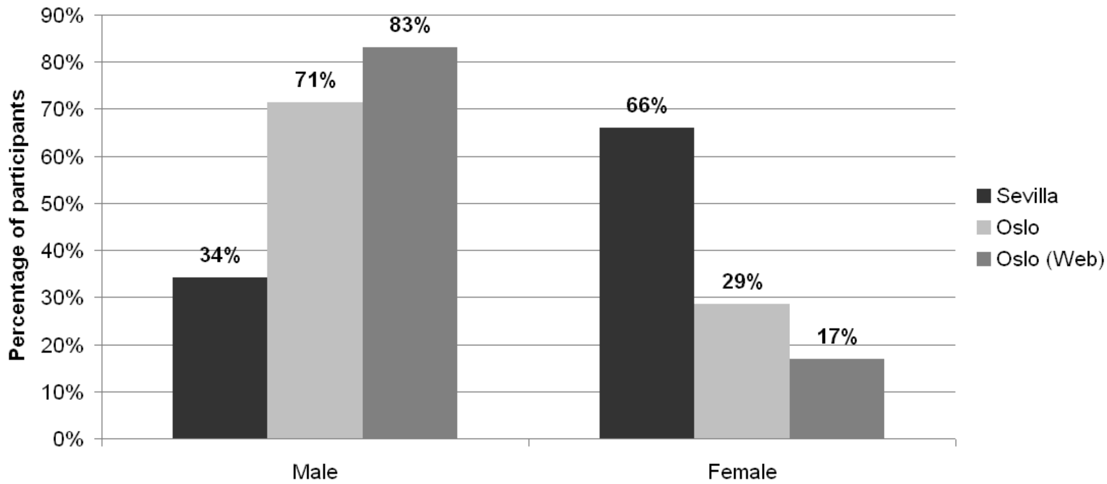


Figure 3.4: *Gender distribution of the 438 participants of the AmbieSense market survey*

distribution is skewed (Pearson's χ^2=23.36, p<.001), as shown in figure 3.4, which may indicate potential gender effects in the data. Note that gender information is not available for data collected by the Lonely Planet web survey[8]

[8]Lonely Planet did not require this information since they were already in possession of sufficient data about their targeted user population. Additionally, technical issues required the questionnaire to be limited

AmbieSense results in this section are presented selectively with respect to three questions that are of particular interest to this thesis. These results are obtained from the AmbieSense technical project reports [Myrhaug and Göker, 2004, Myrhaug et al., 2004b] and its underlying data that have been published in part [Göker et al., 2004, Göker and Myrhaug, 2007]. All AmbieSense questionnaires, with the questions from which these results have been originally obtained, are listed in appendix B.

Firstly, it is important to know what kind of digital content mobile users prefer and expect in order to meet the demands of the user. Every information system that delivers personalised content, mobile or stationary, requires information about the user's content preference in order to target its service. For this reason, both the research effort targeted by the AmbieSense project and the research described in this thesis aimed to find out more about the content demands and expectations of mobile users. Secondly, it is essential to gather knowledge about the information behaviour of users as this might provide insights into how information should be delivered. In AmbieSense, close attention was paid to the ways in which users access information before and during their trip and in which situations they like to access digital content. This is also a relevant question to explore for this thesis since every personalised information system has to target its information delivery to the user and the users situation. Thirdly, it is key to investigate the types of personal information that users are willing to provide as this shapes the possibilities of a personalised information system as described in the previous chapter (see figure 2.2 on page 15). This data was obtained by a question that the author especially contributed to the AmbieSense questionnaires in two occasions[9]. The question was originally inspired by a web survey conducted by the Personalization Consortium in 2000 as described later in section 3.3.3. The following three subsections discuss these three questions in more detail along with selected results the AmbieSense project and the mentioned web survey. An overall discussion follows in section 3.4.

in size for their website.

[9]The question was added to the Seville June 2004 questionnaire (see question C20 in appendix B.3) and, in a slightly refined version, to the Seville September 2004 questionnaire (see question C17 in appendix B.4).

3.3.1 Information Need and Relevant Content Types

One important question that was addressed from several perspectives was what information types tourists and travellers acquire and find useful. Figure 3.5 and figure 3.6 depict the relevance that was assigned for different types of information that are commonly acquired and consumed when travelling. Whereas figure 3.5 shows the relevance ratings in relation to a tourist city scenario, figure 3.6 represents ratings with respect to content relevant in an airport scenario. The tourist city scenario in figure 3.5 furthermore distinguishes between data collected at site (Seville and Oslo) and at the Oslo Airport and Lonely Planet websites (Oslo (web) and LP (web)). Figure 3.5 shows that sites and attractions are highly rated together with people and culture. Events and nightlife scores in the middle field whereas information about research and universities is considered less relevant. Generally, ratings were quite similar with little variation. Travellers surveyed

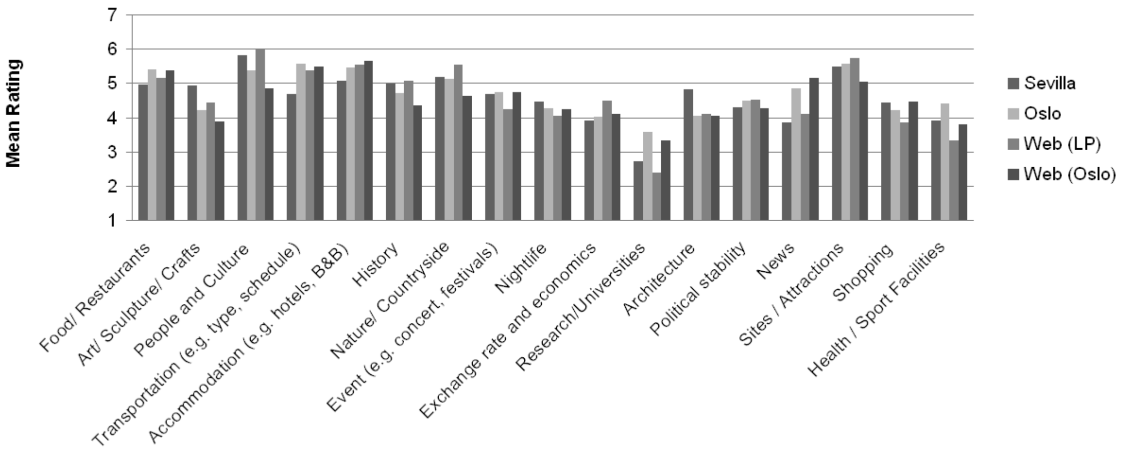


Figure 3.5: *Relevant kinds of information for travellers and tourists (AmbieSense market survey for tourist city scenario)*

in Oslo (see figure 3.6), when asked about the relevant kinds of information in an airport scenario, highly scored personal flight and general destination information, followed by transportation information and airport maps. Generally, there was a strong focus on personal travel plans while other types of information were rated lower (e.g. other flight information and advertisements). Users who participated in the user experiment in Seville in June 2004 also answered one question as part of the pre-questionnaire that inquired what kinds of content users expect from a mobile device (see figure 3.7). In comparison

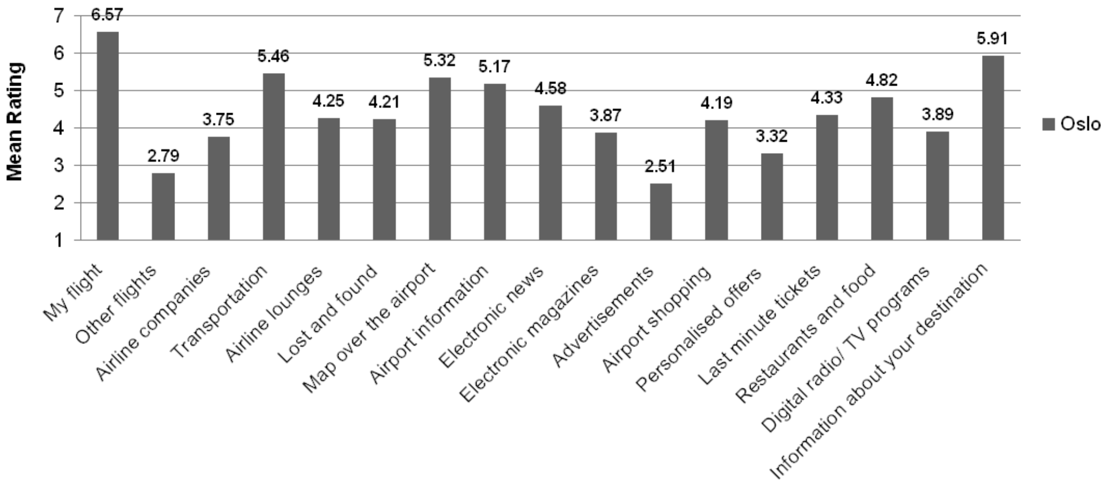


Figure 3.6: *Relevant kinds of information for travellers and tourists (AmbieSense market survey for airport scenario)*

to the previous question from the market survey, this was more focused on the content in relation to the device that delivers the content. The most relevant content types were:

- maps, events, transportation, food/restaurants (85%)
- sites/attractions (77%)
- art, history, nightlife, weather information (62%)
- shopping (46%)
- culture (38%)
- nature, political stability, research and health/sport ($\leqq$ 8%)

The market survey also inquired from travellers and tourists which types of information services would be most useful and fun to use (see figure 3.8). For the tourist city scenario, participants rated consistently that map services would represent the most useful and fun information service followed by an internet service. Little interest was expressed for digital books, digital magazines, financial news and children infotainment. The same question asked in relation to an airport scenario (see figure 3.9) revealed very similar results. Large interest was expressed for flight information followed by internet access.

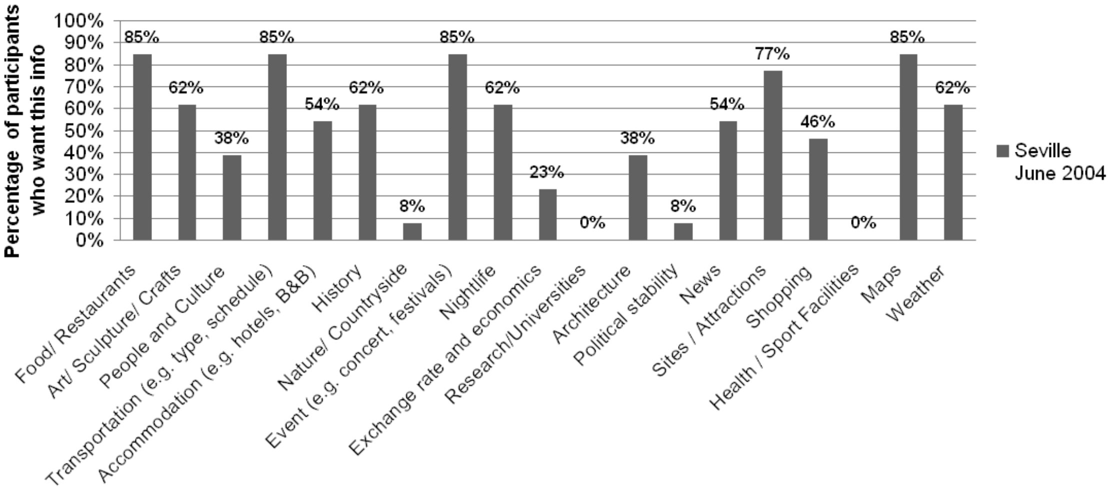


Figure 3.7: *Information types expected from a mobile device (AmbieSense Seville June 2004, pre-questionnaire)*

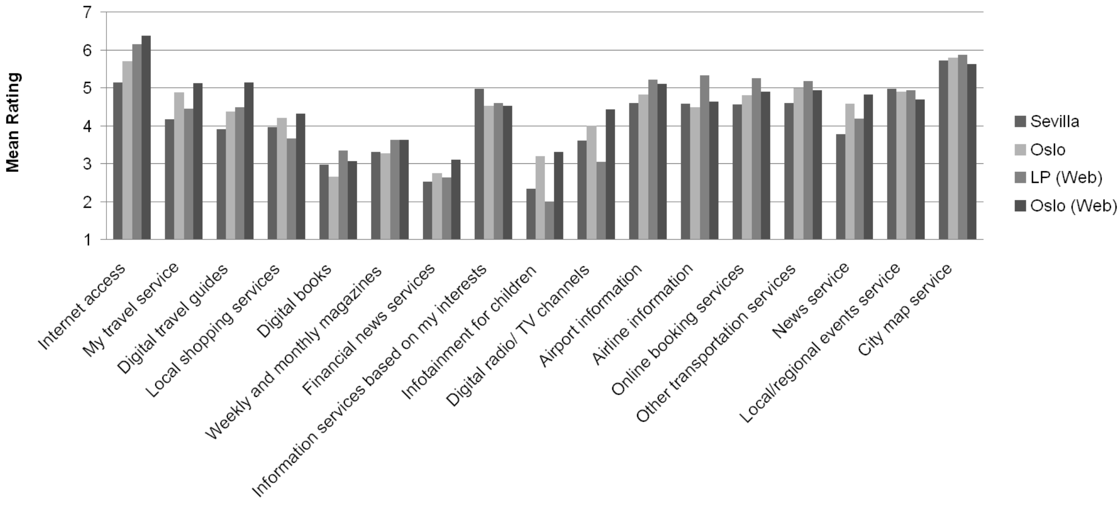


Figure 3.8: *Information service types – usefulness and fun to use (AmbieSense market survey for tourist city scenario)*

People showed generally very little interest in digital books, magazines, finance news, children infotainment but also digital radio and TV.

3.3.2 Information Access and Behaviour

Besides the relevance of certain types of content, it is also important to know when users acquire certain kinds of information and in which situation they prefer to access this

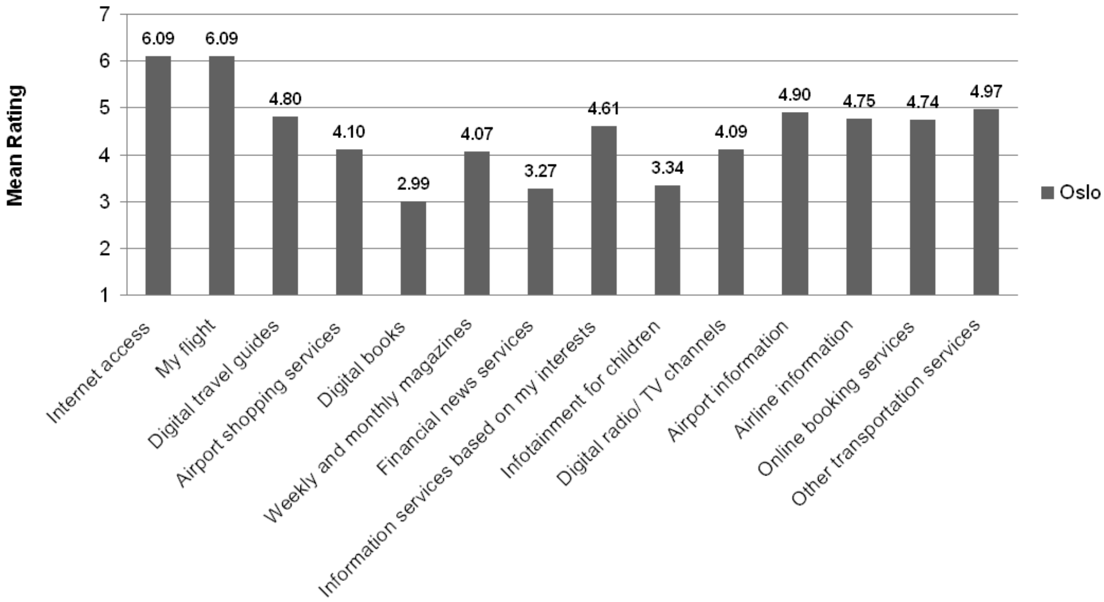


Figure 3.9: *Information service types – usefulness and fun to use (AmbieSense market survey for airport scenario)*

information. Such details about users' information behaviour may offer valuable clues if and how content should be delivered. It may provide insights on how much value mobile information delivery actually has for mobile users such as travellers.

Figure 3.10 depicts the type of information that users gather before and while travelling. Not surprisingly, information about accommodation is generally gathered before the trip. The same with information about transportation and weather; although not as prevalent. Information about restaurants, shops and entertainment events is acquired during the visit. This indicates that these three types of information are potentially best supplied to travellers and tourists while being at the location during their stay or journey. Mobile devices are potentially very useful for this kind of information supply in order to allow users direct and instant access to information about food places, shops and interesting events. Information about sites and attractions as well as maps are more neutral and appear to be acquired almost equally before and during the trip; again, this suggests mobile delivery, but in addition to other forms of delivery such as standard web access from home. An extended version of this question was included in the

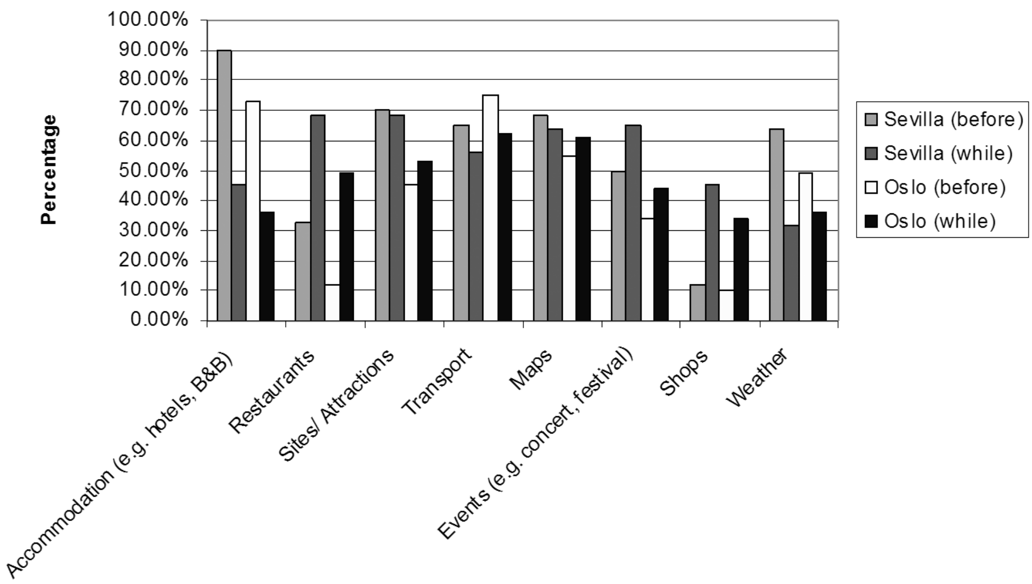


Figure 3.10: *Information types gathered before and during travel (AmbieSense market survey for tourist city scenario)*

pre-questionnaire at the June Seville user experiment (see figure 3.11). Ratings largely confirmed previous findings. Information about accommodation is clearly gathered beforehand by most people whereas transportation information is more neutral. On a moderate level, weather information appears to be more relevant during the trip. As previously seen, travellers gather content about food/restaurants, shops and events during their trip. Content about nightlife, one type of content that has certain relations to event content, is also clearly gathered during travelling rather than before. As seen previously, sites/attractions and maps are considered very important and are both acquired equally before and during travelling. Additionally, information about art and people/culture are acquired beforehand, whereas news and information about architecture are acquired during travelling. Information about history and nature/countryside are relatively balanced. Information about exchange rates, research/universities, political stability and health/sport facilities are considered neutral and are generally not often acquired. A second question inquired about the situation in which travellers and tourists want information. Figure 3.12 shows the results of the market survey for the tourist city scenario where there is a clear difference in which situations people want to access information. Information access was preferred from the hotel, from transport centres and

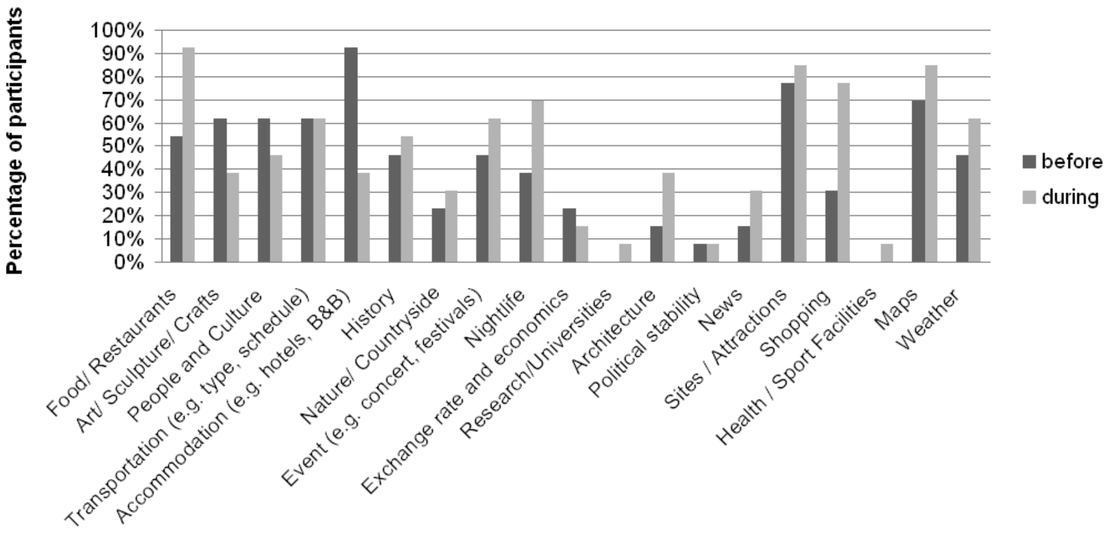


Figure 3.11: *Information types gathered before and during travel (AmbieSense Seville June 2004, pre-questionnaire)*

while not actively engaged in some other activity (e.g. shopping). Travellers did not want to access information while walking, however, people found it important that specific information is provided at specific locations on a tour. Figure 3.13 shows the results of the market survey for the airport scenario and largely confirms previous findings. Large interest in information was expressed while waiting (e.g. for arrival, for departure or at the gate), however, there was little interest in information while being engaged in other activities such as shopping.

3.3.3 Users' Willingness to provide Personal Information

In March 2000, the Personalization Consortium [10] conducted a personalisation survey. This survey showed that most web users are willing to provide personal information in order to receive personalised information from web sites. As part of the Personalization Consortium survey, 4500 web users were asked about their willingness to provide personal information to a standard e-commerce website. The published questionnaire and the data is provided in appendix C. The same question was asked in relation to a website that

[10] In the early stages of this research, the consortium was well presented at `http://www.personalization.org`, accessed October 2004. Now, the website is represented at `http://consortiuminfo.org/links/detail.php?ID=120`, accessed April 14, 2008. The research questionnaire with results has disappeared from the web itself, but a copy has been made and is provided in appendix C.

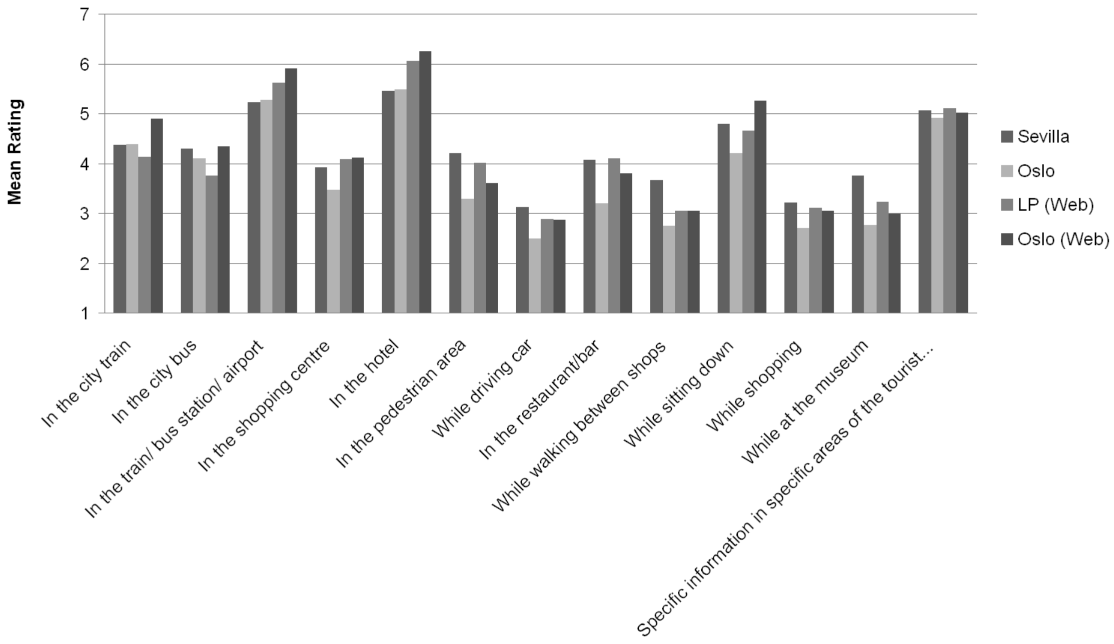


Figure 3.12: *Users' access pattern for travel and tourist information (AmbieSense market survey for tourist city scenario)*

provides personalised service to the user. The results from both questions are provided for comparison in figure 3.14. The Personalization Consortium survey revealed that most users are willing to provide name, email address, address as well as hobbies/interest to a website. Users tend to be more protective about their job and their phone number. Very private information such as the credit card number, the income, the mothers maiden name and social security number are unlikely to be provided. As soon as a website offers personalised services based on user information, users are more likely to provide personal information in almost all cases[11]. The largest difference in the willingness to provide personal information was for interests and hobbies (+25%).

Inspired by the Personalization Consortium survey, the question was adapted and refined to the needs of the project and the thesis and contributed to the AmbieSense questionnaires in two occasions – the two mobile studies conducted in Seville in June and in September 2004. People were asked to state their willingness to provide information about their name, address, age group, information about their educational background,

[11]The social security number is the only exception; although a very reasonable one.

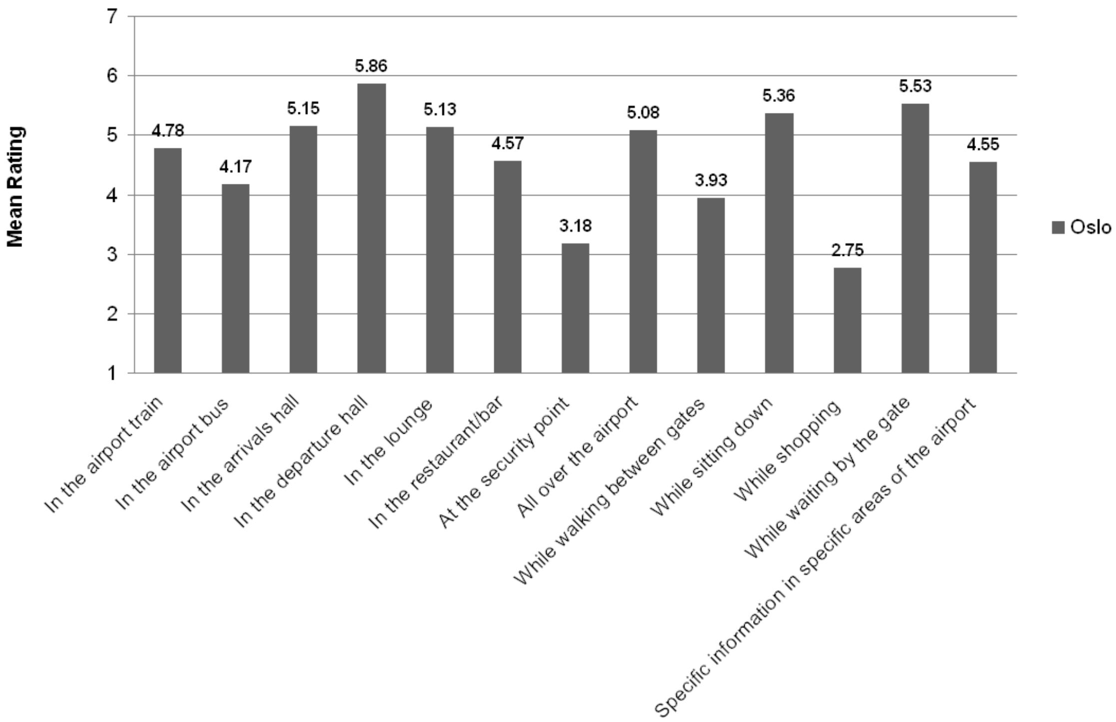


Figure 3.13: *Users' access pattern for travel and tourist information (AmbieSense market survey for airport scenario)*

job, interests and hobbies, and their social network. The question was included in the pre-questionnaire of two Seville user experiments (June and September 2004). In the Seville user experiment in June 2004, 13 travellers and tourists answered the question as part of the pre-questionnaire (see figure 3.15). Most of the 13 participants were happy to provide information about interests and hobbies (12 users; 92%) followed by the age group (9 users; 69%). Most people were more hesitant when it comes to the other types of information such as educational background and job category (4 users; 31%), address (3 users; 23%), name and social network (2 users; 15%). The question was repeated as part of the pre-questionnaire 3 months later in the Seville user experiment in September 2004 (see figure 3.16). Additionally, people were asked if they would be willing to provide their email address. Results turned out to be generally consistent in terms of the priorities. Most users were happy to provide information about their interests and hobbies (62%), followed by age group(49%) and email (39%). People tended to be more hesitant when providing information about the other types of personal information,

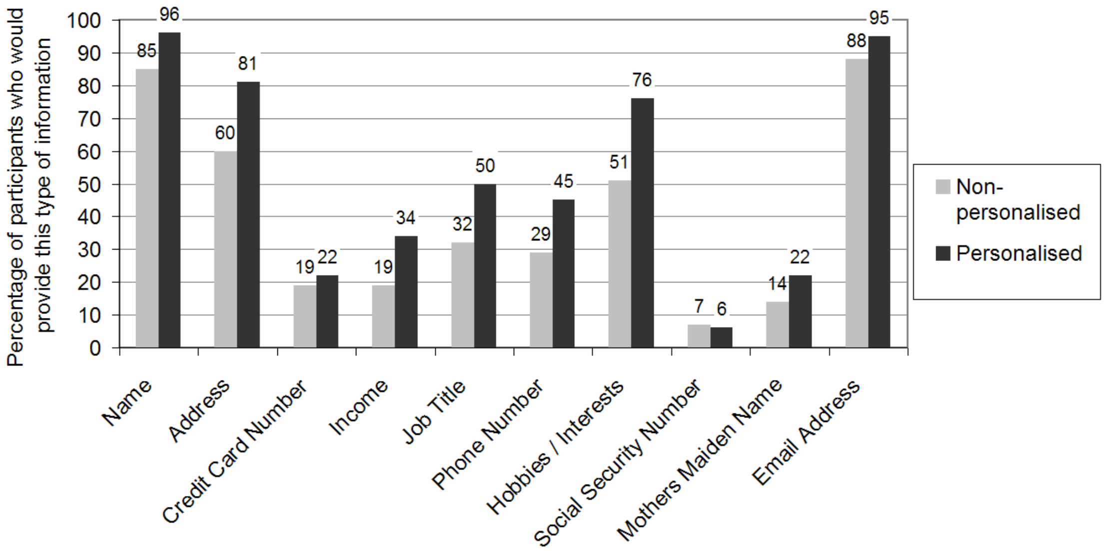


Figure 3.14: *Information types that web users are willing to provide to non-personalised vs. personalised services (Generated from web survey data published by the Personalization Consortium, Inc. in 2000)*

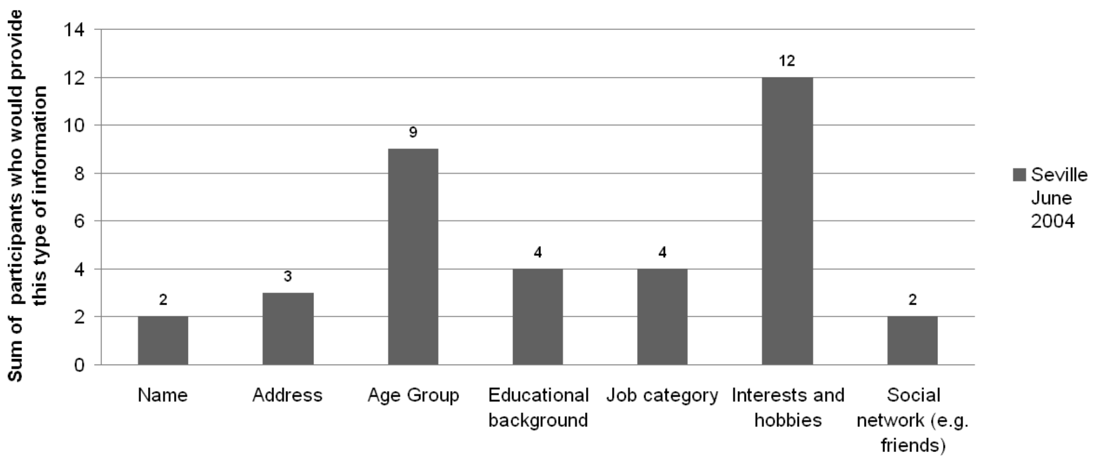


Figure 3.15: *Information types that users are willing to provide for personalisation (AmbieSense Seville June 2004)*

similar to the data collected in June. In comparison to the web survey performed by the Personalization Consortium, travellers and tourists who acted as mobile users were much more hesitant when providing information about their name, address and email address. This could be an indication that users are more concerned about their mobile identity than about their web identity. They were also more careful in stating what

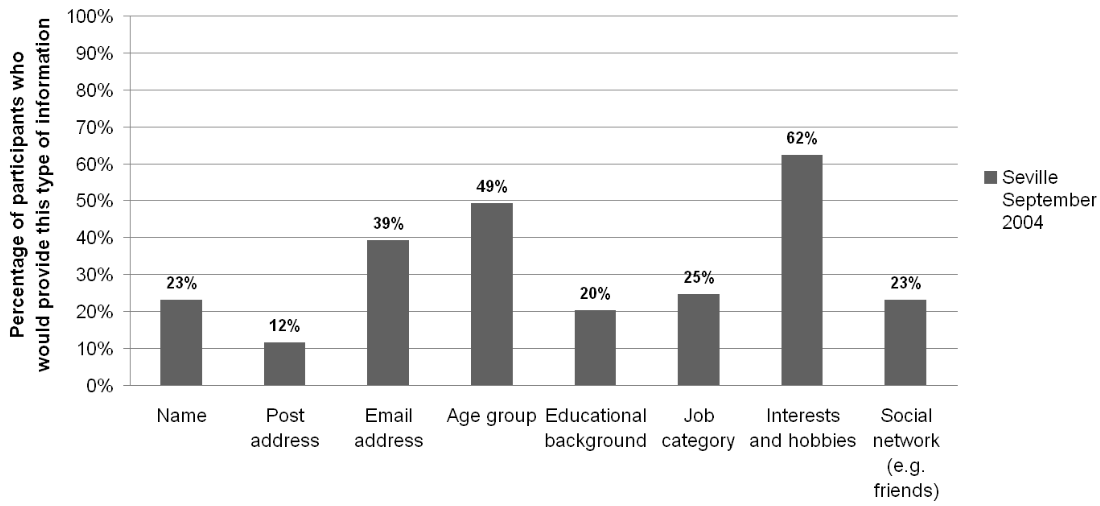


Figure 3.16: *Information types that users are willing to provide for personalisation (AmbieSense Seville Sept. 2004)*

kind of job they have even though this was actually less precise than the job title as it was asked in the Personalization Consortium's web survey; however, results show an overall consistency across all studies when it comes to information about interests and hobbies. The overall strong trend to provide information about personal interests for personalisation services (76%) is mirrored in both; 92% in the June Seville study and 62% in the September Seville study[12]. This demonstrates that there appears to be a consensus across different user groups about the notion of providing information about personal interest in order to receive personalised content. It explains the general acceptance of interest as an attribute in user modelling as it was described in the previous chapter and suggested in [Brusilovsky, 1996]. It indicates that interest is an attribute that is potentially very promising to use as part of a user context model as users are willing to provide information for it, allowing the model to be populated with accurate information that can make a personalised information system perform.

3.4 Discussion

Three particular questions of interest to this thesis were explored in this chapter along data that was collected during the AmbieSense project and data from an earlier web

[12]The September study represents a more reliable result due to a larger sample size of 76 people.

survey published by the Personalization Consortium. Firstly, it was investigated what kind of digital content was demanded by travellers and tourists. Secondly, it was explored at which point in time digital content is acquired by users and in which situations it is accessed and used. Thirdly, results were obtained about how willing people are in general to provide information about themselves in order to allow for effective personalised information services. Further below in this section, results are summarised and discussed with respect to these three questions after a general note regarding the limitations of the presented results.

As a limitation of the results presented in this chapter, note that gender distributions for the market survey and for the Seville June questionnaire are not completely balanced and some gender effect might exist in these data sets. Furthermore, all AmbieSense results shown in this chapter were collected from travellers and tourists. However, it is not unreasonable to assume that travellers and tourists share certain characteristics with general mobile users and therefore provide a reasonable population for initial investigations.

Users want digital content in context and have distinct preferences when asked about the *types of content* they expect; especially in relation to users' personal devices, results have indicated that content about transportation, food/restaurants, sites/attractions, maps and events are among the most demanded types. The AmbieSense project has explored a wide range of different content types including accommodation (e.g. hotels in Seville), food/restaurants (e.g. Tappas places in Seville), shopping (e.g. special offers at Oslo airport), maps (e.g. Seville city map) and transport information. It was beyond the scope of AmbieSense to explore in detail the entertainment sector which means content about festivals and performances (e.g. theatre, exhibitions, live music) was not part of the content collection that was provided and evaluated during the mobile user studies in Seville and Oslo. Based on that, the research in this thesis provides detailed studies for this content domain; an additional and promising type of content that is relevant not only for travellers and tourists but also for mobile users in general.

The studies showed that users want digital content in the right situation thus provided based on their context of use. There is a clear difference in the *acquisition style* of different content types. Information about food/restaurants, events, nightlife and shops is clearly gathered more spontaneously on site whereas sites/attractions, transport and map content is equally acquired during the planning stage of the trip and while travelling. Whereas the Internet PC at home or at work would be best for information gathering before the trip, a mobile device might help travellers to obtain information while travelling. This means that mobile information delivery of content about food, entertainment events, nightlife, shops, sites, transport and maps is a potential extension to standard information access. Travellers prefer content while being at the hotel, at travelling facilities, while waiting or sitting down, at specific and Relevant areas that have a relation to the content but not while they are engaged in other activities (e.g. shopping). This signals a demand for services that respond contextually and provide content based on users' personal situation.

At the same time, users are also willing to provide *personal information* for a system that offers such personalised services. Results showed that personal information such as name and address were not as easily provided by mobile users as it was by web users, however, the likelihood of users providing information about personal interests was high and consistent with previous findings. Overall, this demonstrates that interest is a kind of information that is well supported by users across different populations. It explains the wide-spread use of interest as a major user modelling attribute for personalised services [Brusilovsky, 1996]. As discussed in the previous chapter (section 2.2.2), context models allow modelling of the user and the users' environment. Context modelling is an active area of research and deeper understanding about the nature of individual attributes is necessary to build more effective information systems that are able to provide personalised information services for mobile users such as travellers and tourists.

3.5 Summary

This chapter explored three particular questions of interest to this thesis along results from a range of AmbieSense user studies and findings from an earlier web survey published by the Personalization Consortium. These results provided some useful insights about travellers and tourists as one type of mobile user before proceeding to studies about context and personalisation in the next chapters. The user studies reported in this chapter explored the types of content that mobile users expect, provided insights into how people acquire, access and use content and performed research on how willing they are to provide personal information.

4 A User Experiment on Contextual Usefulness

> The whole is more than the sum of its parts.
>
> *Metaphysics*
> ARISTOTLE

4.1 Motivation

The definition of context, its various interpretations and uses were described and reviewed along relevant literature in chapter 2. It was found that contextual information describes aspects of users and their environment. More specifically, context is modelled as a selection of attributes that describe the user and the environment in which the user interacts with information systems for Information Retrieval and Seeking (IR&S) in order to resolve an information need. The aim of this thesis is to investigate contextual relationships in the entertainment event and map content domain for mobile users. The results of this work can be used to build information systems that can help to deliver useful, personalised content to mobile users.

Context has gained increasing interest in the research areas of IR&S, ubiquitous computing, user modelling, artificial intelligence and adaptive hypermedia amongst others. Context modeling aims to create a better understanding of the contextual structures that is necessary for more effective application of context in information systems. Although several context models and context-aware systems exist, there are few experiments that empirically investigate the nature of individual attributes, connections between attributes and their effects on users' behaviours. Such investigations are desperately needed to better understand and model context and apply it more effectively.

This chapter presents an experiment based on three such attributes – time, location and user's interest – regarding its impact on user's perception of usefulness. The user-centric methodology of the *simulated work task scenario* [Borlund, 2000] is applied and further adapted for a simulated mobile scenario with event content. This content and usage domain appears as a very promising and representative area for context-aware computing. The effects resulting from these attributes were considerable and confirm the importance of context for an information system that aims to deliver personalised services to its users.

This chapter progresses as follows: In section 4.2, background and reasoning are provided for the selection of contextual attributes. Section 4.3 outlines the experiment design and the stimulus material that was provided. In section 4.4 and 4.5, the experiment procedure and the participants are addressed. Section 4.6 presents a detailed account of the results followed by the discussion of the numerous effects in section 4.7.

4.2 The Connection between User, Content and Context

The context model described in [Myrhaug and Göker, 2003] is kept general and categorises a large contextual spectrum into the five aspects of environment context, personal context, task context, social context and spatio/temporal context as described in section 2.2.2 on page 22. The model serves as a guideline that allows for application specific refinement. For this study, two of these aspects were investigated – the spatio/temporal context and the personal context. The application of personal context introduces space for personal variation and allows the model to be used for content personalisation. This choice enabled the further refinement of the model from the perspective of both the user and the content. In section 4.2.1 and 4.2.2, contextual requirements are discussed in the light of mobile usage and two types of content that are particularly interesting in such environments.

4.2.1 Usage Domain

The way in which people use an information system is important for the modelling of contextual attributes. A number of aspects appear especially relevant:

- *Cognitive Challenges:* Being mobile is cognitively intense since users often perform tasks simultaneously to other tasks (e.g. checking messages while walking). This requires mobile applications to provide more adapted and focused content since users cannot spend as much attention to the mobile application [Oulasvirta et al., 2005]. Contextual information can help to achieve this focus. In a mobile situation, users' location is one intuitive context attribute. AmbieSense studies in chapter 3 have shown that users also do not like to receive information while being engaged in other activities (e.g. shopping).

- *Spontaneousness:* In [Tamminen et al., 2004], it is identified that mobile usage is generally more spontaneous than desktop usage. Within a planned activity, users allow for spontaneous sub-activities. However, users' temporal and spatial flexibility is usually limited by these activities. A wider plan might provide the framework for all potential sub-activities (e.g. a business trip limits time to a few hours and a part of the city). AmbieSense results in chapter 3 also revealed that certain types of content are more likely accessed spontaneously (e.g. events) whereas others are planned before their use (e.g. accommodation). This suggests to use time and location as potential context attributes.

- *Device Limitations:* Mobile devices suffer from limitations in screen size, performance and storage capacity, and ways in which users can interact with the device. Whereas performance and storage limitations will eventually resolve, the problem of screen size and interaction will most likely remain. This calls for more focused content provision. Considering users' interest in the adaptation process is therefore especially important. This is also suggested by the wide application of users' interest within personalised information systems for content adaptation [Brusilovsky, 2001].

4.2.2 Content Domain

As well as the usage environment, the content is another important factor that has implications on the context model. A mobile computing scenario was selected as the focus of this study. Two particularity interesting types of content when being mobile are entertainment events and geographic maps as results from AmbieSense studies in the previous chapter have shown.

Event Content and its Special Features:

The Reuters Kalends event corpus was used in the study of this chapter. The collection consists of 10500 leisure time events divided in 39 topical categories. The following listing provides one example that was used as part of the user experiment reported here. Appendix E provides further information about the precise structure of the collection and more examples of Reuters' event content.
From this collection, a number of interesting features were identified with potential implications for a context model:

1. *Event Location/Venue:* Events usually occur in at least one place (e.g. a play in a theatre). Popular events are repeated at several locations to be available to a larger audience. This indicates that event content has a strong connection with location. The relation between the user's current location and the event venue has potentially a strong influence on the usefulness of an event.

2. *Event Performance Time:* Events usually occur at least one time (e.g. a book signing event by the author of a novel in a local bookshop). Popular events are likely to be repeated or appear even periodically on a regular basis. They can be planned in advance. There are many similarities between event content and news content. However, news content mostly describes incidents in the past that are reported afterwards. Based on that, event content can be associated with a particular time or time period. Furthermore, events can be planned and predicted. This indicates a strong connection between event time and current time with potential influence on the usefulness of the event content.

```
<event workflowID="3">
    <title>Ken Dodd - The Happiness Show</title>
    <description>Ken Dodd is much more than a comedian. He is a comedy
        genius and showbiz legend whose humour has made him one of Britain's
        best-loved entertainers. For his Diddymen, jam-butty mines and black
        pudding plantations, the Professor of Giggleology and Master of
        Applied Tickleology has been awarded The British Comedy Awards
        highest accolade - the Lifetime Achievement Award. Come and join the
        King of Comedy for more quick-fire gags than you can shake a tickle-
        stick at!</description>
    <startdate>2001-11-04T00:00Z</startdate>
    <enddate>2001-11-04T22:59Z</enddate>
    <sourceeventid>L249820425</sourceeventid>
    <allday>false</allday>
    <phonenum>+44 (0)1908 606 090</phonenum>
    <faxnum />
    <emailaddress>info@mktgc.co.uk</emailaddress>
    <uncertaintyperiod>0</uncertaintyperiod>
    <category categoryid="ENTF" />
    <location street="900 Midsummer Boulevard" postcode="MK93NZ">
        <pointofinterest
            poiName="Milton Keynes Theatre"
            poiPhone="+44 (0)1908 606 090"
            poiURL="http://www.mktgc.co.uk"
            poiEmail="info@mktgc.co.uk">
            <description>A new apound;30million theatre is being built -work
                 started in 1997 and it opened in 1999. The name was chosen
                as the result of an extensive local survey. Flexible seating
                capacity 950-1400. An ATG member.</description>
            <poiservices>Theatre Tokens, Infra-red system,
            Wheelchair access, Disabled toilets</poiservices>
            <image
                imageType="Exterior Photo"
                imageFile="http://www.dynamiclisting.com/uktw/venues/ex754.
                    jpg" />
        </pointofinterest>
        <country countryISO2code="GB" />
    </location>
</event>
```

Listing 4.1: *Example from Reuters event Kalends collection*

3. *Event Category:* Each event of the Reuters Kalends collection always belongs to exactly one of Reuters' topical categories. These categories are derived from the broad collection of available events and cover topics such as 'Musical', 'Dance' or 'Theatre'. The category naming and description is important in the sense that it serves as an important source of information for the end user. There may be a connection between these categories and the users' interest. For this reason, event categories could be matched with the personal interest of individuals. The existence of such a match would likely increase the usefulness of this event for this user.

Map Content and its Connection with Events:

Geographic maps have two interesting aspects that connect them with event content:

1. *Spatial integration:* Any content that is associated with a geographic location[1] can automatically be integrated in a map. This is given for events, since they are usually attached to at least one venue location.

2. *Temporal integration:* Besides the geographic integration of information in the map, it is also possible to integrate content based on their association with time. About 15 years ago, the field of GIS has started first efforts in this direction. Most GIS solutions included temporal information simply through snapshots of maps over time. Various data models have been developed as well as numerous prototypical systems for specific needs. However, today's best known and widespread GIS solutions such as Google Earth[2] or NASA WorldWind[3] still do mostly not support temporal features at all. Such methods, that integrate maps with time dependent map features, are still early in development with only few initiatives such as TimeMap[4]. Event content allows for temporal integration since an event has at least one particular performance time.

In addition to these two major reasons, maps provide intuitive visualisation allowing users to explore content spatially as an extension to the traditional way of text presentation.

Based on the considerations of the specific usage and content domain, the following context model is proposed as depicted in figure 4.1. This model consists of a promising set of relevant attributes based on the domain that was chosen for this experiment. The model contains the three attributes which are *location*, *time* and *interest*, as discussed above. These attributes have originated from two of the five broader contextual categories suggested in [Myrhaug and Göker, 2003] and confirmed by a closer look into the special requirements that emerge from the usage and content domain. This model covers the

[1]The method that associates valid geographic locations to information is called geocoding.
[2]`http://earth.google.com`, accessed April 14, 2008
[3]`http://worldwind.arc.nasa.gov`, accessed April 14, 2008
[4]`http://www.timemap.net`, accessed April 14, 2008

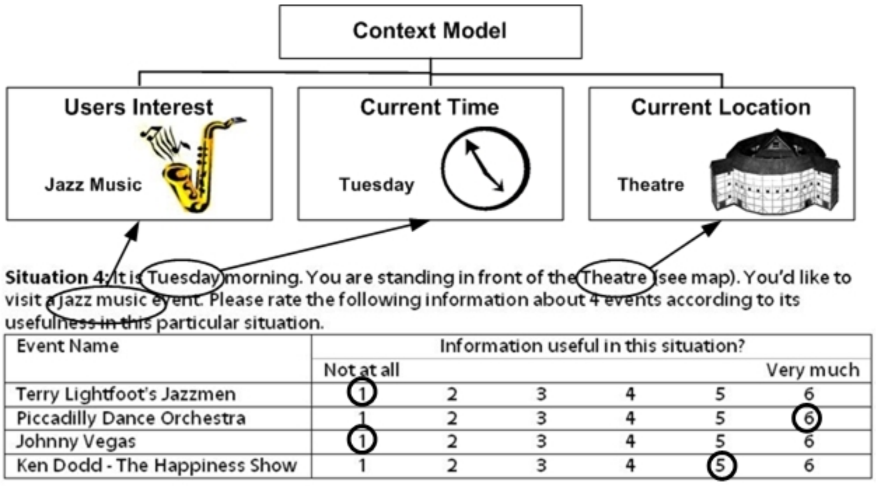


Figure 4.1: *Example of one instance of the context model with the three attributes for user's current interest, current time and current location as provided to participants. 'Map' refers to figure 4.2.*

user's personal context with one attribute (*interest*) and the spatio/temporal context with two attributes (*time* and *location*). These three contextual attributes are expected to influence the usefulness of geographically related event content in a mobile scenario. With these in mind, the following **four research hypotheses** emerge:

- H_1: Time has an effect on users' perception of event usefulness.
- H_2: Location has an effect on users' perception of event usefulness.
- H_3: Interest has an effect on users' perception of event usefulness.
- H_4: Time, Location and Interest interact with each other regarding users' perception of event usefulness.

It is not suggested that these attributes are finite or absolute. More and different attributes could have been proposed, such as the season of the year, the financial budget of the user or various aspects of users' physical or psychological states. However, it is assumed that these attributes provide a manageable set. The focused choice of attributes allows for a full investigation of all their effects. In future work, more attributes could be studied and combined with this basic set.

Although the context model is strongly motivated by mobile usage, it is important to point out that this model may also be applied in non-mobile usage scenarios. However, it is expected that a mobile environment offers a higher challenge and a better utility for the attributes.

4.3 Experiment Design and Method

To investigate the four hypotheses about the main effects of the three attributes and their potential interactions, a repeated-measures experiment with a full factorial design was applied. Participants in repeated-measures experiments perform in a range of experimental conditions (also called treatments). In this experiment, participants where asked to rate content items about entertainment events based on a range of different situations. Repeated-measures experiments minimize natural differences between participants and allow for high statistical power even from relatively small and moderate sized samples [Murphy and Myors, 2003]. In other words, choosing this experiment design allowed to collect enough ratings from a limited set of participants and not automatically violating the statistical meaningfulness of the data.

4.3.1 Experiment Setup and Scenario

For this experiment, simulated work task situations [Borlund, 2000] were applied in order to establish an informative environment that helps participants to create information needs and provides them with a framework for their judgements. Each participant received a background scenario together with a list of contextual situations. Whereas the background scenario described the broader setting of the experiment, the situations represented more detailed information. It was decided to present the background scenario as a festival – a typical "hotspot" for entertainment events of all kind.

Festivals as Event "Hotspots":

During festivals, large varieties of events for leisure time entertainment are provided. Activities and performances usually run within a relatively short period of time. Event locations are usually held within a shorter distance; most likely an dedicated area or a small number of places that provide appropriate space, facilities and general infrastructure. The Aberdeen Jazz Festival 2006 for example performed five days, providing a total of 30 performances at five different venues close to the city centre. The much larger Edinburgh International Festival is usually held over two weeks at a few main (and some additional smaller) venues.

Background Scenario and Contextual Situations:

To create an equally realistic setting for the tasks, a small fictional festival was chosen as a background scenario with a small set of events about *Jazz Music* and *Comedy Performance* distributed over a time of three days (*Monday*, *Tuesday* and *Wednesday*). In this scenario, events are performed at three different fictitious places (the *Theatre*, the *Community Centre* and the *Gallery*) located along a long street next to each other. All locations were only accessible by walking. To support the scenario visually, each user also received a simplified paper map with the event locations (see figure 4.2). The choice of values for

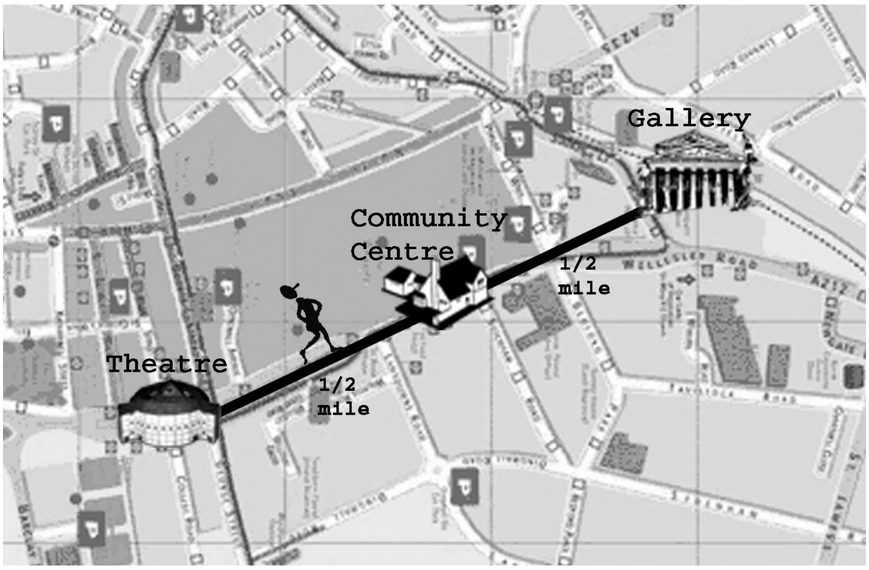


Figure 4.2: *Simplified paper map provided together with the background scenario*

time, location and interest were based on the following considerations:

- *Time:* Three days of the working week were selected instead of the weekend in an attempt to limit bias towards or against particular days. Saturdays tend to be more popular for entertainment and socialising than other days which could skew the experimental data. Although an event festival on weekdays might cause generally lower rates of usefulness, it will will more likely produce stable results.

- *Location:* The background scenario is more general and does not refer to any particular town. The three chosen locations resemble three generic locations that exist in most towns. This was necessary in order to obtain results that are not bound to one particular city but instead support generalisation. It also solves the problem that participants do not need any specific knowledge about a town and particular places in order to participate in the experiment.

- *Interest:* In real life, peoples' interest tends to be a very personal and dynamic variable. Since it was decided to investigate interest as one of the attributes, it was necessary to control its variation as part of the experiment. The two kinds of interest (*Jazz Music* and *Comedy Performance*) were inspired from the Reuters Collection. Both types of interest are distinct which is expected to allow participants to differentiate them.

Event Calendar:

Participants also received an event calendar with four different events that were extracted from the Reuters Kalends event collection. The content for each event consisted of a title, a short description, performance time, venue and the interest category as shown in figure 4.3. The event titles and descriptions were taken from Reuters' collection. The

Event no:	Comedy Event 2
Title	Ken Dodd - The Happiness Show
Description:	Ken Dodd is much more than a comedian. He is a comedy genius and showbiz legend whose humour has made him one of Britain's best-loved entertainers. For his Diddymen, jam-butty mines and black pudding plantations, the Professor of Giggleology and Master of Applied Tickleology has been awarded The British Comedy Awards highest accolade - the Lifetime Achievement Award. Come and join the King of Comedy for more quick-fire gags than you can shake a tickle-stick at!
Where?	Gallery 
When?	Wednesday, 8 pm

Figure 4.3: *Example of a event as provided to participants during the experiment procedure*

event performance time was one of the evenings of the three days (*Monday*, *Tuesday* or *Wednesday*) and the event location was one of the three places (*Theatre*, *Community Centre* or *Gallery*). Two of the events were about *Jazz Music* and two were *Comedy Performance*. In a real festival, this material could be handed out to people as part of an information brochure that describes the basic outline of the highlights together with a detailed account of the programme.

Eighteen different contextual situations (i.e. 3 possible times x 3 possible locations x 2 possible interests) were given to each participant. A situation is comprised of one of each of these attributes. Time being either *Monday*, *Tuesday* or *Wednesday*, location being either at the *Theatre*, at the *Community Centre* or at the *Gallery* and personal interest being either *Jazz Music* or *Comedy Performance*. The example in figure 4.1 shows one situation where the participant, located at the Theatre on a Tuesday, is interested in Comedy Performance. The eighteen situations expressed therefore eighteen different instantiations of the context model.

4.4 Experiment Procedure

The experiment was performed at various locations at the university – mainly offices, lecture facilities and the cafeteria area of the business school – where people were approached and invited to take part in the experiment. The background scenario, contextual situations and the event calendar, as described in section 4.3.1, were explained and handed out to participants on paper. They were asked to rate the usefulness of the four different leisure time events for each of the 18 different situations thus providing a total of 72 ratings. See figure 4.4 for a general overview about the procedure and an example rating in the lower part of the figure. In preparation for the task, participants were introduced to usefulness as situational relevance. According to Borlund, situational relevance "...expresses the relationship between the users perception of usefulness of a retrieved information object, and a specific work task situation" [Borlund, 2003a, p. 922]. The situations contained information about the current time, the current location and the current focus of interest. Each situation was embedded in the festival scenario as it was described in the last section. For all these situations, participants were required to rate the usefulness of each event. The rating was scaled along a 6-point rating scale that was ranging from 1 ("Not at all") to 6 ("Very much"). The scale forced a decision from participants based on the missing middle score. The order of situations was randomised in an attempt to limit potential ordering effects. Participants completed the ratings on

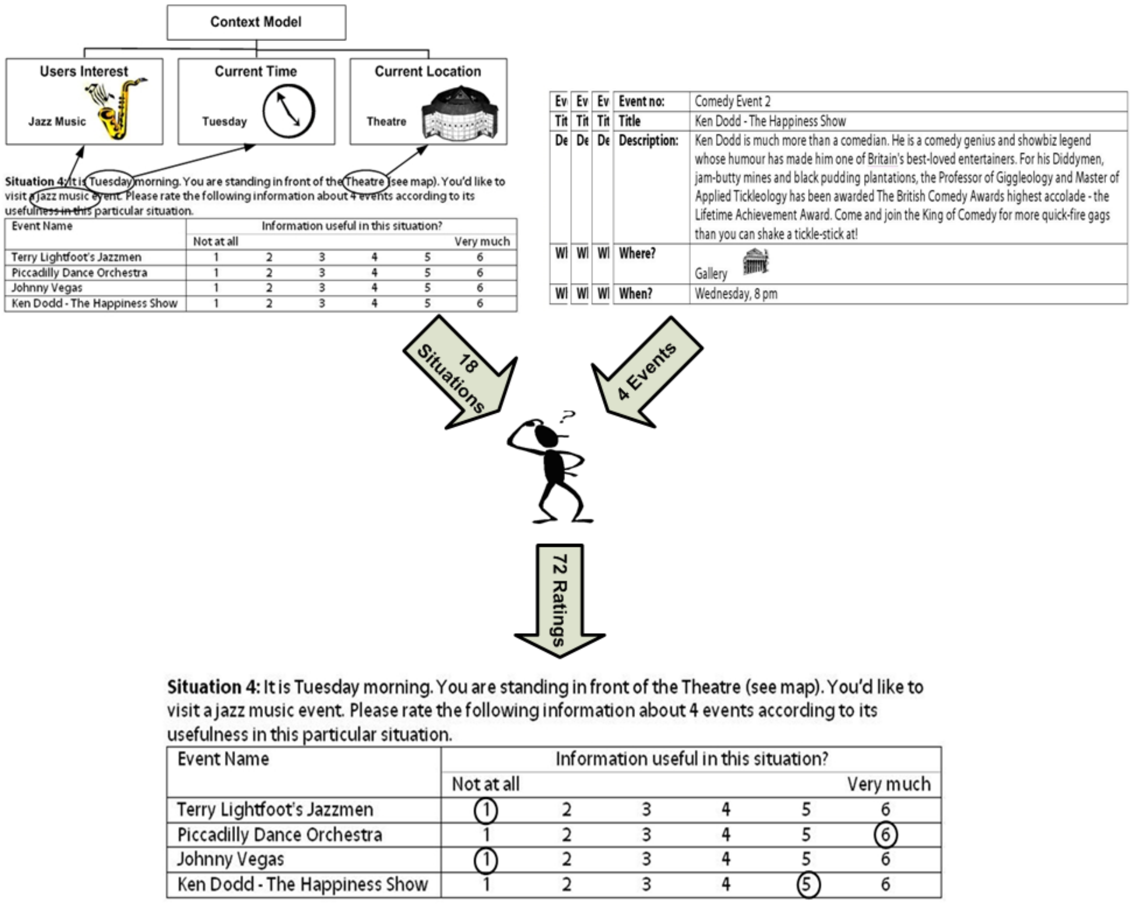


Figure 4.4: *Overview of the experiment on contextual usefulness*

paper by themselves and returned the forms to the experimenter afterwards in exchange for a coupon for a free drink at the university cafeteria.

4.5 Participants

The participants for this user study were 32 individuals chosen from the student and staff population at various faculties of the Robert Gordon University in Aberdeen/Scotland. They were 20 to 49 years old (75% were 18 to 29 years), 14 male and 17 female[5] as shown in figure 4.5. In an ideal setting, it would have been desirable to include also people from outside the university. The limitations in financial resources however did not allow for this given that each participant had to spend an average of 45 minutes in exchange for a coupon for a free drink. Nevertheless, it should be appropriate given that the target

[5]One participant did unfortunately not provide demographic information.

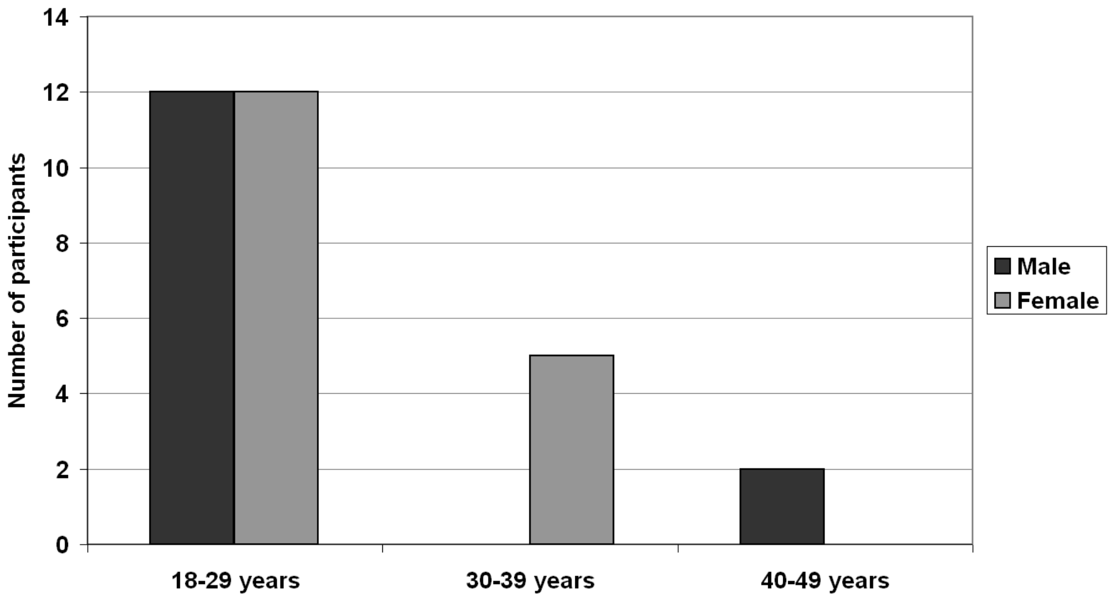


Figure 4.5: Age and gender distribution of 31 participants of the experiment (1 participant did not provide demographic data)

population are mobile users in general. Also, participants were sampled from different faculties which helps to minimise the possible bias from computer science students who might be familiar with mobile and context-aware computing.

4.6 Results

This section provides an overview to the data, a graphical representation of contextual effects, and a table with the degrees of statistical significance and corresponding effect sizes.

4.6.1 Data Overview

The participants provided ratings of usefulness for event content based on a set of situations that were defined along time, location and interest. Participants provided these ratings on paper that was given to them at the beginning of the experiment.

For each user rating, the differences were calculated between the time of the situation and the event performance ($time_{situation} - time_{event}$), the location of the situation and the event venue ($location_{situation} - location_{event}$), and determined if there was a match or a non-match between the interest of the situation and the event interest category. This resulted in 5 possible differences for Time (-2, -1, 0, 1 or 2 days of difference), 3 possible distances for Location (0, 1 or 2 places of distance) and a binary possibility for Interest being either matching (0) or non-matching (1).

Figure 4.6 shows the usefulness ratings of all participants for all situations they were presented with. In other words, it represents the summary of all effects of Time (T), Location (L), and Interest (I) on participants' usefulness ratings for event content. In particular, it shows the magnitude and change of all three attributes graphically on the mean event usefulness as it was rated by participants. The graph shows the mean usefulness (average rated values of usefulness) assigned by participants for matching interest (I = 0) in the upper part of the graph and for non-matching interest (I = 1) in the lower part of the graph. The error bars indicate standard errors. As mentioned before,

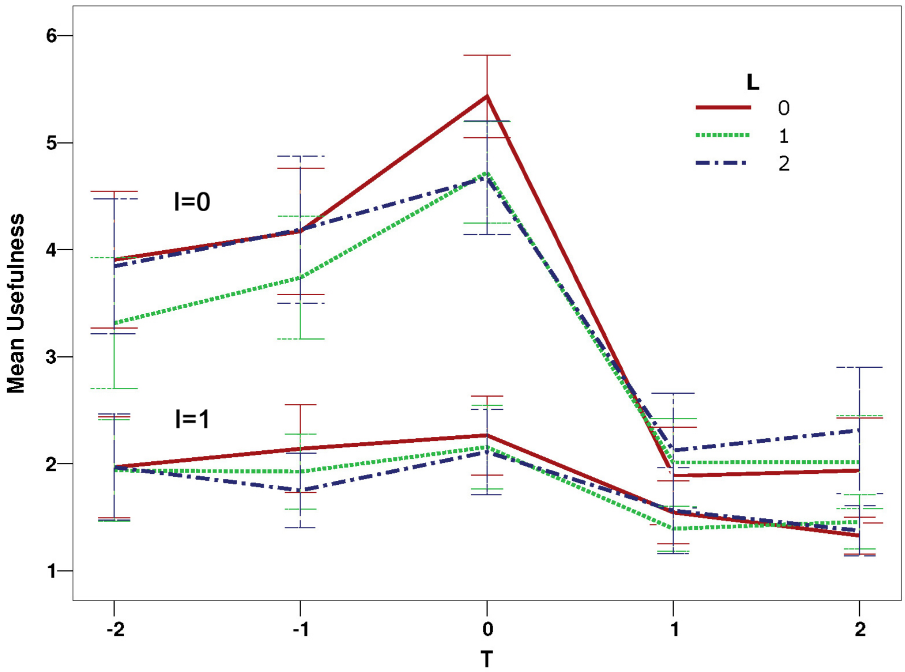


Figure 4.6: *Shows the impact of all combinations of contextual attributes on usefulness ratings. Mean rated usefulness for matching interest (I=0) and non-matching interest (I=1) for 5 levels of Time (T) and 3 levels of Location (L). Error bars indicate standard errors.*

the study was composed of 5 differences for time, 3 distances for location and a binary interest being either matching or non-matching. The time levels (T-Levels) are provided in days and 4 different level changes are possible between the 5 distinct time differences. Similarly, the location attribute provides two different level changes (L-Levels) from 0 to 1 and 1 to 2 places of difference. The interest attribute can only change from matching interest (0) to non-matching interest (1) which means that the repeated contrast is equal to the overall effect of interest (I-Levels).

Effect	**T-levels** (difference in days)	**L-levels** (difference in places)	**I-levels** (match)	**F**	**effect** (η_p^2)	**sig.**
T(overall)				**38.299**	**.553**	**.000***
T(contrasts)	-2 → -1			2.957		.095
	-1 → 0			28.185		.000*
	0 → +1			81.991		.000*
	+1 → +2			0.128		.723
L(overall)				**3.872**	**.111**	**.042***
L(contrasts)		0 → 1		9.495		.004*
		1 → 2		5.459		.026*
I				**95.388**	**.755**	**.000***
T x L(overall)				**3.379**	**.098**	**.007***
T x L(contrasts)	-2 → -1	0 → 1		0.006		.939
		1 → 2		0.406		.527
	-1 → 0	0 → 1		0.372		.546
		1 → 2		1.332		.257
	0 → +1	0 → 1		9.502		.004*
		1 → 2		1.518		.227
	+1 → +2	0 → 1		2.362		.135
		1 → 2		0.056		.814
T x I(overall)				**26.497**	**.461**	**.000***
T x I(contrasts)	-2 → -1		0 → 1	3.513		.070
	-1 → 0		0 → 1	12.264		.001*
	0 → +1		0 → 1	63.717		.000*
	+1 → +2		0 → 1	8.641		.006*
L x I(overall)				**2.217**	**.067**	**.139**
T x L x I(overall)				**2.420**	**.072**	**.049***
T x L x I(contrasts)	-2 → -1	0 → 1	0 → 1	1.343		.255
		1 → 2	0 → 1	0.077		.783
	-1 → 0	0 → 1	0 → 1	2.845		.102
		1 → 2	0 → 1	3.596		.067
	0 → +1	0 → 1	0 → 1	8.817		.006*
		1 → 2	0 → 1	0.041		.841
	+1 → +2	0 → 1	0 → 1	1.596		.216
		1 → 2	0 → 1	4.922		.034*

Table 4.1: *ANOVA results with contrasts for time (T), location (L), interest (I) and their interactions. Statically significant effects are labelled with asterisks (*).*

4.6.2 Detailed Account on Context Effects

The ANOVA results and relevant repeated contrasts for the main effects of time (T), location (L), interest (I) and their interactions (T x L, T x I, L x I and T x L x I) are listed in table 4.1. Contrasts offer more detailed information about statistical significance between individual factor levels. The table provides only contrasts for statistically significant overall effects ($p<.05$). In other cases, contrasts are not statistically significant and are therefore not listed. The table lists the F-values (F) and the p-values of statistical significance (sig). Whereas the F-value represents a distance measure between individual data distributions, the p-value expresses the significance of this difference in statistical terms. Both values are related and as a rule, the larger the F-value the smaller the p-value and the more significant the two distributions. The p-values of overall effects are corrected by the Greenhouse-Geisser method since the data did not provide equal differences between treatments levels – quite common when using factors with more then two levels. Furthermore, the table provides effect sizes for all main effects and interactions (expressed in the common partial eta squared measure η_p^2). Effect sizes express the strength of an attribute or an interaction in relation to the rated usefulness.

Based on the ANOVA table, the effects caused by the attributes of the model are now discussed. When necessary, references will be provided to the participants' ratings as depicted in figure 4.6.

- The effect of *time* on the judgement of usefulness (T) has high statistical significance ($p<.001$) and an effect size of η_p^2=.553. The contrasts show that significant effects only exist for the time difference between the day before and the same day as well as between the same day and the day after the event. These two changes can be seen as a strong rise directly before the day of the event and a very sharp decline after the event performance in figure 4.6 although much less pronounced with non-matching interest (I=1).
- The effect of *location* (L) is also statistically significant ($p<.05$) and its contrasts confirm this between each level pair. The effect size of location is of smaller magnitude (η_p^2=.111) in comparison to time and interest. When viewing the graph, the effect of location can be seen as the distance between the individual lines. Each

of the lines represents the change of usefulness over time at one location.

- The effect of *interest* (I) also revealed to be highly significant with $p<.001$ and the strongest of all main effects with η_p^2=.755. This can be viewed in the graph when comparing the upper part of the graph (I=0 or matching interest) with the average ratings visualised in the lower part of the graph (I=1 or non-matching interest).

Further, there is rich interactive behaviour between all three contextual attributes:

- The **three-way interaction between all three factors** (T x L x I) showed up statistically significant with $p<.05$ and an overall effect size of η_p^2=.072. Its contrasts are statistically significant in two cases. The first interaction ($p<.05$) exists between the same day and one day after the event when the location difference changes from the same place to one place and matching interest changes to non-matching. In this case, matching interest causes a stronger decline in usefulness in comparison to non-matching interest. This effect can be viewed graphically when comparing the two groups of lines in figure 4.6. Whereas for matching interest (I=0) the difference between matching place and one place difference collapses shortly after the event, for non-matching interest (I=1) it remains almost constant. The second interaction occurs at the end of the curve between one day and two days after the event where the usefulness rises for interested participants being two places away from the venue but not for those whose interest does not match.

- The **two-way interaction between time and location** (T x L) is statistically significant with with $p<.05$ and an effect size of η_p^2=.098. The contrasts show significant interactions when time changes from the same day to one day after and the location difference from the same place to a one place difference.

- The **two-way interaction between time and interest** (T x I) is statistically significant with $p<.001$ and shows in almost all cases statistically significant interactions. It also represents the strongest interaction effect with η_p^2=.461. When participants time context changes from one day before to the matching day, the usefulness rises faster when interest is met then when interest is not met. The opposite happens shortly after the event when the time context changes from the matching day to one day after. In this occasion, usefulness declines much stronger when participants' interest was met in comparison to the case when participants' interest is not matching. Also there is a statistically significant interaction at the end of the time line when the location changes from one place to a two place distance and matching interest changes into non-matching.

There is no statistically significant interaction between location and interest (L x I) as a direct result of the experiment. Consequently, contrasts for this interaction are not included in table 4.1. Although not significant in the strict sense, there seems to be a trend for the existence of interactive behaviour between location and interest. However, the effect size is consistent with the significance level indicating only a very small effect imposed by this interaction. It is possible that such an interaction might exist in other experimental settings that involve users' knowledge and long term behaviour.

4.7 Discussion

This user study was conducted to obtain a better insight into contextual attributes and their effects on people. In particular, the following four hypotheses were tested in order to investigate the impact of three carefully selected context attributes on users' perception of usefulness (or situational relevance as described in [Borlund, 2003a]).

- H_1: Time has an effect on users' perception of event usefulness.
- H_2: Location has an effect on users' perception of event usefulness.
- H_3: Interest has an effect on users' perception of event usefulness.
- H_4: Time, Location and Interest interact with each other regarding users' perception of event usefulness.

The study was focused on event content provided with a geographic map since it was decided to select a setting that is relevant for mobile computing.

All three attributes revealed statistically significant effects. It also turned out that all three context attributes have high order interaction effects between them. Also results show that there is a priority between the three attributes; interest being the strongest of all attributes, followed by time and then location. In the following, all context attributes and contextual interactions are discussed based on the order of the hypotheses:

- *Time* caused an overall large effect on users' perception of usefulness. The attribute was expected to cause higher degrees of usefulness before the performance time of the event in comparison to the time after the event. Furthermore, its peak usefulness was expected to be when the time of the situation matches with the performance time of the event. These expectations are confirmed by the findings and show strong evidence for time having an effect on users' perception of usefulness (H_1). It is very interesting that the style with which event usefulness is rising and declining seems not to be linear. Between two days and one day before the event rising is much slower in comparison with one day before and the matching day. After the event, usefulness is declining very strong between the matching day and one day after the event. This continues between one day after and two days after the event, however in much smaller magnitude. This pattern indicates a nonlinear effect of time on event usefulness and is confirmed by the shape of the data at every location and interest level.

- *Location* showed statistically significant effects on the amount of usefulness (H_2). However, the attribute has generally a much lower impact than the other attributes. This is both indicated by the weaker level of statistical significance as well as the associated effect size. One reason for this can be the distances between locations which have not caused participants to consider remote events of much lesser use. Also, the rather large time frame of several days must have caused location to be less influential when the event was still in the very far future. It is also possible that the simulated nature of the experiment caused people to underestimate the importance of location. The structure of the effect of location is also interesting. The assumption was that people would favour an event being local in comparison with the same event being more distant. Location-aware systems generally work on this premise when extracting and processing information. This expectation can only be partially confirmed by the data. In cases where the user's interest is met and the event has not yet been performed, local events are rated highest.

Usefulness drops when the distance to the event location increases by one place in this condition. However, events that have the maximum distance surprisingly increase again in their usefulness instead of dropping. After the event performance, the effect of location is actually reverted in respect to the original expectation. Events at maximum distance are rated highest and at the current location lowest. When the user's interest is not matching the event, the effect of location follows the normal pattern (further distant implying less useful). However after the event, this is not the case. One obvious reason for this must be that the event location has more relevance before the event than after. This has potentially also caused the rise of usefulness between one and two places of distance. Participants might have paid only limited or no attention on the precise quantitative rating after the event because of its low use. This shows evidence for a connection of location with the other two attributes; particularity with time.

- *Interest* has a profound effect on participants' rating of usefulness (H_3) indicated by the largest of all effect size. Matching interest produced a strong rise of usefulness indicating the intense impact of the attribute on participants' opinion. This confirms with the literature, in which interest was frequently used as one of the main attributes for content personalisation [Brusilovsky, 2001].

- As already described in the last three points, it was possible to obtain comprehensive statistical evidence about strong and manifold *interactions between all three attributes* in almost all cases (H_4). The data confirms a very strong interaction between time and interest as well as smaller interactions between the other two-way as well as the three-way interaction. This clearly indicates that a model of context cannot be derived by the simple combination of time, location and interest. It is necessary to include interactions between attributes. The strong connectivity between this rather focused set of attributes shows an example of how complex a context model can evolve with only three components.

This study is limited to some degree with respect to realism. The impact of the interest attribute, for example, might even be stronger if participants would have been sampled from particular interest groups (e.g. members of a comedy club or people from the audience of a jazz performance). This could be explored in a separate study perhaps embedded in a real festival stetting. Also, as already mentioned above, the impact of the location attribute might have been weakened by the fact that participants were provided with an artificial map rather than a map with associated meaning to real distances. It would be worthwhile to repeat this experiment in a real setting engaging participants that are fully aware of the space that a realistic map represents. This is one of the reasons why the mobile experiment reported in chapter 6 has been taken to the field and was equipped with a real map of Aberdeen/Scotland where it was performed. Additionally, participants had to walk to two of these locations to perform their information searches in the exact place their context described.

4.8 Summary

Starting from a broader and more general context model, three promising contextual attributes were investigated – time, location and interest – based on their impact on users' perception of usefulness. Since mobile computing (e.g. as in location-based services) is one very promising application area for context, the four hypotheses were evaluated along a simulated, mobile scenario as the basis for a task in which participants rated content about entertainment events. The study showed that time, location and interest matter to users in mobile situations and data analysis showed statistically significant effects. There appears to be a priority emerging in the relative importance of these attributes for the mobile user. Also, the results show high order interaction effects between the attributes. This experiment has provided an insight into the dynamics of a context model along a relevant and promising scenario for the application of context. In the next chapter, the data will be further explored in order to develop a quantitative, predictive model about the influence of contextual information on perceived usefulness.

5

A Personalisation Model from Context

All models are wrong, but some are useful.

Science and Statistics
George E. P. Box

5.1 Introduction

The previous chapter established an initial context model based on how people use an information system and on the content this system processes for its users. A more general context model [Myrhaug and Göker, 2003] was refined along a subset of attributes – taken from personal and spatio/temporal context. During this refinement, three context factors were identified – *time*, *location* and *interest* – and investigated in a laboratory user study. In this study, 32 participants rated leisure events based on a set of simulated work task situations as described in [Borlund, 2000]. These situations were part of a simulated jazz and comedy festival. Each situation represented one instance of the context model and was defined by a combination of a time, a place and a topic of interest. The situations together covered all combinations of a set of representative times, locations and interests in a full factorial repeated-measures design (see section 4.3). Results showed that time, location and interest had important effects on people's judgements individually as well as in interaction. This confirmed the validity of the factors as part of the context model and provided an interesting insight into the dynamics of such a model.

In this chapter, findings from the previous user study are used to create a predictive model of context for personalised search. The model is connected to some of the theory that describes the human process of explanation finding, or attribution theory. Its basic concepts and methodology are related and applied to context modelling. The model may serve as a guideline for the development of applications that process information about events and provide them to users in a mobile setting or similar application areas. The context model can be useful to predict the likelihood with which a particular event would satisfy the need of a user in a particular situation.

The next section introduces attribution theory as a theoretical framework that is one possibility to represent context modelling as the human process of causal explanation, thus viewing it as a human activity that appears to have similarities with context modelling. The theory of attribution is highlighted from the perspective of Harold H. Kelley, who significantly enriched the theory with key models and tools. Kelley's principle of covariation links the theory with the statistical technique of factorial ANOVA as applied in the previous study and also relates the theory with the method of multiple regression that is applied in this chapter. Section 5.3 further describes this regression technique that is used to determine a score of contextual relevance for the usefulness of event content items based on a situation. Section 5.4 shows how the model was integrated with one common information retrieval algorithm that allows the re-ranking of information search result lists into personalised result lists based on the user's current situation.

5.2 Causal Attribution: Context as an Explanation Process

One property of humans is that we constantly seek plausible explanations for the events and phenomena that happen around us. When receiving good results for an exam, we might account the quality of the results to our skills or to the amount of time we spent in preparation. The same exam with poor results however might be explained by its difficulty, a bad lecturer or simply bad circumstances. Psychology is the research field that is primarily interested in understanding humans by developing and testing theories about their mental processes. One of these theories investigates the human process of finding explanations for phenomena. It is called causal attribution theory or attribution theory for short. The theory particularly "deals with the information [people] use in making causal inferences" [Kelley, 1973]. Its origin dates back to the 1950s and Fritz Heider who wrote a book on "The Psychology of Interpersonal Relations" [Heider, 1958]. Among the many branches of attribution theory, Kelley's view [Kelley, 1967] is particularly compelling as it offers a more detailed framework of procedures and templates. Kelley distinguishes two cases of how people attribute causes to an effect – Covariation and Configuration. The next two sections look into these two types of attributions in more detail[1].

5.2.1 Covariation with ANOVA

In this case, people are confronted with a number of known potential causes and a single effect that was observed repeatedly over a period of time. Kelly himself describes the principle of covariation as an "effect [that] is attributed to the one of its possible causes with which, over time, it covaries." [Kelley, 1973, p. 108]. Kelley explains this behaviour with the assumption that people basically act as naive scientists who employ basic statistical methods to explain the world around them (i.e. perform attributions). This statistical model allows people to validate their explanations by measuring the covariation between the effect and each cause. Kelley postulates the use of the classic fixed model Analysis of Variance (ANOVA) to be used to measure the covariation of presence or absence of causes with the effect. Kelley identified a number of classes for possible causes that were able to explain a wide range of different attribution problems. These classes were *persons*, *entities* and *times* often depicted as the three dimensions of a cube (see figure 5.1). The 'persons' dimension of the cube represents how many people experienced an effect based on the same stimulus (also called consensus). The 'entities' dimension expresses the level of uniqueness of an effect with the stimulus (also called distinctiveness). The cube's 'time' dimension represents the effect over time or at different points in time (also called consistency). The following example, adapted from [Frieze and Weiner, 1971], attempts to demonstrate Kelley's cube in more detail. The example consists of three attributions where each of them refers to one cube in figure 5.1 (a, b and c):

[1] Later it will become clear that these two cases are related.

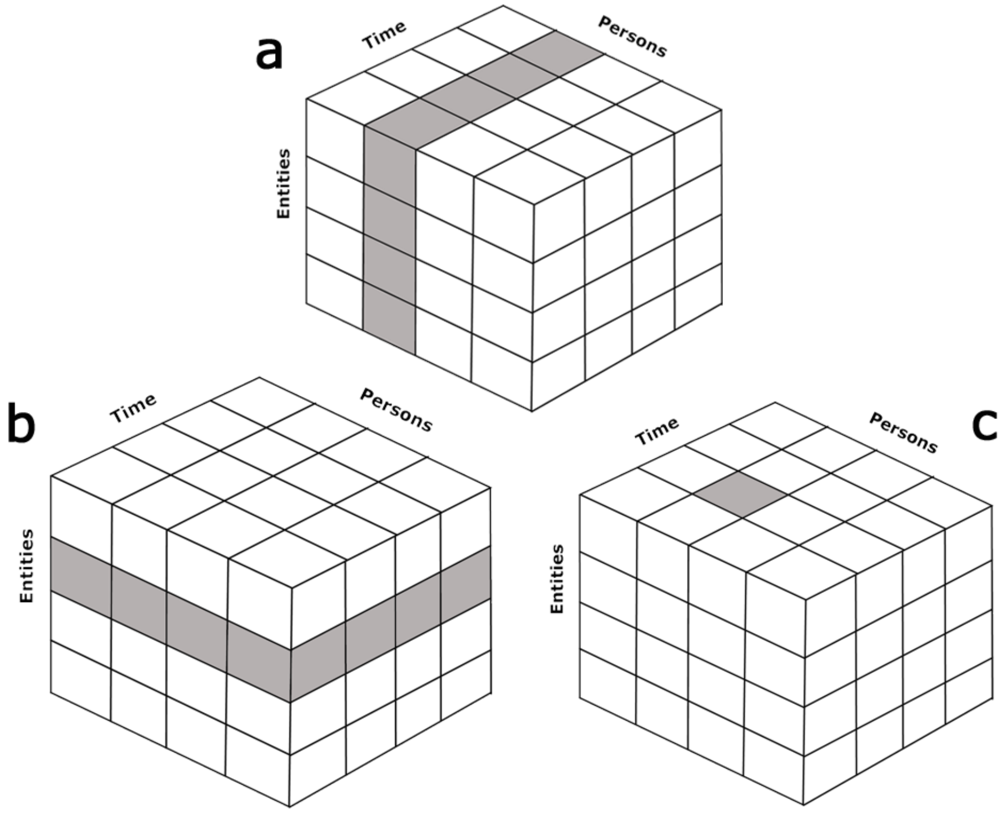


Figure 5.1: *Three different attributions as adapted from [Frieze and Weiner, 1971]. One effect each attributed to the person (a), the task (b) and the situation/circumstances (c). Effects highlighted in grey.*

1. In the first cube (a), a person experienced success (the effect) with one task (Entities) and had also successfully solved similar tasks in the past (Time), but only few other people (Persons) had equal success. The effect in this case was therefore explained within the person.

2. Cube (b) shows a situation where a person succeeded with a single task repeatedly at some specific point in the past together with other people, but has now failed in a more recent task. In such situations, the person explained the failure with the task and not within the person.

3. Cube (c) depicts a situation where a person has succeeded with a single task only at one particular point in time and did not repeat this success in similar or other tasks nor did other people do so. This situation was attributed to other reasons such as

bad luck and other variable circumstances.

Generally, a person validates the quality of an attribution through confidence. In particular, confidence is built when the person's response is distinctively associated with the stimulus, the person's response is similar to other peoples' responses and the person's response is consistent over time. Thus, the process of finding explanations is validated based on the variables modelled by the cube.

Nevertheless, the covariation principle is idealised as it is based on multiple observations of the same effect. It is not always possible or feasible to observe effects repeatedly. People may sometimes lack the opportunity, the time and the motivation to consider multiple observations before deriving explanations. The next section introduces causal schemata – small, simplified cause/effect templates that operate on the basis of the covariation principle as a shortcut in cases of limited data such as single observations.

5.2.2 Configuration with Causal Schemata

Sometimes the outcome may follow from less predictable causes where multiple observations were not possible or simply not feasible. Here, Kelley introduced causal schemata – hypothetical matrices that relate the presence or absence of factors to the effect.

Causal schemata are simple and pragmatic tools that represent a rule-of-thumb about how factors cause particular human behaviour. It is "an assumed pattern of data in a complete analysis of variance framework" [Kelley, 1973, p. 115]. Schemata are therefore specialised templates that operate on the more general framework of covariation described in subsection 5.2.1. Although schemata operate on the covariation principle, they do so with less data. However, even if a person attributes certain factors to an effect only based on a single observation, it is unlikely to be completely random. It is reasonable to assume that this person has experienced similar effects in the past, has some common knowledge about possible causes or simply acts intuitive.

One basic causal schema is shown in figure 5.2; it has two possible causes and it represents a case where an attribution is only performed when both causes are present (i.e. a logical AND). Schemata are not restricted to two causes. The schema depicted

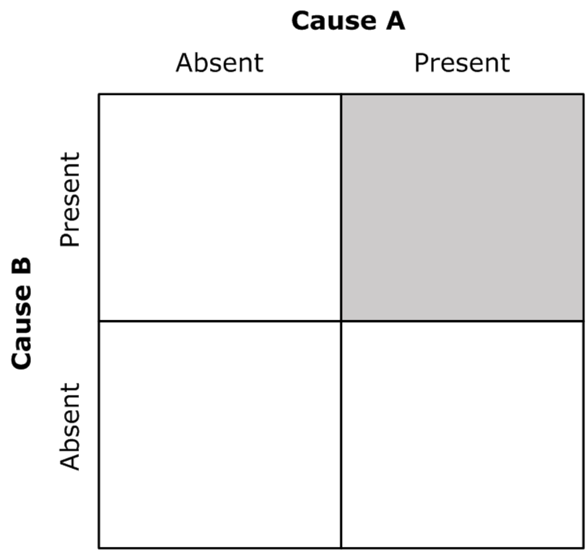


Figure 5.2: *Causal schema with two possible causes representing an attribution only if both causes are present. Effect highlighted in grey.*

in figure 5.3 presents three causes and, like before, represents an attribution only if all three causes are present. In practise, schemata will tend to be restricted in terms of dimensionality as they tend to be focused. Since each causal schema represents only one way of relating an effect to causes, it is possible that a person has and uses various schemata for the same problem. It is also not unreasonable to expect that these schemata may sometimes conflict or contradict each other.

5.2.3 Relation between Attribution Theory and Context Modelling

At first sight, the process of context modelling and the process of causal attribution may appear very different. Context modelling originated in computer and information science and is mainly concerned with the adaptation of information systems to users particularly in very dynamic environments such as the mobile application domain. Attribution theory, on the other hand, originates from social psychology and focuses mainly on understanding how people explain observed effects from their environment; a closer look reveals a number of interesting similarities worth further discussion:

- *Mental model:* A schema is a personal view about the connection between an effect with its possible causes. In other words, the causal schema serves as a mental model

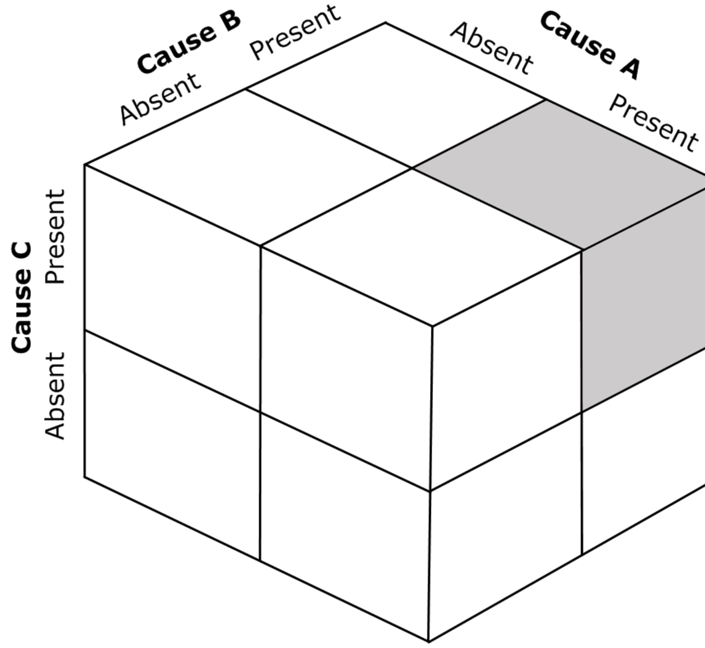


Figure 5.3: *Causal schema with three possible causes representing an attribution (highlighted in grey) only if all three causes are present.*

for context. Different schemata can be applied to explain one given effect in different ways and a single person might use them interchangeably. As such, a causal schema can be directly related to a context model. Viewed in this light, Kelley's cube therefore serves as one example of a context model with three attributes – *entities* (e.g. pictures in an art gallery[2]), *persons* (i.e. people, for example visitors of an art gallery) and *time* (e.g. different viewing times). These dimensions are equivalent to the attributes of a context model. In fact, a range of existing context models use one or more of these attributes in different degrees of granularity. The cube directly addresses time as an important contextual dimension. The personal aspect is viewed on a social dimension that represents an effect across different people. An entity refers to the matter of concern (e.g. the exhibit) that could also be related with one or more attributes about the situation around the paining (e.g. temperature in the exhibition room or number of visitors) or could refer to different places (e.g. exhibition rooms in the museum or across museums). Furthermore, the entire cube is viewed in relation to one particular effect (e.g. the person's perception of joy

[2]This example has been used in [Kelley, 1973] that was adapted from [McArthur, 1972]

about art).

- *Focus:* Most context models usually operate on a focused set of attributes. They are generally constructed based on a small arrangement of factors relevant for the information problem at hand and obtained from past empirical evidence from similar situations. This is equally represented in the schema, that focuses on a small set of possible causes. One major ambition in context modelling is to identify relevant contextual factors equally to the human explanation process that constantly seeks plausible and good causes to explain effects in our environment.

- *Internal and external components:* Both attribution theory and context modelling distinguish between internal and external factors. In attribution theory, internal attributions refer to causes within the person whereas external attributions mean causes in the person's environment[3]. Similarly, context modelling relates internal context attributes to the user model (as discussed in section 2.2.1) and external attributes to the environment of the user (as discussed in section 2.2.2). This demonstrates a structural connection between the process of human explanation and the process of context modelling with respect to its entities.

- *Causal pattern:* Attribution research, such as [Cunningham and Kelley, 1975, Kun and Weiner, 1973], discovered patterns in human explanation about the connection between different causes and effects. Normal phenomena often create a patterns similar to a logical 'OR'[4] that requires only one of many causes to be present. Exceptional phenomena however often appear as patterns similar to a logical 'AND'[5] that requires all causes to co-occur for the effect to happen (see figure 5.2 and 5.3). Questions about such patterns and associated investigations are also the theme in context modelling to further understand the relationship between context attributes (cause) and their effect on people.

Overall, causal attribution provides a theory that is not necessarily restricted to social psychology but might also prove helpful as an underlying theory for context modelling.

[3] Since attribution theory originates from social psychology, environment here usually refers to the social environment of the person.

[4] Also called a 'Multiple Sufficient Cause Schema'

[5] Also called a 'Multiple Necessary Cause Schema'

Kelley's view on attribution is particularly compelling due to its applicable framework and tools. The human acting as a naive scientist, reasoning about their environment by covariation over a set of causes and simplifying this process through causal templates is a plausible mental model. Another example of an alternative theory for human explanation is abductive inference, as defined in [Wirth, 1998], that has close similarities to attribution theory. Unlike attribution theory, its origin is based in logic with applications in artificial intelligence. Abductive inferences is one way of describing the process of finding possible explanations for an effect. The process is specifically focused on selecting the 'best' explanation from a number of possible explanations. It is not the scope of this thesis to investigate abductive inference with respect to context, however, the theory has been comprehensively examined for its relation to relevance feedback [Ruthven, 2001]. The rest of this chapter continues to focus on attribution theory as one possible way of viewing context modelling as a human explanation process.

As such, causal attribution also reflects the ideas of a context model that defines a set of attributes relevant for a user group, models attribute interrelations and expresses their importance for an application area.

When viewing the context model as a causal schema, it can be represented as a 3-dimensional cube, similar to Kelley's, along the chosen dimensions of time, location and interest (see figure 5.4). As already indicated, Kelley's assumption is that humans behave like naive scientists. He assumes further that people employ a simplified form of the scientific method of covariation to test for relations between an observed effect and potential causes. For that reason, ANOVA, the formal equivalent to the covariation principle, has been used as the statistical method for exploring and testing of attribution data. Examples include the pioneer work on attribution reported in [McArthur, 1972] and the study on the perception of unemployment conducted in [Hesketh, 1984]. This is also in line with the statistical analysis that is reported in chapter 4[6].

Based on this, a causal schema can therefore be interpreted as a simple form of multiple regression analysis that predicts a causal relationship based on assumed

[6] Also published in [Bierig and Göker, 2006]

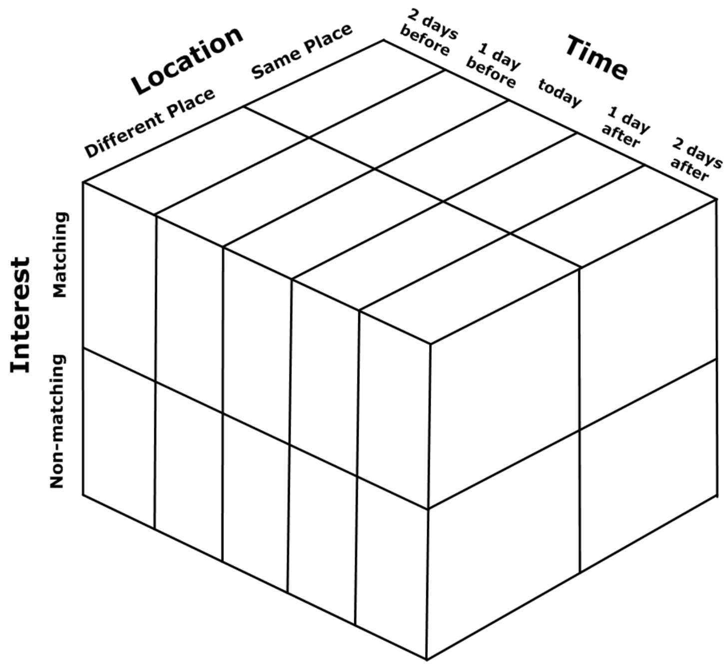


Figure 5.4: *Causal schema for the arrangement for the causes of time (5 levels), location (2 levels) and interest (2 levels)*

correlations between variables [Surber, 1981]. Surber's paper examined the effects of exam difficulty on the prediction of grades based on attributions of effort and ability. The paper reformulates a causal schema as a regression equation, a statistical method that is based on the same principles as ANOVA and widely used for prediction. The next section provides more information about the method and develops a model based on regression along the data of the user study that has been reported in the previous chapter.

5.3 Multiple Regression for Context Modelling

Based on [Kelley, 1973], the human process of attribution is linked with the statistical method of ANOVA. This is based on his assumption that people act as naive scientists and use basic statistical models to find explanations. Surber's paper [Surber, 1981] expanded on Kelley by further indicating that the explanation process is facilitated through building a simple form of a multiple regression model. In this chapter, we extend on Surber's paper by applying multiple regression to create a model of context based on the data from the study in chapter 4.

5.3.1 Overview to Regression

Multiple regression was originally developed in the area of behavioral sciences around 1900. In [Cohen and Cohen, 1975], multiple regression analysis is defined as "a highly general and therefore very flexible data-analytic system that may be used whenever a quantitative variable (the dependent variable) is to be studied as a function of, or in relationship to, any factors of interest (expressed as independent variables)". In other words, multiple regression defines a predictive function which quantitatively describes the relationship between one (or more) independent variable(s) and a dependent variable. Regression models relationships between variables with any functional form, does not constrain variables and allows data to be modelled in a holistic way thus also including their interactive behaviour. Based on these features, it appears as a very suitable and flexible method for the purpose of multi-variable context modelling.

The data from the user experiment that has previously been tested and analysed with ANOVA, is now regressed to estimate a more precise model as a functional description. This regression allows to predict usefulness of event content based on different levels of time, location and interest. The final form of the regression model is presented in formula 5.1.

$$Y = f(x) = \begin{cases} e^{1.564} e^{0.217T - 0.106L - 0.885I - 0.147TI} & -2 <= T <= 0 \\ e^{1.460} e^{-0.628T - 0.114L - 0.807I + 0.362TI + 0.088TLI} & 0 < T <= 2 \end{cases} \tag{5.1}$$

The remainder of this section justifies the regression in more detail. The standard form of multiple regression is a linear equation of the kind

$$Y = B_0 + B_1 X_1 + B_2 X_2 + ... + B_n X_n + \varepsilon \tag{5.2}$$

where B_k are constant weights for the predictors X_k, Y represents the prediction, and ε the error of the model. Formula 5.2 models not only linear relationships that already exist in this form but also all those that can be transformed into the shape of formula 5.2. This is done by first determining the functional form that best describes the data and then transforming the data by the inverse of that functional form. Many different functions may be suitable to describe the data and no certain, straightforward process exists that determines its shape, consequently, the search for the 'best' functional description is a challenging task. Generally, this is performed with support from theories that provide a hypothetical explanation for the data. With no theories at hand, the only other way is to manually investigate the data and determine the function empirically.

The data from the previous user study has a number of interesting aspects that empirically hints at its functional form. The most obvious effect is that the amount of usefulness is highest when the situation matches with the content information. Furthermore, usefulness strongly increases when the current time approximates to event performance time. After the event, usefulness declines strongly at first and to a lesser extent later on. This behaviour exists between all levels of location and interest difference. Although the effect is much more pronounced when the situation matches the content, the basic shape is generally consistent in all arrangements (see figure 5.5). The shape of

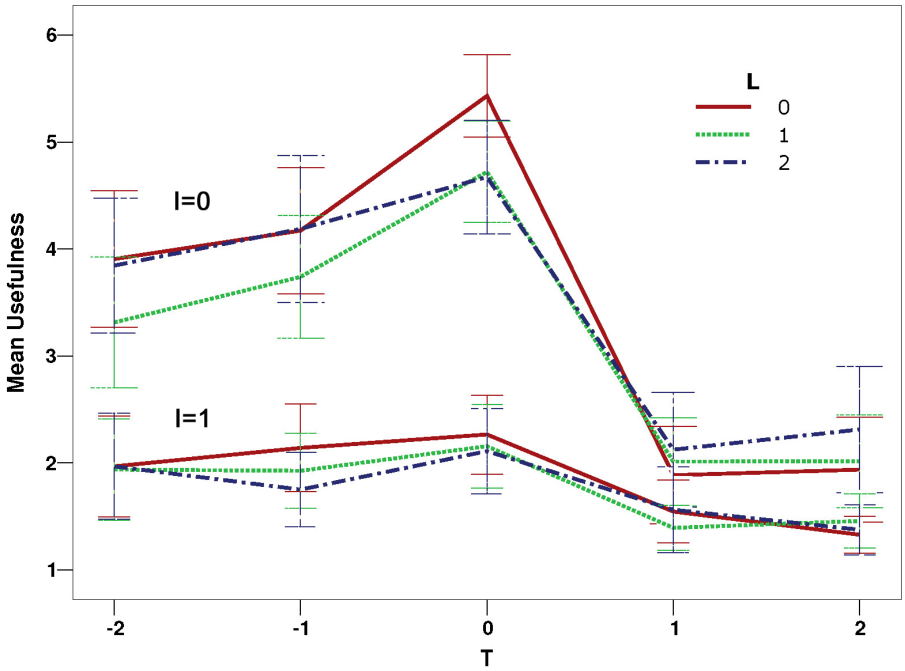


Figure 5.5: *Mean usefulness for matching interest (I=0) and non-matching Interest (I=1) at three levels of location difference (L) and 5 levels of time difference (T).*

the data suggests an exponential trend in the data. Before the event performance time, there is a strong rise, after the event there is a sharp decline. This indicates the presence of a power law relationship between at least one of the three contextual attributes and the perceived usefulness. This suggests that the shape of the data does most likely follow a power function of the form

$$Y = Ae^{BX} \tag{5.3}$$

what is considered as the standard form of an exponential function. Euler's constant e serves as the base and the two constants A and B determine the intensity and shape of the curve. When substituting A with $e^{Constant}$ and expressing BX as a linear combination of amounts of predictors (implicitly expressed as a sum), the formula becomes

$$Y = e^{Constant} e^{\sum_{i=1}^{n} B_k X_k} \tag{5.4}$$

A regression model that considers all effects and all possible interactions therefore contains 7 possible different predictors – the three main effects for Time (T), Location (L), Interest (I), as well as all its interactions (Time x Location (TL), Time x Interest (TI), Location x Interest (LI) and Time x Location x Interest (TLI).

The model places all attributes in the exponent. It therefore assumes some degree of basic power law behaviour for all main contextual factors and all their interactions. This is done since the model should represent the dominant trends in the data. Such a dominant trend was contributed by the time attribute (besides interest) and its power law behaviour. Furthermore, the model should also be based on the strength of attribute interactions. Although the interest attribute is stronger than the time attribute, it strongly interacts with time and is also the strongest overall interaction; even more effective than location. This suggests the time attribute as the major factor of the model and its power law effect as a potential major trend of the model.

One intuitive way to describe the rising and declining shape of the data is through a discontinuous split function. With this, the two parts of the split function model the two aspects in the data. One part describes the rising trend of usefulness before the event actually performs. The other part of the split function defines the declining trend after the event has performed and becomes less useful. Formula 5.5 shows its explicit form[7].

$$Y = \begin{cases} e^{Constant_i} e^{B_{1_i}T + B_{2_i}L + B_{3_i}I + B_{12_i}TL + B_{23_i}LI + B_{13_i}TI + B_{123_i}TLI} & -2 <= T <= 0 \\ e^{Constant_j} e^{B_{1_j}T + B_{2_j}L + B_{3_j}I + B_{12_j}TL + B_{23_j}LI + B_{13_j}TI + B_{123_j}TLI} & 0 < T <= 2 \end{cases} \quad (5.5)$$

The standard form for regression is obtained when transforming this formula with the natural logarithm – the inverse of the exponential.

$$ln(Y) = \begin{cases} Constant_i + B_{1_i}T + B_{2_i}L + B_{3_i}I + B_{12_i}TL + B_{23_i}LI + B_{13_i}TI + B_{123_i}TLI & -2 <= T <= 0 \\ Constant_j + B_{1_j}T + B_{2_j}L + B_{3_j}I + B_{12_j}TL + B_{23_j}LI + B_{13_j}TI + B_{123_j}TLI & 0 < T <= 2 \end{cases} \quad (5.6)$$

5.3.2 Results and Discussion of the Regression Model

The data points from the previous user study were split into two groups, each group of data was used to model one part of the split function.

- The first part describes the time before the event performance, is statistically highly significant (p<0.001) and explains 35.2% of the variation in the data (R^2=.352)[8].
- The second part describes the time after the event, also shows high levels of statistical significance with p<.001 and explains 36.2% of the variation in the data (R^2=.362).

Through the process a number of non-significant predictors were removed from the model as they did not contribute to its accuracy. Based on the remaining significant contextual predictors (e.g. p<.001 for interest), the regression model is

$$Y = f(x) = \begin{cases} e^{1.564} e^{0.217T - 0.106L - 0.885I - 0.147TI} & -2 <= T <= 0 \\ e^{1.460} e^{-0.628T - 0.114L - 0.807I + 0.362TI + 0.088TLI} & 0 < T <= 2 \end{cases} \quad (5.7)$$

The coefficient of each contextual predictor expresses the trend of the attribute and its quantitative strength. The positive coefficient for the time predictor (T) in the first part of the model indicates a rising trend of usefulness towards performance time. Similarly, the negative coefficient for time (T) after the event indicates a declining trend. The coefficient for the location predictor (L) shows that increasing location difference lowers usefulness almost equally in both parts of the model. Similarly, the negative coefficient for the interest predictor (I) signals that mismatching interest lowers the degree of usefulness.

[7]This is the complete form of the model. In the next section, it will be demonstrated that some predictors are not required because they do not contribute to the accuracy of the model.

[8]The intuition of R^2 is, the larger the magnitude of R^2 the better the model explains the data.

The interaction between time and interest causes a small declining correction in the first part of the model, and a slightly larger positive correction in the second part. The three way interaction also contributes a very small positive correction for the second part of the model. The graph in figure 5.6 shows the regression curves along different values

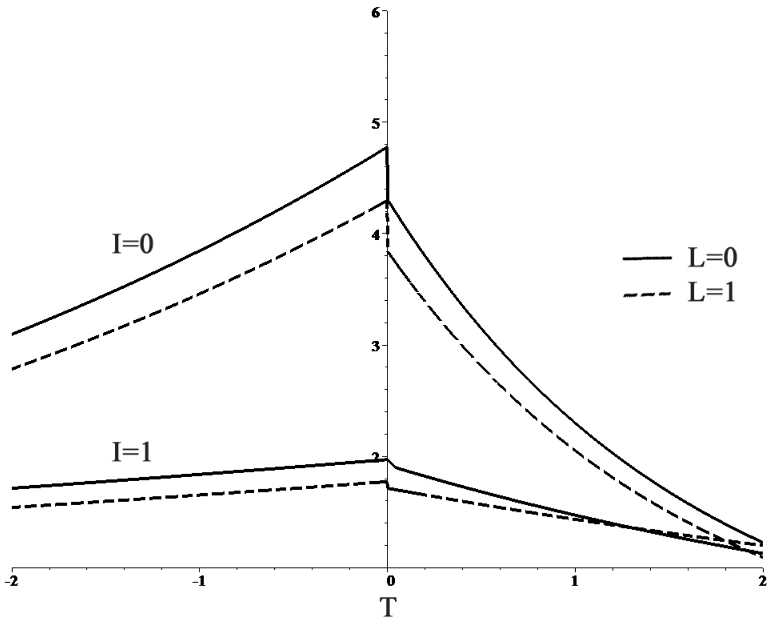


Figure 5.6: *Combined regression model that predicts usefulness based on time(T), location(L) and interest(I)*

for time, location and interest. The model represents 5 different levels for time (T={-2, -1, 0, 1, 2}), 2 levels of location (L={0,1}) and 2 levels of interest (I={0,1}). T <= 0 represents the time before and concurrent with the event, whereas T > 0 represents the time in days after the event. The location of the situation is either matching (L=0) or not matching (L=1) with the event location. Interest is also either matching (I=0) or not matching (I=1) with the event category.

As shown in figure 5.7, this split regression function can now be mapped to the causal schema as discussed earlier. The regression adds the precise quantities of the predicted usefulness of an event (the effect) for a user in a particular situation composed of a time, a location and an interest (the causes). The figure shows the regression only as a projection to the front of the cube due to the limitations of 2-dimensional visualisation.

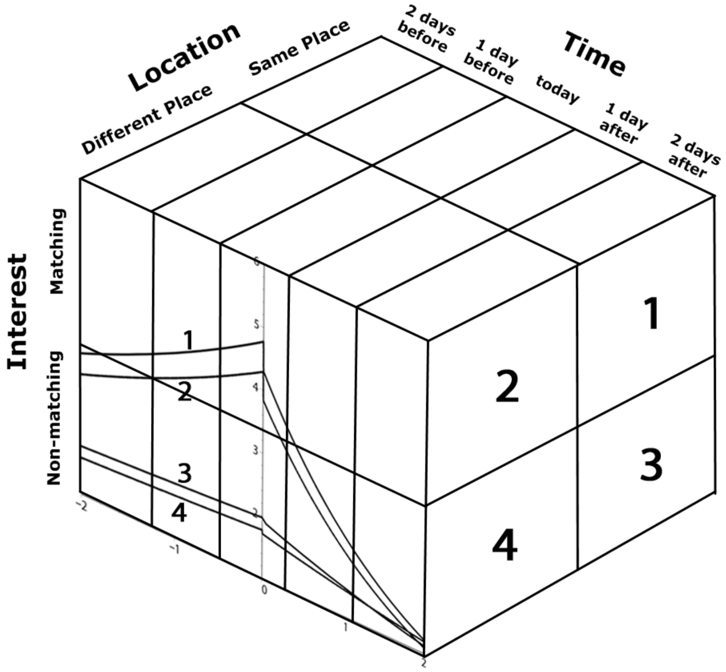


Figure 5.7: *The context model mapped to a causal schema*

Ideally, each of the 4 regression lines[9] would be in one of the four quarters of the cube as enumerated on the right side.

This opens the discussion of how causal schemata are structured in human perception. As suggested in [Kelley, 1973] and discussed in section 5.2, people might use a range of simple schemata for one kind of problem (e.g. determining if content is useful). In other words, simple schemata might be used interchangeably for a single problem; figure 5.7 only shows the most comprehensive form. This representation assumes that people have a single complete schema that describes the relation among three context attributes and the usefulness for event content, however, there are other, equally possible ways of how this single schema might be divided into smaller and simpler schemata. Figure 5.8 shows three simple possibilities of how people might differentiate the relation between the attributes of the context model and usefulness into different causal schemata – one distinguishes based on location (a), one on interest (b) and one on time (c). From these three possibilities, the third one is the most likely case based on the two opposed

[9]The individual lines in the regression function of figure 5.7 are enumerated across the split

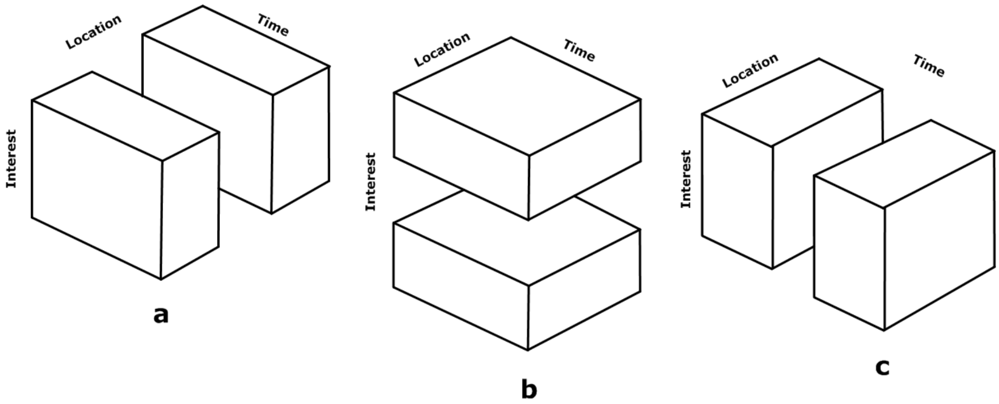


Figure 5.8: *Different possibilities of how people might group context into causal schemata*

power-law trends in the data. As mentioned earlier, before the event, usefulness rises first moderately and later strongly up to the event performance time; afterwards, it declines first strongly and lesser later on. This indicates that different causal templates may be responsible for the explanation process – one that describes the attribution up to the event performance and one that explains the relation afterwards. At this point, however, the question of how many schemata co-exist is left to future studies. The remainder of this chapter continues to investigate the causal schema further and then describes its integration with information retrieval.

The quality of a regression model is generally assessed by a number of assumptions as discussed in [Schroeder et al., 1986] and [Field, 2005]. These assumptions are mostly based on the errors (also called residuals) that the model produces based on the same data set. A highly reliable model should have unrelated errors (low autocorrelation). These errors should be constant in their variation (homoscedasticity), fit the linear model (linearity) and be normally distributed (normality), furthermore, the contextual predictors should produce differentiable effects (low multicollinearity) on usefulness. Tests[10] showed that the regression model does indeed provide low autocorrelation, an overall low multicollinearity and homoscedasticity; however, the assumption of linearity

[10] Autocorrelation was evaluated based on the collinearity diagnostics produced by SPSS. Homoscedasticity and linearity were evaluated using plots of standardised errors against the standardised predicted values of the model. Normality was tested with a Kolmogorov-Smirnow normality test. Multicollinearity was tested using collinearity statistics (using the tolerance/VIF measure) as provided by SPSS. All mentioned methods are a standard practice and described in [Field, 2005, p.202-206,258-263]. All evaluations and tests were performed on the errors (also called residuals) of the regression model.

and normality do not hold.

This indicates that the model does not explain all relevant effects that play a role in the way users judge usefulness with respect to the provided context. It is likely that more hidden attributes exist but are not part of the investigation and not included in the regression model. Nevertheless, this first evaluation based on the data set the model is built from allows a first look into the relationship between the data that was measured and the prediction. Despite the necessity of this evaluation, a far better method to evaluate the quality of a regression model is through testing with new data as suggested by Cohen [Cohen and Cohen, 1975]. The next chapter performs such an evaluation where the model is applied as part of a personalised information system and used in a mobile application environment. This allows for an overall evaluation in a more realistic scenario based on new data. The next section prepares this evaluation by creating a personalisation model that combines the regression with a common information retrieval model for personalised search.

5.4 A Personalisation Model for Situational IR

5.4.1 A Brief Review on Information Filtering and Retrieval

The previously described regression model describes a functional relationship between context and usefulness as a refinement of a causal schema. Such a model can be used for different kinds of content personalisation. Recalling from section 2.4 on page 35 in the related work chapter, there are two basic kinds of content-based personalisation techniques covered by the literature – information retrieval and information filtering. Although the context model could be applied for both approaches, it is preferred to use information retrieval for two reasons:

- *Including the participating user:* The user is the main focus of any information system. Rather than a pure focus on the effectiveness of the context model, it is preferable to investigate its effectiveness in line with the user in an interactive process. An information retrieval system allows to create a link between system and user. By doing so, it turns the user into an information participant rather than an information consumer.

- *Including the content:* Filtering based on a context model would only allow to filter the event content by its contextual match with the user's current situation. The text content (i.e. term statistics) would not be directly involved like in Letizia [Lieberman, 1995] and Fab [Balabanovic and Shoham, 1997]. Using an information retrieval system therefore allows to include the content as an important element of a system.

Whereas both approaches allow investigation into the value of context, information retrieval is preferred as it includes the user of the system in an active manner as well as the content with which the system is dealing.

5.4.2 A Combined Score of Information Retrieval and Context

Recalling from the previous sections, the regression function represents a causal schema of how people explain the usefulness of event content based on time, location and interest. More precisely, the model describes the amount of situational relevance based on the degree of match between the user's current situation and the contextual information contained in the event content. Figure 5.8 again shows the context score as it has been introduced and justified in section 5.3.

$$Score_{ContextModel} = \begin{cases} e^{1.564} e^{0.217T - 0.106L - 0.885I - 0.147TI} & -2 <= T <= 0 \\ e^{1.460} e^{-0.628T - 0.114L - 0.807I + 0.362TI + 0.088TLI} & 0 < T <= 2 \end{cases} \quad (5.8)$$

The aim of this thesis is to apply this score to extend standard information retrieval into personalised information retrieval. For that, the regression ($Score_{ContextModel}$) is combined with a score that represents the content-based relevance by information retrieval ($Score_{IRModel}$). This content-based score is determined by the degree of match between the user's query terms and the terms contained in the event information. For that, an IR algorithm[11] is applied as described in section 2.4.

To enable personalised information retrieval, both scores are combined into a single

[11] The Lucene IR library was employed for the computation of the IR score. Its scoring formula is described more comprehensively in section 2.4. The scoring formula is presented in more detail in appendix D.

score. In other words, both types of relevance are merged into a single personalisation model that combines information retrieval and context. The personalisation model is represented by the following formula.

$$Score_{PersonalisationModel} = \alpha Score_{ContextModel} + \beta Score_{IRModel} \tag{5.9}$$

The two scores are combined by addition. Unlike multiplication, addition is a common way of combining elements that are regarded as independent. The content score is assumed to be independent from the context score. Indeed, both scores originate and operate on different types of data and different methods are used to generate them. The use of multiplication would combine the two scores on a logarithmic scale thus strongly tying them together. This is not desired based on their distinct origin and nature, and therefore addition appears as a more accurate way for combining the two. The basic additive combination between the two scores is enriched by an extra weight for each of the two parts of the model. The two constants α and β determine the weight for each score thus allowing to shift emphasis between content-based and situation-based relevance. Similar strategies have for example been applied in IR evaluation. The F measure [van Rijsbergen, 1979] combines two different evaluation measures; precision and recall. The E measure, its generalised form, allows to parameterise between these two evaluation measures, likewise, the above formula provides a basic parameterised method for personalised information retrieval with two different scores.

In this thesis, both constants have been balanced thus giving equal weights to both forms of relevance as shown in formula 5.10. This personalisation model is applied in the mobile user study that is described in the next chapter.

$$Score_{PersonalisationModel} = 1.0 Score_{ContextModel} + 1.0 Score_{IRModel} \tag{5.10}$$

Figure 5.9 shows an example[12] of the scoring of two events using the personalisation model as defined with formula 5.10. As shown in the formula, each of the two events are scored based on two individual scores. One score is based on the closeness of the

[12] This example has been obtained from the data of the mobile user experiment that is described in the next chapter that provides more details about the conduction of this experiment with results.

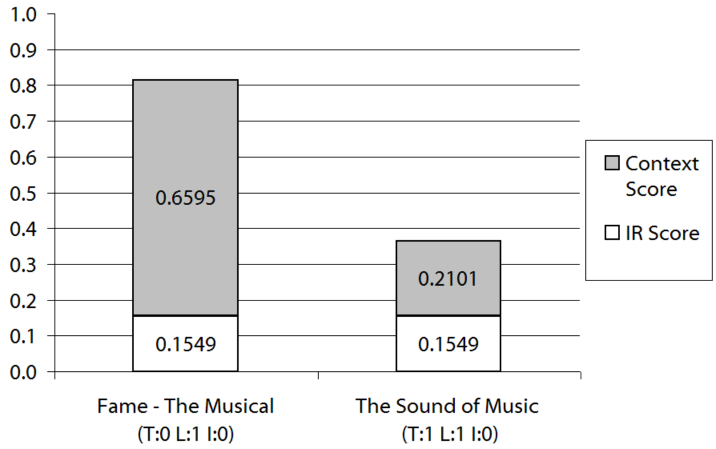


Figure 5.9: *Example of two scored event content items consisting of one IR score (lower part) and one context score (upper part). Both IR scores are equal and would create a weight block. The context score resolves this weight block.*

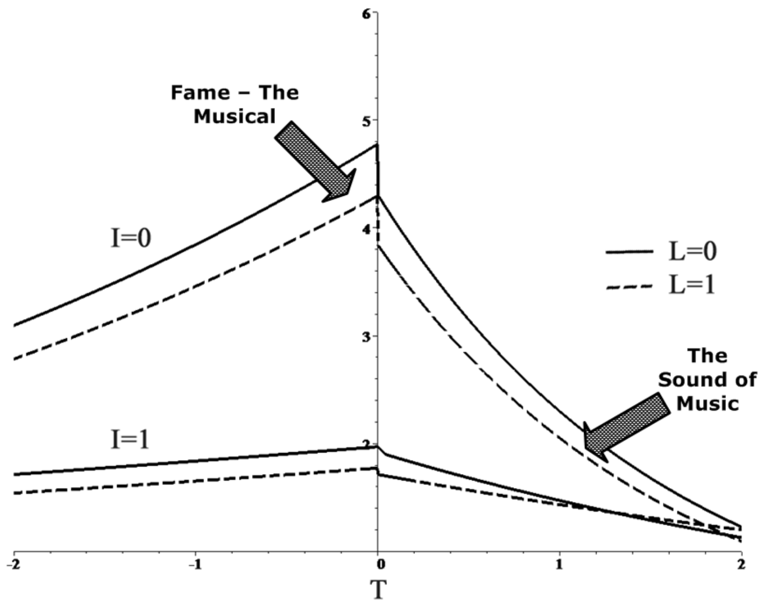


Figure 5.10: *Example of two scored event content items pinpointed in the regression model*

search query to the event text content (IR Score). The other score is determined by the regression model that is based on the closeness of the user's situation to contextual information contained in the event content. The regression represents a causal schema between three context attributes (possible causes) and the user's perception of usefulness (the effect) in the form of a more fine-grained and more detailed, functional relationship.

Figure 5.10 pinpoints the two events in the regression function. Both content items have an equal IR score although both are, in fact, different. This condition is generally referred to as a weight block, as, for example, described in [Göker, 1994]. Weight blocks are search results with equal relevance scores but different content. A ranked list with a weight block contains two or more content items with equal scores. Equally scored results mean that these should also occupy an equal position in the ranked list. As this is not possible, retrieval systems 'rank' those items even though the score would suggest otherwise. In practise, items end up in arbitrary rank positions dependent on how the search engine works, forming a weight block; however, the user perceives different degrees of relevance dependent on these arbitrary rank positions. From a statistical viewpoint, weight blocks are a lack of variance caused by the scoring function. Context can help with this problem since it introduces a second source of variance in the form of another score. This score is based on different data and generated by a different method which adds additional variance from another perspective – the perspective of the contextual situation. As the personalisation model combines both types of relevance it reduces the likelihood of weight blocks[13].

A balanced weight between the two scores might not be optimal, however, equal weights are plausible for initial investigations since the effect of a changed weight balance is difficult to predict. Nevertheless, the personalisation model presented with formula 5.9 allows for adaptations based on specific demands. A system that requires more emphasis on the content and less on the context would use a larger β and a smaller α (e.g. β=1.0 and α=0.5). Likewise, a personalised system that aims to focus more on the contextual side would increase α and reduce β instead. Figure 5.11 shows an example of how different weights for context and IR scores affect the total scoring of two events for an example query ("alice")[14] and therefore their rank. The example represents the results of a query with 3 hypothetical personalised information systems that make use of a personalisation model as described in formula 5.9. System A puts more emphasis on the content hence weighting the IR score twice as strong as the context score (IR:1.0, Context:0.5). System

[13] Although the likelihood of weight blocks has been significantly reduced in the implementation of modern information retrieval algorithms (e.g. the BM25 algorithm), it is worth noting that a weight block of only two results may be enough to occupy half of the screen of a mobile device.

[14] The figure shows non-normalised scores. However, the research prototype normalises all scores to a maximum of 1.0. This normalisation does not affect the rank.

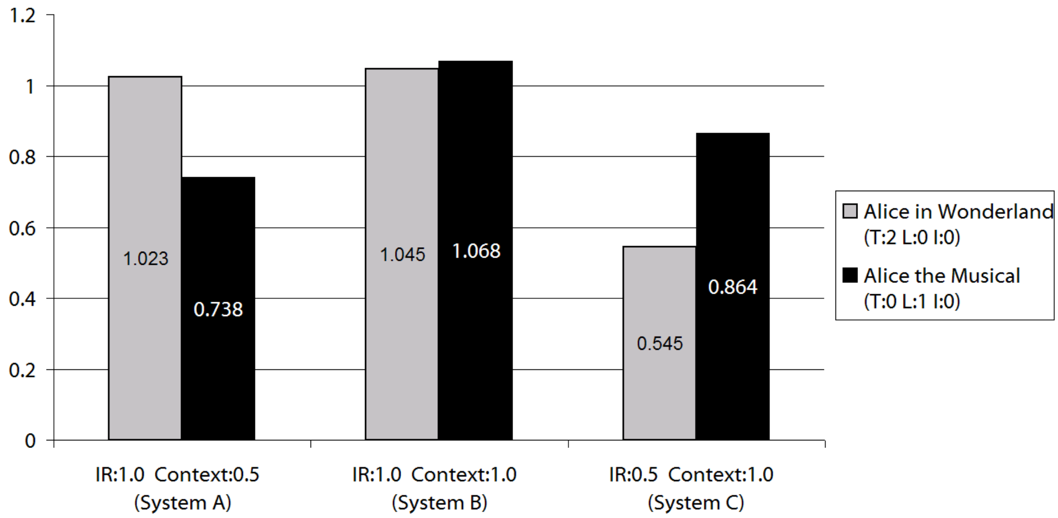


Figure 5.11: *Search result for example query "alice". Two event content items scored with different weights for context (α) and IR (β).*

B represents formula 5.10 and maintains a balanced relation between the two scores (IR:1.0, Context:1.0). System C weights context twice as much as the IR score (IR:0.5, Context:1.0). The result list consists of two events – *Alice in Wonderland* (event 1) and *Alice the Musical* (event 2). Despite the difference in content title and description, the performance time of event 1 has passed by two days (T=2), matches the user's current location (L=0) as well as the user's interest (I=0). However, event 2 performs on the same day (T=0), at a venue different from the user's current location (L=1) and also matches the user's interest in musicals (I=0). Recalling the context regression model it is evident that event 2 is contextually more relevant, however system A still ranks event 1 higher. This happens despite the fact that context information is considered on a moderate scale. System B, including equal content and context scores, changes the rank between the two events; an effect that is further strengthened with system C. This example shows the effect of context as an adjustment for a system that operates mainly content-based.

An alternative to a personalisation model of the form of formula 5.9, a personalised information system could also extend a conventional information retrieval system with additional context as part of the query. This would touch a number of issues some of which have been addressed previously in this thesis:

- **Contextual knowledge:** The main issue when integrating context with queries is the problem of implementing contextual knowledge with the means of information retrieval. One possible way would be to extend the retrieval process with extra Boolean rules that operate on the result list. These rules could filter or promote results based on their contextual information. An example implementation could for example filter out past events, group results based on their closeness to the user and put events on top that match the users interest. However, this ad-hoc style would solely be based on intuition. When dealing with several attributes at once, it would also become increasingly difficult to coordinate the individual effects of attributes into an overall consistent effect. Interactions between the different contextual factors would also need to be addressed with equal care. The study in chapter 4 clearly showed that significant interactions between the attributes occur and are part of the context model; any ad-hoc, fixed rule based system would lose part of its quality if attribute interactions were not considered and integrated.

- **Query construction:** The query construction would be with the user who would need to incorporate time, location and interest information into the query. For simplification, this part of the query could be controlled through extra user interface elements. Many advanced searches of popular search engines [15] and digital libraries[16] offer this feature, however, its application remains the burden of the user. Based on experience with search engines, it is known that users hardly ever use advanced searches [Jansen et al., 2000]. Nielsen reports similar observations pointing to on a recent study on search behaviour [Nielsen, 2005]. Additionally, the limits of mobile usage would play an additional role as discussed previously in section 4.2.1. The screen of most mobile device restricts the application of such extra user interface components. Research has demonstrated that the attention span of the mobile user tends to be small and difficult to maintain [Tamminen et al., 2004, Oulasvirta et al., 2005]. Complex query construction on a

[15] Such as `http://www.google.com`, `http://www.altavista.com` and `http://www.ask.com`, all accessed April 14, 2008

[16] Such as `http://www.sciencedirect.com`, `http://www.emeraldinsight.com` and `http://portal.acm.org`, all accessed April 14, 2008

mobile device is therefore a challenging task.

- **Feedback loop:** Besides the need for users to construct the contextual query, they would also need to evaluate the effects resulting from the extra contextual information. This would be independent from the contextual query being constructed through a user interface or with an advanced query language. It would not be surprising if such a solution would turn out to be generally weak in terms of this feedback loop. Many parameters would rest in the hands of the user to be adjusted and tweaked in order to solve an information need, however, in reality users are generally unwilling to repeat and refine queries as described in [Nielsen, 2005]. This becomes even more critical in a mobile usage environment, as stated in the previous point, where users tend to be more limited in their cognitive resources [Tamminen et al., 2004, Oulasvirta et al., 2005].

5.5 Summary

In this chapter, results from the previous experiment on contextual usefulness were applied to create a predictive model of context. The model was connected to some of the theory that describes the human process of explanation finding, also called attribution theory; an area of psychology that investigates how people relate effects to their potential (contextual) causes. The theory was highlighted based on Kelley's covariation principle that links human explanation with factorial ANOVA and causal schemata that connect attribution with a basic form of regression modelling. Regression was applied to develop a predictive context model where a score expresses the amount of usefulness (situational relevance) based on time, location and interest. Within a personalisation model, this context score was combined with a traditional information retrieval score. The next chapter reports on a comprehensive field experiment that tests the personalisation model in a realistic mobile scenario. Results from the validation of the personalisation model are presented and discussed in more detail.

6

Personalising Events with Context: The Field Evaluation of a Model

> Therefore, having obtained the opportunity from these sources, I too began to consider the mobility of the earth.
>
> NICOLAUS COPERNICUS

6.1 Introduction

In the previous chapter, results from the laboratory user study were applied to create a predictive model of context. The context model was connected with the natural human process of explanation finding; an area of psychology called attribution theory that investigates how people relate effects to their potential (contextual) causes. The theory was highlighted based on Kelley's covariation principle, that links to human explanation with factorial ANOVA, and causal schemata that connect attribution with a basic form of regression modelling. Regression was then applied to develop a predictive context model where a score expresses the amount of usefulness (situational relevance) based on time, location and interest. Within a personalisation model, this context score was combined with the score of an information retrieval system.

This chapter reports on an experiment that applies this personalisation model to evaluate its overall effect on users in a more realistic mobile usage environment. This is especially valuable, as some of the formal assumptions – linearity and normality – of the context model were not fully met as described in section 5.3.2 on page 113. According to [Cohen and Cohen, 1975], the best way to handle this issue is to obtain a new data set that tests the model from a different viewpoint – thus validating the previous findings. The evaluation takes place as a mobile field experiment with a mobile information retrieval system that embeds the personalisation model. A different and larger event content subset is selected from the Reuters Kalends collection[1]. The content described events from topical categories different from those applied in chapter 4. In this evaluation, the performance of the context model is measured holistically with respect to three performance measures – rated usefulness as a measure of content quality, task time and the amount of submitted queries as a measure of users' search effort. This allows for an evaluation from a different perspective on the effect of context on mobile users.

Besides evaluating the the personalisation model, this chapter also investigates the mobile search behaviour of participants. A closer look is taken at the amount of queries and query terms people submitted during the experiment. The query formulation process is investigated with respect to the amount of context stimulus that was included. Also, this chapter reviews how participants rated their retrieved event content with the search system.

This chapter is structured in 6 further sections. Section 6.2 discusses relevance and value of a mobile field experiment with respect to its ability to evaluate contextual personalisation. Section 6.3 describes the design and method of the experiment. Section 6.4 describes the participants recruited for the study. Section 6.5 contains the details of the experiment procedure. Section 6.6 presents the results – usefulness and search effort as well as participants' more general search behaviour – and in section 6.7 the results are discussed.

[1] Appendix E provides a comprehensive overview of the Reuters content collection.

6.2 A Mobile Environment for the Evaluation of Contextual Effects

Evaluation of information systems and their underlying theoretical concepts have started with a tradition in laboratory experiments[Saracevic, 1995]. With relatively little effort, users can be placed into similar conditions, trained jointly and equally on a particular information system, receive similar tasks and perform them in an equal and (relatively) unchanging environment. However, there have been significant changes over the last decade due to the arrival of powerful mobile computing equipment. Today, people use an entire range of different, small computers that coexist with the personal computer: notebooks, tablet computers, PDAs and mobile phones. According to the NetSize 2006 Guide[2], the mobile industry aims to transform the mobile phone as we know it today into a portable and truly personal minicomputer. This is evident in the growing abilities of mobiles to handle secondary software and rich multimedia supported by constantly increasing performance in memory, processing power and communication abilities (such as Bluetooth and Wireless LAN). This development has several effects. Firstly, it has added new, more integrated, types of usage for computing equipment into peoples' lives. Secondly, it has changed the requirements for evaluation.

The following two sections focus on the above two points. Section 6.2.1 revisits the importance of information system usage with respect to evaluation. This is followed by section 6.2.2 that discusses the current call to enhance the use of classic laboratory experiments with field experiments as an alternative that more closely matches with the new requirements of mobile computing. Some of the issues are highlighted that need to be considered when evaluating context and the choice of experiment methodology is clarified. Then, evaluation measures and hypothesis are stated before continuing with the experiment design and method in section 6.3.

[2] The NetSize Guide is an annual report that provides in-depth analysis based on statistics from over 100 mobile operators operating in over 30 countries and covering over one billion mobile phone users. It is available from `http://www.netsize.com`, accessed April 14, 2008.

6.2.1 Importance of Usage for Context Evaluation

Chapter 4 stressed the importance of how an information system is used when establishing a context model. Now this issue is revisited in more detail with respect to the evaluation of contextual concepts.

During the last decade, a large amount of powerful handheld computing equipment became widely available; mainly PDA's and mobile phones able to assist users with daily life tasks in all kinds of diverse situations. These situations can be categorised into three groups based on the amount of mobility they support:

1. In *stationary usage* information systems are applied in a particular place only; usually on a desktop computer or otherwise stationary computing device (e.g. an information system on the computer at the local library or an information point in a museum).
2. In *semi-mobile usage*, an information system is applied from a mobile device but mostly while the user is not busy moving (e.g. reading text messages on the phone or updating the PDA calender while waiting at the airport).
3. During *mobile usage*, users are simultaneously mobile and utilising mobile computing equipment (e.g. phoning, texting and searching for information while walking).

These three categories of usage are obviously not strictly separated. For example, somebody using his PDA during a train journey could view the situation as semi-mobile since the device is generally mobile but used while the person is at rest within the boundaries of the train carriage; the train is nevertheless constantly moving. A more precise categorisation between the three types of usage is based on the way location is modelled for an application. A scheduling application may model location based on a GPS signal to predict arrival times and inform the user about time constraints. Internet access provided in the train may only model users' location based on their relative location in the train (e.g. the carriage).

It is the device that generally determines whether it can be used stationary or also (semi-) mobile. Usability and cognition may also determine whether a system is applied semi-mobile or mobile. Most devices are cumbersome to use on the move: Tablet computers are difficult to use one-handed and PDA's often require the use of a stylus due to their limited display sizes. Even mobile phones, designed to be used on the move, often force their users to withdraw from mobile usage and go back to semi-mobile as this frees necessary cognitive capacities as described in [Oulasvirta et al., 2005].

The experiment in this chapter is conducted in a semi-mobile usage scenario based on the categorisation above. This is based on the widespread occurrence of this usage style; results might, nevertheless, also be applicable for the other usage categories mentioned.

6.2.2 Importance of User-Centred Evaluation for Context

According to [Saracevic, 1995], "Evaluation means assessing performance or value of a system, process (technique, procedure...), product, or policy" as "a critical necessity in science". Information science, mainly driven by information retrieval, has traditionally focused on laboratory evaluation. This tradition was originally created and shaped by the Cranfield experiments [Cleverdon and Keen, 1966] half a century ago. These experiments consist of a collection of documents (i.e. content), a set of queries and a set of relevance judgements. The Cranfield experiments set the standard for other major evaluation efforts such as SMART [Salton, 1971] and TREC [Voorhees and Harman, 2005]; purely system-centred forms of evaluation that apply precision and recall as performance measures. The focus on system-centred evaluation has recently been challenged as an approach too narrow. Extensions into more user-centred forms of evaluation have been proposed. In [Saracevic, 1995] it is argued that both forms of evaluation are equally required. New evaluation frameworks have been proposed such as in [Borlund and Ingwersen, 1997] that uses simulated work task situations for more realistic tests in information seeking. In [Reid, 2000] and [Vakkari, 2001], tasks were proposed for more user-centred forms of evaluation.

The emergence of mobile, context-aware and personalised information systems has pushed the demand for alternative forms of evaluations even further. Different usage styles allow for other forms of evaluation outside the laboratory setting. The field experiment in section 6.3 represents such an alternative form with respect to experimentation. A field experiment is a quantitative evaluation that is conducted in the natural environment of the effects under investigation. The field experiment is intuitive since mobile applications are created for these environments and people use these applications in mobile settings.

Over the last decade, several mobile systems have been developed and evaluated with field experiments – examples include the system for the CRUMPET EU-IST project [Schmidt-Belz et al., 2003], the AmbieSense EU-IST project [Myrhaug et al., 2004a, Göker and Myrhaug, 2007] and MOBILEWARD [Kjeldskov et al., 2005]. The field of mobile and ubiquitous computing also hosts specialised conferences such as Mobile HCI [Nieminen and Röykkee, 2006] that generally encourages the methodology of evaluations in the field. Nevertheless, field experiments still remain relatively sparse compared with laboratory studies. In [Kjeldskov and Graham, 2003], it is reported that evaluation in mobile human computer interaction is mostly performed in very basic and intuitive ways using trial and error and driven by requirements. The paper also highlights the general focus on laboratory studies in comparison to the sparsity of field studies. This lack of realism in the evaluation of information systems is also identified in [Scholtz, 2006]. The paper distinguishes between three forms of evaluation – laboratory evaluation, simulated evaluation and operational evaluation: [3].

1. *Laboratory evaluation* often uses the methodology of scientific experiment that makes it generally easy to control the variables under investigation, allows the experimenter to observe/measure the experiment effect and to repeat the experiment in (relatively) unchanged conditions in order to verify findings. A laboratory evaluation usually investigates isolated and more fine-grained aspects of a system, model or theory [Scholtz, 2006]. As a trade-off, it is possible that certain holistic properties of a system (e.g. usability) cannot easily be identified in an laboratory environment. An article published in 1992 on information retrieval evaluation by Robertson and Hancock-Beaulieu also highlights the conflict between laboratory

[3] This overview represents a selection of evaluation methodologies. It does not claim to covers all possible forms of evaluation nor addresses all variations that exist for of laboratory, simulated and operational evaluation.

evaluation and operational evaluation as a trade-off between control and realism [Robertson and Hancock-Beaulieu, 1992].

2. *Simulated evaluation* explores a middle ground between control and realism. It employs the method of experimentation that exercises certain amounts of experiment control but also introduces some degree of realism. The methodology presented in [Borlund, 2003b] uses simulated work task situations to contextualise peoples' information needs and provide a reference for users' relevance judgements. This allows for certain amounts of control within an otherwise realistic process of information seeking.

3. *Operational evaluation* usually makes experiment control unobtrusive or removes control from the setting by introducing the system into a real application environment. Real users apply the system in a real surrounding solving real problems. The advantage of this type of evaluation is a fully realistic data set. As a disadvantage, effects may not occur or may be difficult to observe or measure. Repeated evaluations might also produce very different results. This type of evaluation also tends to produce more coarse-grained results [Scholtz, 2006] in comparison to laboratory and simulated evaluations.

An article published in economics [Harrison and List, 2004] also discusses the differences between traditional laboratory experiments and field experiments; the field experiment being composed of all experiment settings that are not conducted in strictly controlled laboratory environments thus including the simulated evaluation and the operational evaluation mentioned above. Robertson and Hancock-Beaulieu's article distinguishes between laboratory and operational tests but was published before simulated work task approaches were developed. Harrison and List identify a range of relevant criteria for field experiments. Three of these criteria are of a general nature and therefore relevant for the approach taken in this thesis – *users*, *environment* and *task*[4] [Harrison and List, 2004, p. 1012]. It appears advantageous to discuss these three criteria in preparation for section 6.3 that describes the mobile experiment.

[4]Besides the mentioned three criteria, Harrison and List's article mentions the criteria of the commodity and the stakes that are limited to economics and therefore excluded from this discussion.

1. Real *users* are often difficult to recruit. Nevertheless, it is important to assess the appropriateness of the quality of a user sample and how well they represent the target population of the study. The application domain of this research is targeted for the general public who uses mobile devices (e.g. mobile/smart phones) and generally enjoys entertainment events as part of a modern lifestyle. Based on that, no specific sampling for users was applied but instead a range of members of university staff, mature students and professionals was selected in order to maintain a variety of people. This brings experience to the experiment task and introduces a wider range of demographic characteristics as described in [Harrison and List, 2004]. The sample was gender balanced (9 male, 8 female) with an emphasis toward the younger age (10 users between 18 and 29). Participants were very familiar with PCs, mobile phones, paper and electronic maps, search engines and generally the city centre of Aberdeen. They were less familiar with Personal Digital Assistants (PDA's). Furthermore, participants preferred searching events electronically rather than in the newspaper. More statistical details about the demographic structure of the sample are reported in section 6.4 on page 138.

2. A natural *environment* provides an extra degree of realism when collecting data. According to [Harrison and List, 2004, p. 1013], the "environment can provide context that suggests strategies and heuristics that a lab setting might not". Users may respond to controlled experiment stimuli, but also to implicit, contextual information originating from the environment. Experiments that include the natural environment require a more careful experiment design, may consume more resources and may provide only incomplete findings as stated in [Kjeldskov et al., 2004, p. 9]. However, it also contextualises findings and strengthens results with additional degrees of realism. Simulated evaluation performed as a field experiment can introduce necessary levels of control and should be preferred over fully realistic and unrestricted evaluations in the field.

3. Realistic *tasks* introduce an extra level of realism into a field experiment. This can be achieved with simulated work tasks [Borlund, 2000] that integrate realistic tasks under conditions of a real usage environment into a laboratory setting. This methodology was applied in this study and was additionally moved into the field

of a semi-mobile usage environment. This allows for a controlled experiment in an otherwise realistic setting.

Under consideration of the previous points, the experiment reported in this chapter is a mobile experiment. It is organised as a simulated evaluation that fulfills the requirements of a field experiment considering the previous three criteria of having real users, operating in a natural environment and using realistic tasks. The experiment is closely related to the methodology of simulated work tasks situations as described in [Borlund, 2000].

The effect of the personalisation model is measured based on the judged usefulness (understood as situational relevance as described in [Borlund, 2003a]) of event content, the time users need to solve a task and the number of submitted queries for solving a task. Whereas the measure of usefulness evaluates the content quality, task time and the amount of submitted queries investigate users' effort to fulfil their information needs. All three measures are expected to be affected by personalisation as expressed in the following three hypotheses:

- People find more useful event content with a system that provides personalised event search results based on context compared with a system that provides non-personalised results. (H_5).
- People solve search tasks faster with a system that provides personalised event search results based on context compared with a system that provides non-personalised results. (H_6).
- People solve search tasks with fewer queries with a system that provides contextually personalised event search results based on context compared with a system that provides non-personalised results. (H_7).

6.3 Experiment Design and Method

The study was conducted as a field experiment following a simulated evaluation based on the discussion from the previous section. Participants were brought to real locations and solved search tasks based on a simulated scenario with a mobile device provided to them.

Each participant also completed a short questionnaire before and after the search tasks.

The experiment is a repeated measures design where every participant performed on the same set of tasks – one training task and two experiment tasks. The order of tasks was counterbalanced in an effort to limit learning and boredom effects. The repeated measures design was chosen since it allows for the collection of a comprehensive data set from a moderate sample of participants in an attempt to obtain statistically powerful results [Murphy and Myors, 2003].

The experiment consisted of three parts; a pre-questionnaire, a set of search tasks and a post-questionnaire. For these parts each participant received the following[5]:

1. A *pre-questionnaire* that was first completed before proceeding with the search tasks.
2. Three *search tasks* (one training task and two experiment tasks) provided in connection with one overall background scenario. Each search task was comprised of a situation description and a task statement.
3. A *post-questionnaire* that was completed after all search tasks were finished.

Each participant also received a Sharp Zaurus 5500 Personal Digital Assistant (PDA) (see figure 6.1) with the running mobile application that was able to operate in two different system modes – a personalised mode and a non-personalised mode[6]. In the personalised mode, the system performed as a personalised information retrieval system. Search results were scored by the personalisation model based on the combined score of the IR model and the context model as described in section 5.4 in the previous chapter. In the non-personalised mode, the system ranked results using an IR system. Appendix H provides a more detailed description of the mobile application in relation to the experiment procedure. The remainder of this section describes the three parts of the mobile experiment in more detail.

The *pre-questionnaire* collected demographic and other descriptive information from each

[5]More details about the precise experiment procedure is provided in section 6.5
[6]The experiment setting used neutral code names for the two system modes to avoid bias.

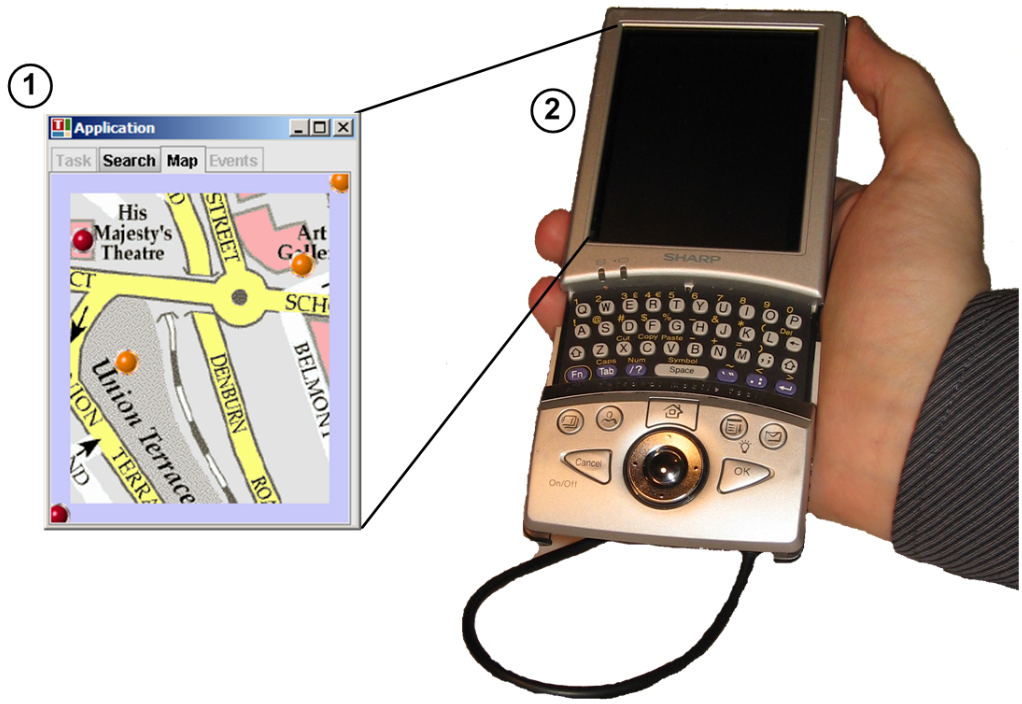


Figure 6.1: *Mobile application (1) running on a Sharp Zaurus 5500 Personal Digital Assistant (PDA) (2)*

participant. In particular, participants were asked for prior knowledge, experience and habits in the areas of mobile computing and search engines, as well as map and event information usage. For the *search tasks*, participants used the mobile application. They received a background scenario description and three *search tasks* (one for training and two for the experiment) each consisting of a situation description and a task statement. The background scenario described a simulated Musical and Dance Festival in Aberdeen offering a large number of events in and around the city centre. Each participant also received a small map of Aberdeen's city centre as part of the paper handout. The map contained details about streets and buildings as well as 6 event locations (see figure 6.2) that were highlighted inside the paper map using grey points (see figure 6.2). The same map was also provided in electronic form with the mobile application; however without the visual highlighting for potential event locations. Instead, events were displayed based on a colour schema (as shown in figure 6.3) as soon as participants searched for events and viewed them in the electronic map. Note that this colour schema was only used to

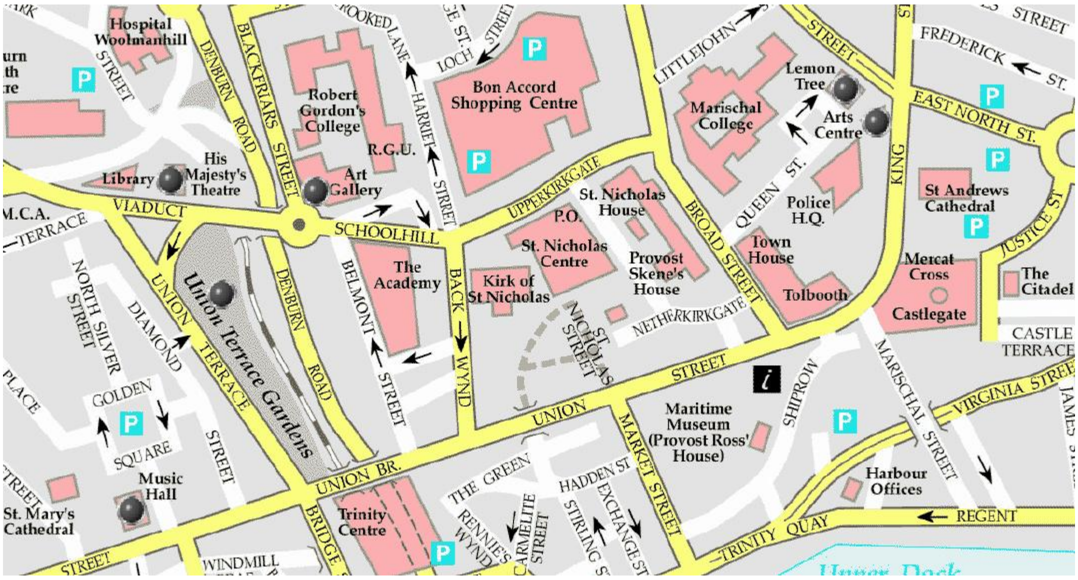


Figure 6.2: *Aberdeen city centre map with the 6 event locations as provided in paper handout. Grey points indicated all potential event locations when searching with the mobile application.*

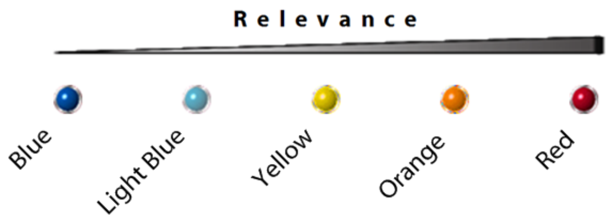


Figure 6.3: *Colour schema for more simplified representation of system relevance scores in the mobile application (red=highly relevant, blue=less relevant). This schema was also part of the colour handout (colour names only provided here for b/w support).*

visualise system relevance to the user but not for the user feedback[7]. The background scenario set the overall motivation for the experiment, prepared all participants equally for the experiment setting and allowed them to contextualise the more detailed task situations as shown below. Each search task consisted of a situation and a task statement. The following two textboxes show the two experiment tasks. The first box displays task 1 that asks participants to search for suitable musical events. The second box represents task 2 that requires participants to retrieve information about dance events[8].

[7]More details about the map visualisation are provided in section H.3 in the appendix on page 269.

[8]Context attributes are only highlighted (italics) in the example text here but have not been emphasised in the handout material. Furthermore, the text in the second box (task 2, dance events) has minor syntactical corrections compared with the original handout provided to participants. However, these changes do not alter the meaning of the text. The reader may find the original text of these handouts in

Task 1 (Musical Events)

Situation: *This morning*, you and your friend arrived in Aberdeen. After you found your hotel, you and your friend are roaming around in the city centre – now you are at the *Art Gallery*. It has come to your attention that there is currently a musical and dance festival in town offering a large amount of events in and around the city centre. That is a fortunate incident as you both share a common interest for good *musicals*.

Task: Bearing in mind the given situation, please find one or more suitable musical events which you would consider.

Task 2 (Dance Events)

Situation: A friend has phoned you *today* telling you about the currently ongoing Aberdeen Musical and Dance Festival which offers a large amount of events in and around the city centre. Your friend asked you to find information about *dance event performances* in order to select something. Your busy working day did not allow you do that so far; now you are at *His Majesty's Theatre* where you are going to meet your friend for a drink. You want to use the waiting time to search up some information before you meet.

Task: Bearing in mind the given situation, please find one or more suitable dance events which you would consider.

The situations offered contextual information for the participant to perform the task. Note that participants were neither instructed about the existence of contextual information, a context model nor where they informed about its use in the mobile application that was running on the PDA[9]. Each situation nevertheless provided

appendix G.

[9] This does not imply that participants only operated based on the provided stimulus. Participants obviously had different levels of contextual knowledge about musicals and dance entertainment and might have used this knowledge during the experiment.

information about participants' current location, time and focus of interest – all elements of the context model as described earlier.

Each situation was also provided with a task statement describing what the participant was supposed to do. The task statement had a generic structure that suggested searching for one type of event (e.g. musical events) in compliance with the situation provided. Participants were free to search for any event(s) within the limits of this task statement; thus only restricted in terms of the event type (i.e. musical events or dance events). It was up to the participant's personal interest to further specify searches for events based on this event type. No restriction was imposed regarding which events to view and to rate. There was also no additional guidance other than the background scenario, the situation description and the task statement for participants to base their judgements on. This was done in an attempt to encourage participants to behave realistically and accomplish a search task as naturally as possible. Section 6.5.3 provides more details on how the search and rating process proceeded based on the user interface of the mobile application. The mobile application as a whole is described in appendix H.

The *post-questionnaire*, completed after the end of the search tasks, ascertained information about participants' holistic impression of the experiment procedure. This included one question each about the suitability of the task situations, the suitability of the experiment locations, the level of interest / number / ease of the tasks, the level of interest in events and the overall usability of the software.

6.4 Participants

For this study, 17 people were recruited – mature students and members of staff from the Robert Gordon University as well as professional people from outside the university; a sample with a range of diverse people. The kind of people recruited for the sample seemed sufficient given that the application area of this research is the general public who uses mobile devices (e.g. mobile/smart phones) and enjoys entertainment events as part of modern lifestyle. This diversity of different people brings experience to the experiment task and introduces a wider range of demographic characteristics as described

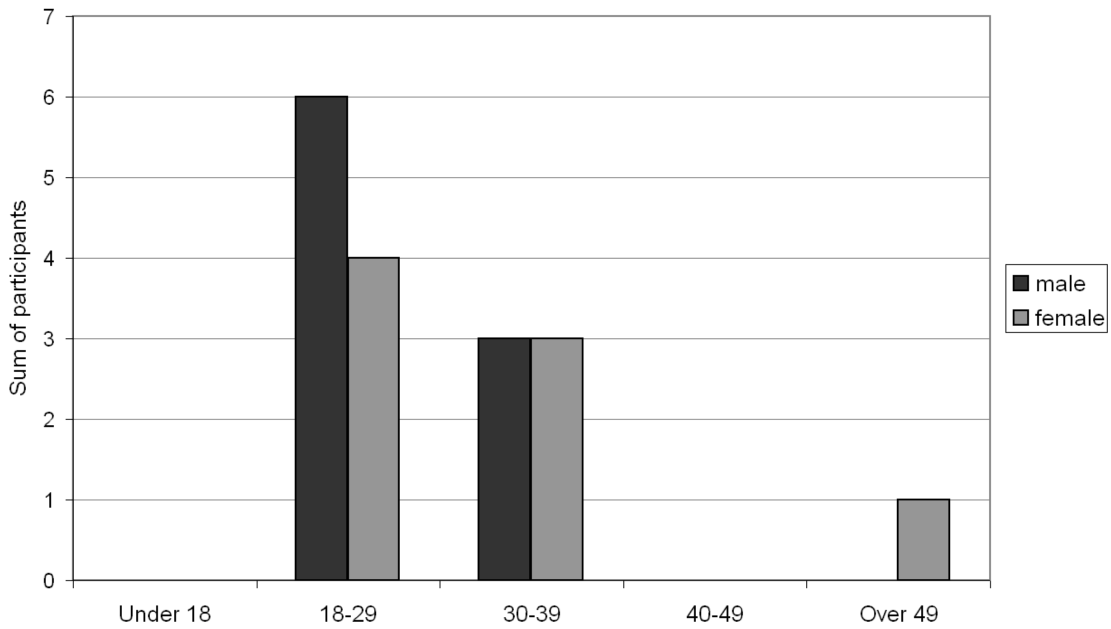


Figure 6.4: *Age and gender distribution of the 17 participants of the mobile experiment*

in [Harrison and List, 2004]. The sample was gender balanced (9 male, 8 female) with an emphasis toward the younger age (10 participants between 18 and 29) as shown in figure 6.4. As shown in figure 6.5 and 6.6, participants were very familiar with PCs, mobile phones, paper and electronic maps, search engines and generally the city centre of Aberdeen. They were less familiar with Personal Digital Assistants (PDA's). Furthermore, participants preferred searching events electronically rather than in the newspaper.

6.5 Experiment Procedure

This section describes the three parts of the experiment procedure as shown in in figure 6.7. Subsection 6.5.1 describes the pre-questionnaire that participants were required to complete. Subsection 6.5.2 provides details about the three search tasks (one training task and two experiment tasks) that participants were asked to perform using the mobile application. Subsection 6.5.3 additionally provides more details by representing an experiment task in close connection to the mobile application. Subsection 6.5.4 reports on the post-questionnaire that participants were required to complete.

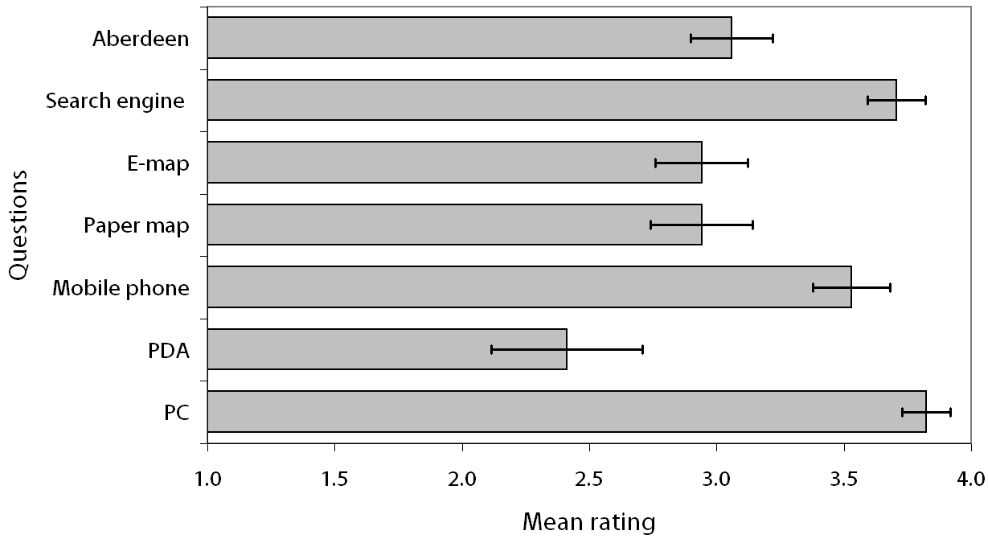


Figure 6.5: *Questions from the pre-questionnaire of the mobile experiment about participants' familiarity with PCs, PDA's, mobile phones, paper and electronic maps, search engines and the city centre of Aberdeen.*

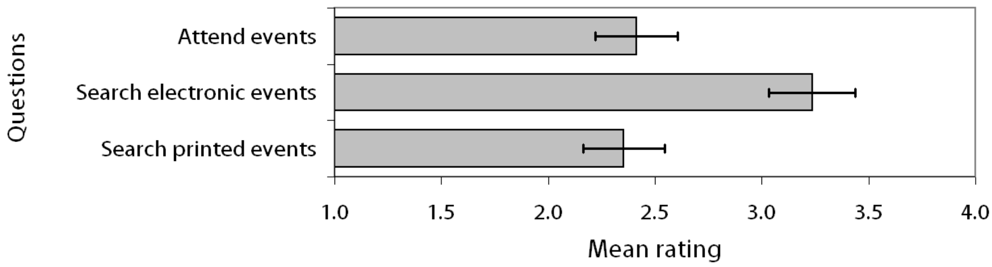


Figure 6.6: *Questions from the pre-questionnaire of the mobile experiment about participants' event search behaviour and frequency of attendance.*

Participants were appointed individually and each mobile experiment was carried out on an one-one basis. The mobile experiment was performed at two different places – the Art Gallery and His Majesty's Theatre in Aberdeen. Figure 6.8 shows the pictures of the two locations and relates them to the map[10]. The overview diagram in figure 6.7 further relates the two places with the two tasks.

Note that no pilot study was performed prior to the experiment, however, all event locations were well known and visited by the experimenter. The AmbieSense studies reported in chapter 3 also helped informing the experiment design and guiding its

[10]As stated previously, this map was part of the handout and also incorporated in the mobile application that was provided to participants on a Sharp Zaurus PDA (see page 39).

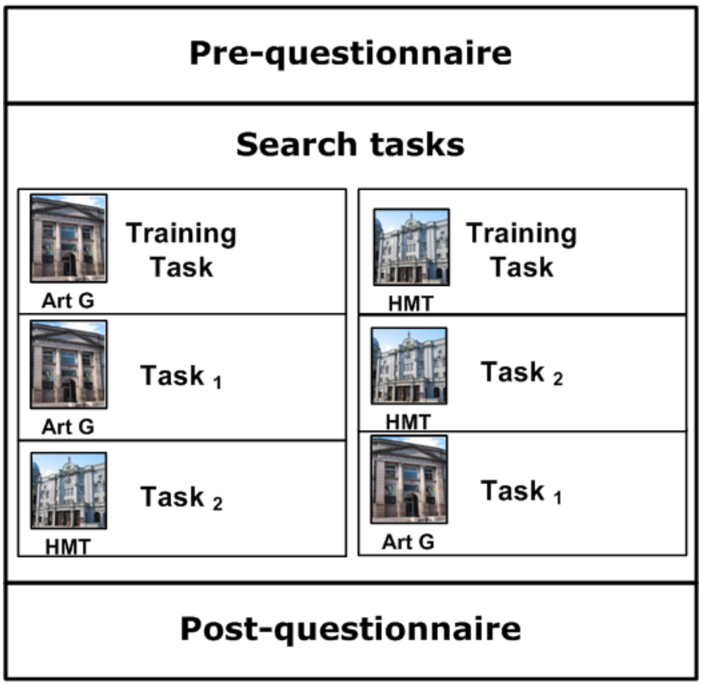


Figure 6.7: *Procedure of the mobile experiment based on the two cases whether a participant begins at the Art Gallery ('Art G.', left side) or at His Majesty's Theatre ('HMT', right side). People that started at the HMT first performed task 2 to avoid a change of location. For reasons of simplicity, the diagram does not show the counterbalanced order of system use at these places - refer to appendix G for these.*

conduction in the field.

6.5.1 Welcome and Pre-questionnaire

The experimenter met the participant at the first event location. Depending on the order of the tasks, this was either the Art Gallery or His Majesty's Theatre. Figure 6.7 visualises the experiment procedure based on whether the participant starts at the Art Gallery (left side) or at His Majesty's Theatre (right side). A short overview introduced the participant to the three parts of the study. The previously described stimulus was provided as a handout. The participant was informed that the procedure involved one change of location; one half of the study was performed at the current location; namely the pre-questionnaire, the training task and the first experiment task. The second half of the study was completed at the second location; namely the second experiment task and the post-questionnaire. In other words, participants who completed their first task at the Art Gallery relocated later to His Majesty's Theatre for the second task. Participants

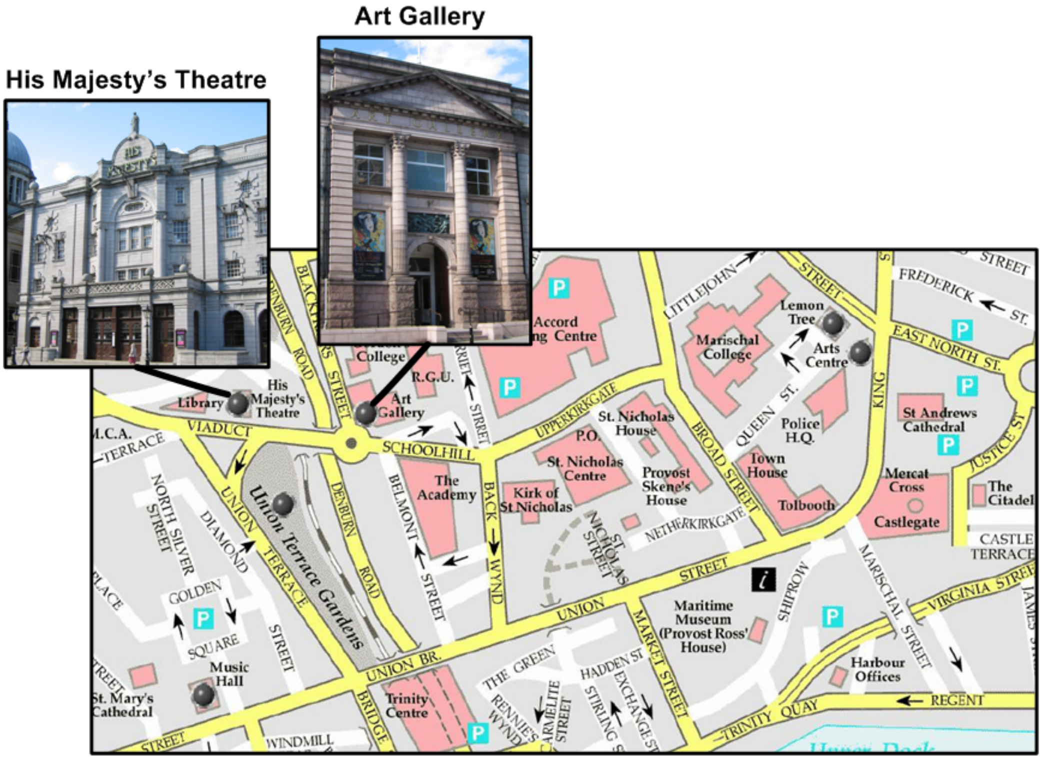


Figure 6.8: *Aberdeen city centre map with highlighted event locations visited during the mobile experiment (His Majesty's Theatre and Art Gallery).*

who started at His Majesty's Theatre changed their location later to the Art Gallery.

Participants started with the pre-questionnaire answering basic demographic questions and some questions about their background knowledge in mobile computing, search engines, map and event information usage.

6.5.2 Search Tasks

After completing the questionnaire, participants were directly introduced to the search tasks. The experimenter read the background scenario to the participant and explained that all tasks (training and experiment) were situated within the festival background scenario. Participants were informed that the mobile application on the PDA produces relevance scores for the event content based on the query submission to the system. They were also told the system represented search results using colours based on a colour schema; a simplified visual representation of the scores (see figure 6.3 on page 136).

Five colours represent the range of possible system scores – either personalised or non-personalised based on the current system mode. Participants were not told how these scores were determined thus leaving the two system modes opaque. They were instructed to rate individual events for their usefulness, based on provided background scenario and task situation[11]. Usefulness was explained to participants as situational relevance, as described in [Borlund, 2003a]. Participants first solved a training task under supervision of the experimenter. The training task allowed participants to become familiar with all parts of the system and all features of the user interface; particularly the different ways of navigating between the various parts of the system. Note that data from the training task was not used for analysis. After the training task, the participant solved the first experiment task once in the non-personalised system mode and once in the personalised mode. Experiment tasks were not supervised, unlike in the training task. The experiment location was physically changed after the first experiment task. Participants who performed their first task at the Art Gallery, were brought to His Majesties Theatre and vice versa. At the second location, participants performed the second task also in both system modes. A counterbalanced experiment design over tasks and systems was used in an attempt to limit effects caused by learning and boredom.

6.5.3 The Search Task with the Mobile Application

Figure 6.9 shows an example of search task 1 (musical events) in the personalised system mode[12] of the mobile application visualised by its user interface screens. The example is reproduced from one of the automatic user logs that were recorded during the mobile experiment. The user interface of the mobile application consisted of four different views – "Task"(1), "Search"(2), "Map"(3) and "Events"(4)(5)[13]. These views provided a wide range of personal freedom by offering (relatively) unrestricted navigation between the functionalities of searching, browsing (both geographic and by using ranked lists) and viewing of events.

[11]Note that users did not rate other types of relevance (e.g. topical relevance).

[12]The two system modes of the mobile application were opaqued by naming. The "Blue" system presented the non-personalised system mode and the "Green" system the personalised mode.

[13]For the chosen example, screen (5) expands over two display lengths. This is highlighted with the two border lines in figure 6.9.

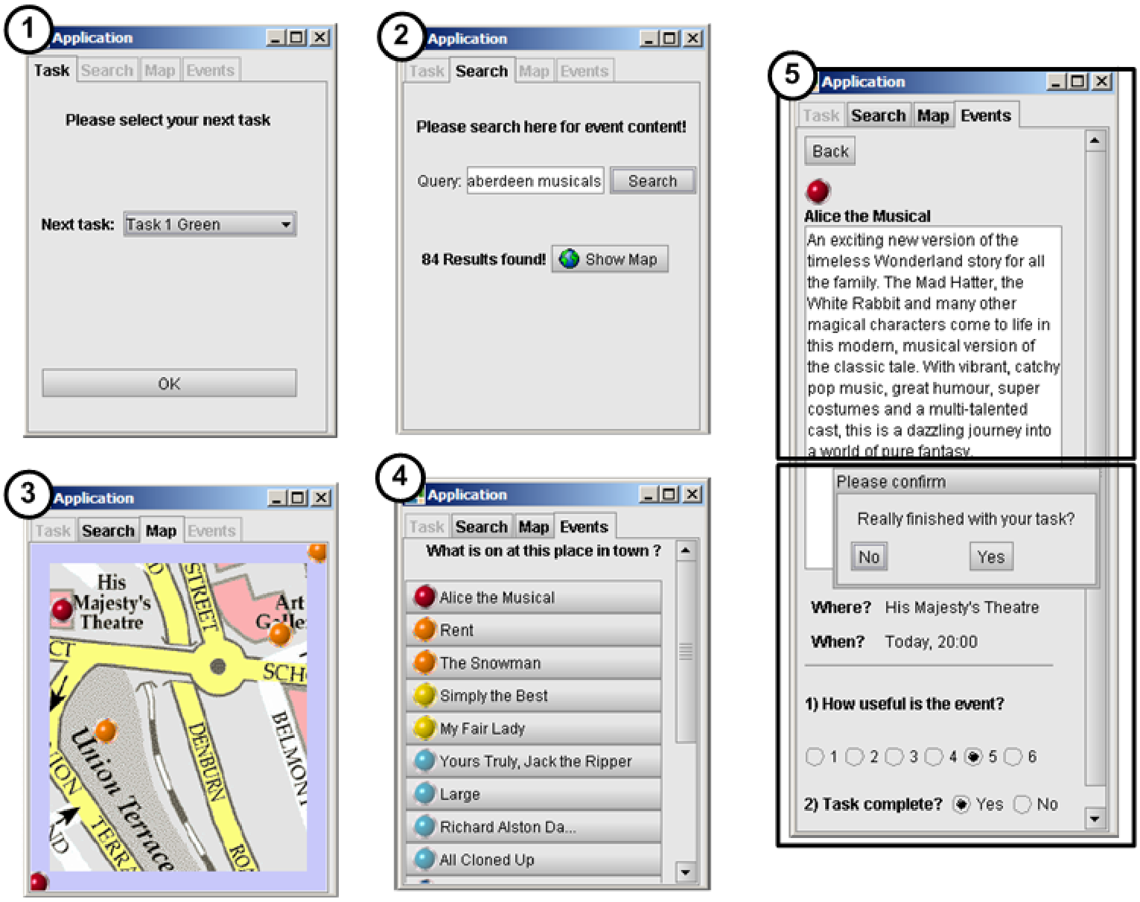


Figure 6.9: *User interface views of the mobile application based on an example of search task 1 (musical events); represented with different views for task selection (1), search (2), map browsing (3) and event viewing (4)(5).*

The process started when the experimenter selected a new task together with either of the system modes (1) (see figure 6.9). After this selection, the application switched into the search view (2). At this point, the device was handed over to the participant. Here, the participant submitted one or more search queries (e.g. "aberdeen musicals") to retrieve events based on the provided stimulus (background scenario, situation description and task statement). After every query submission, the number of results was displayed at the bottom of the search view (e.g. 84 in the example) together with a button to switch to the map view (3), a small geographic browser. The map displayed all retrieved events as geographic points based on relevance as determined by the current system mode. In this view, the participant freely navigated the map via drag-and-drop. Event locations with results were represented by the top ranked event at this location (i.e. the event with the highest score) based on the colour schema as shown in figure 6.3 on page 136.

Event locations without results did not appear in the map. All those events that were located outside the viewable area of the map were visualised using an extended viewing feature; a small border surrounding the viewable area of the map on which nearby events were displayed. This viewing aid allowed people to identify events from the entire map even if the display was limited to a small part. Section H.3 in the appendix describes this feature in more detail. The participant eventually selected an event location on the map and the application switched to the 'Events' view (4). This view presented a ranked list of all events available at the selected event location. Events were shown with titles and scoring information based on the colour schema as described in section 6.3. Upon selection of one entry, more detailed information about the event was displayed (5); the event description, the performance time and the venue. At this point, the participant was required to provide a rating of usefulness (understood as situational relevance as described in [Borlund, 2003a]) on a 6-point scale. After that, participants were free to continue browsing and searching until task completion. Navigation was supported with tabs that allowed participants to freely switch between search (2), map (3) and event views (4)(5) – the only condition being that every newly viewed event had to be rated[14]. The completion of the task was declared by the participant and could be set every time when viewing an event as shown in (5). An additional dialog box was installed to prevent task completion being selected by mistake. After the task was completed, the view changed back to the task view (1). At that point, the PDA was returned to the experimenter who selected the next task to continue with the experiment procedure. The participant completed one training task in one of the two system modes as well as two experiment tasks for both system modes, meaning that the procedure was repeated five times.

6.5.4 Post-questionnaire

The experiment procedure finished after the completion of the *post-questionnaire* that requested feedback about the overall impression of the study. In particular, participants were asked to rate the suitability of the task situations, the suitability of the experiment locations, the level of interest / number / ease of the tasks, the level of interest in the events and the overall usability of the software. The post-questionnaire's main purpose

[14] This small restriction was to ensure that participants provided enough data.

was to sample participants' opinion about the quality of the main experiment parameters. The next section describes the findings from the experiment before they are discussed in greater depth in section 6.7.

6.6 Experiment Results

This section presents results from data automatically logged during the experiment and from the post questionnaire completed by participants after finishing the experiment. This range of different types of data allows drawing an initial picture about the differences between conventional search and personalised search with context. The remainder of this section is divided into four subsections. The first subsection describes the findings concerning participants' experience of usefulness and search effort between the two systems; this directly tests the three hypothesis as proposed in section 6.2. Subsection 6.6.2 presents descriptive data about participants' search behaviour. These results are compared with other research conducted on public search engine query logs of various sizes – the long-term, billion log AltaVista[15] study reported in [Silverstein et al., 1999], the million log Excite[16] study [Spink et al., 2001] and the comparative study from AlltheWeb[17] on two one million logs from 2001 and 2002 [Jansen and Spink, 2005]. Subsection 6.6.3 provides more details on participants' rating behaviour. The last subsection reports on the data collected in the post-questionnaire.

6.6.1 Usefulness and Search Effort

Non-parametric Wilcoxon signed-rank tests [Wilcoxon, 1945] were applied to compare the two systems of the mobile experiment. Unlike parametric statistical tests, they do not rely on distributional assumptions in the data but still provide strong and reliable results. In addition, they complement the data analysis from chapter 4 and chapter 5 with a different type of statistical validation. The Wilcoxon signed rank test was selected since the same participants used both systems in the experiment procedure. Note that participants completed each task twice – one time on each system – which is considered a design limitation. For this reason, a careful look is taken at system order effects.

[15] `http://www.altavista.com`, accessed April 14, 2008
[16] `http://www.excite.com`, accessed April 14, 2008
[17] `http://www.alltheweb.com`, accessed April 14, 2008

Table 6.1 shows descriptive statistics and table 6.2 presents the statistical test for the three measures that were introduced and discussed in section 6.2. Both tables provide statistics for individual tasks and for both tasks combined. The assigned usefulness obtained from participants' relevance judgements (usefulness) represented a measure of content quality. The amount of time (task time) and the number of submitted queries (query number) were used as a measure of user effort. In particular, the usefulness that the participant assigned to the retrieved event was measured, the time that a participant took to finish a task was assessed and the number of queries the participant submitted until the task was completed was measured. Table 6.1 shows the measures with means and standard deviations (sd) for both tasks (individually and combined) for the personalised system (P) and the non-personalised baseline system (NP). Table 6.2 presents the z-scores, significance values (sig) and effect sizes expressed in Pearson Correlation Coefficients (effect(r)). For the three measures, tests were performed for significant differences between the two systems.

Users of the personalised system found more useful event content than with the system that did not personalise search results contextually (z=-5.995, p<.001, r=.25).

measure	**mean (NP)**	**sd (NP)**	**mean (P)**	**sd (P)**
task time (task 1)	275s	153s	254s	96s
query number (task 1)	1.94	2.90	1.47	0.62
usefulness (task 1)	3.18	2.06	4.20	1.83
task time (task 2)	356s	305s	317s	187s
query number (task 2)	2.59	2.98	1.71	1.11
usefulness (task 2)	2.75	1.82	3.69	1.72
task time (both tasks)	315s	241s	285s	150s
query number (both tasks)	2.26	2.92	1.59	0.89
usefulness (both tasks)	2.96	1.95	3.95	1.79

Table 6.1: *Means and standard deviations (sd) for the musical event task (task 1), the dance event task (task 2) and both tasks on the non-personalised system (NP) and the personalised system (P) for task time (in seconds), usefulness (6-point scale between 1 (lowest) and 6 (highest)) and query numbers per task.*

Figure 6.10 shows the mean difference of usefulness rated by participants. On average based on both tasks, content provided by the personalised system was rated 19.8 % more useful than content that was searched on the non-personalised system. The rated

measure	z-score	sig.	effect (r)
task time (task 1)	-.260	.795	.04
query number (task 1)	-.577	.564	.10
usefulness (task 1)	-4.024	.000	.24
task time (task 2)	-.308	.758	.05
query number (task 2)	-1.364	.172	.23
usefulness (task 2)	-4.256	.000	.25
task time (both tasks)	.316	.752	.04
query number (both tasks)	-.752	.452	.09
usefulness (both tasks)	-5.995	.000	.25

Table 6.2: *Test for significant differences between the two systems. Z-scores and Wilcoxon signed-rank significance tests (sig) with effect sizes (effect(r)) for the musical event task (task 1), the dance event task (task 2) and both tasks for task time (in seconds), usefulness (6-point scale between 1 (lowest) and 6 (highest)) and query numbers per task between the two systems.*

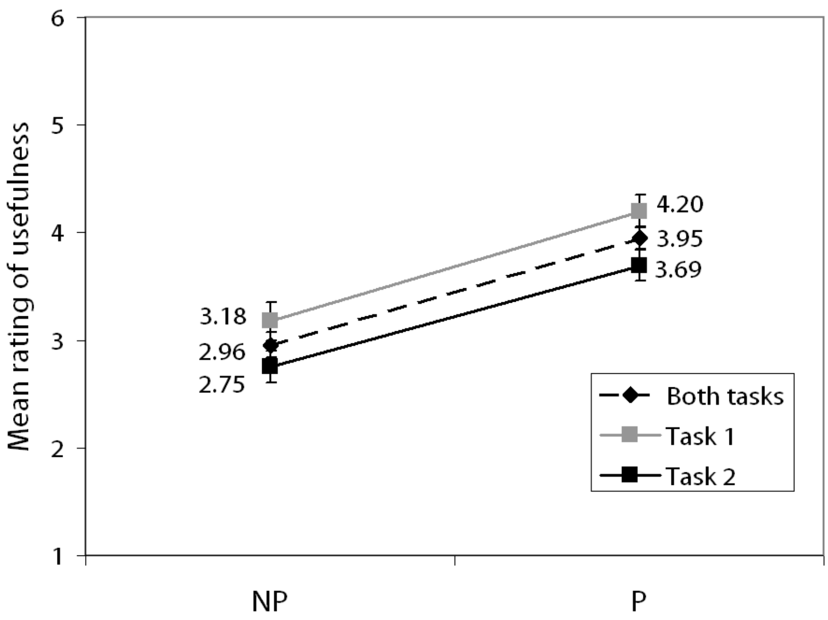


Figure 6.10: *Mean difference of rated usefulness between the personalised (P) and the non-personalised (NP) system for both tasks and individually (error bars indicate standard errors) based on logs from the mobile application.*

difference between the two systems for task 1 was slightly smaller (18.8%), than for task 2 (20.4%). Subsection 6.6.3 provides more details on how participants rated event content.

As mentioned earlier, every participant performed each task one time with the personalised and one time with the non-personalised system and was assigned to one of two possible orders based on the experiment design. Participants either performed a

system order on measure	z-score	sig.	effect (r)
system order on usefulness (task 1)	-1.067	.286	.10
system order on usefulness (task 2)	-3.182	.001	.27
system order on usefulness (both tasks)	-2.637	.008	.16

Table 6.3: *Test for significant differences between system orders for usefulness ratings with z-scores and Wilcoxon signed-rank significance tests (sig) with effect sizes (effect(r)) for the musical event task (task 1), the dance event task (task 2) and both tasks.*

measure	mean (NP)	sd (NP)	mean (P)	sd (P)
task 1 (NP→P)	3.51	1.95	4.20	1.85
task 1 (P→NP)	2.84	2.13	4.20	1.82
task 2 (NP→P)	3.01	1.78	4.15	1.64
task 2 (P→NP)	2.45	1.83	3.16	1.67
both task (NP→P)	3.24	1.87	4.18	1.75
both task (P→NP)	2.64	1.99	3.65	1.81

Table 6.4: *Means and standard deviations (sd) for task 1, task 2, and both tasks on the non-personalised system (NP) and the personalised system (P) for usefulness grouped by the order in which users applied the two systems. (NP→P) refers to ratings collected from participants that first applied the non-personalised and then the personalised system. (P→NP) refers to ratings collected from participants that first used the personalised and then the non-personalised system.*

task first with the non-personalised and then with the personalised system (NP→P) or vice versa (P→NP)[18] These two orders were counterbalanced across the user population, however, the repetition of tasks with the same participant is considered as a disadvantage. For this reason, a closer look was taken at potential ordering effects caused by people performing the same task twice on both of the systems. As shown in table 6.3, tests revealed that participants of task 2 rated usefulness different based on the order in which they applied the two systems ($z=-3.182$, $p<.01$, $r=.27$). The table also shows a significant order effect for both tasks in combination ($z=-2.637$, $p<.01$, $r=.16$). Table 6.4 presents the usefulness measure with means and standard deviations (sd) for both tasks (individually and combined) differentiated by the two possible system orders. Note that task 1 was no significant regarding system orders and is therefore not discussed.

The figure below shows the mean differences in usefulness ratings for the non-personalised (NP) and the personalised system (P) based on table 6.4. Usefulness ratings

[18]Appendix G on page 256 provides all details on the task orders used for this experiment.

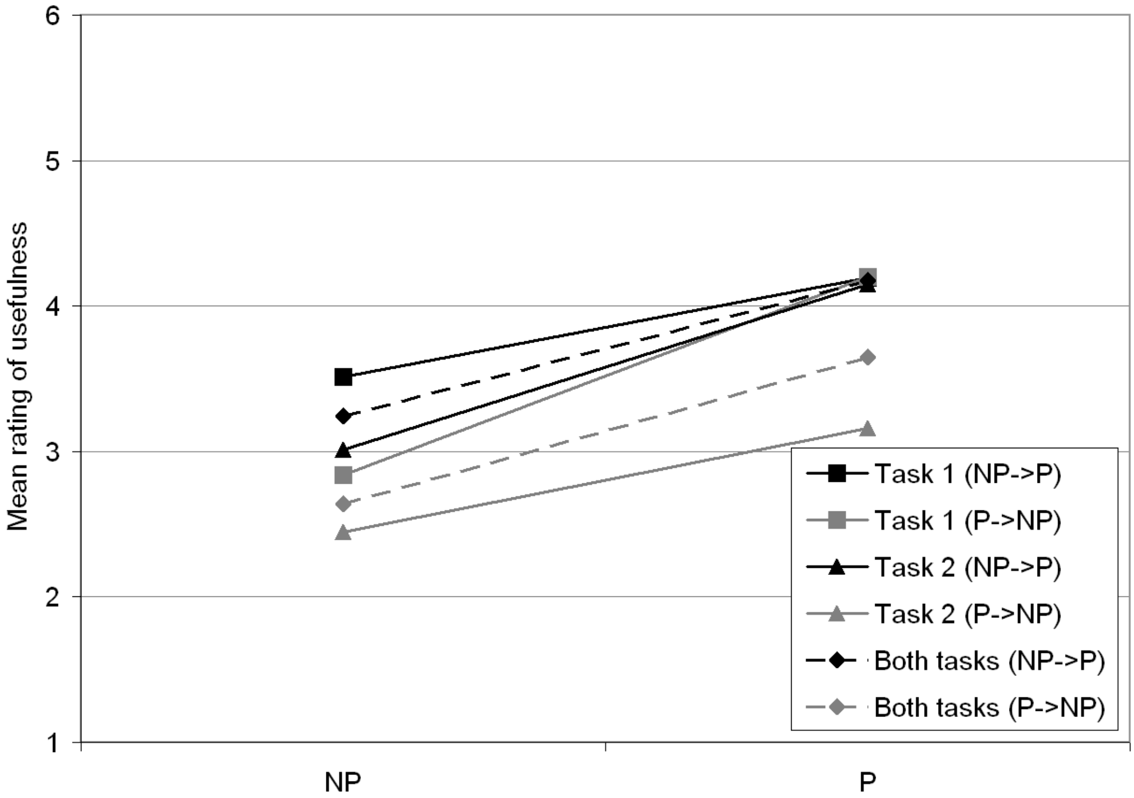


Figure 6.11: *Mean difference of rated usefulness between the personalised (P) and the non-personalised (NP) system for both tasks and individually. Differentiated display based on two groups of users; (NP→P) refers to ratings collected from participants that first applied the non-personalised and then the personalised system. (P→NP) refers to ratings collected from participants that first used the personalised and then the non-personalised system.*

are differentiated by task (individual and in combination) and the two possible system orders (NP→P or P→NP) as described before. Ideally, each pair of lines (NP→P and P→NP) for each of the tasks should match; this would indicate that all system order effects were completely eliminated. However, the figure reveals a small difference between the pairs for task 2 and for both tasks; system order effects for task 1 are not significant and are therefore not considered. In particular, people rated slightly higher when they started with the non-personalised system and slightly lower when they started with the personalised system[19]

Despite this, events retrieved with the personalised system were still more useful

[19]This can be seen when comparing the grey lines with the black lines in the figure.

on average in all cases as demonstrated by the rising trend in all of the lines in the figure. On average, based on both tasks, personalised content was still rated about 20% higher on the scale (18.6% for the the NP→P group and 20.1% for the P→NP) indicated by its parallel lines[20]. However, the P→NP had a generally lower magnitude for ratings. Task 2 had a larger difference with personalised events rated 22.8% more useful by participants that used the non-personalised system first but only 14.3% by participants that used the personalised system first. Task 1 did not have significant system order effects based on table 6.3 and is therefore not considered.

Participants spent on average 30 seconds less with the personalised system than with the non-personalised system to accomplish a task. This difference was slightly smaller with task 1 (21 seconds) than with task 2 (39 seconds). The time difference however was overall not statistically significant (z=-.316, p=.752, r=.04). Participants submitted on average less queries with the personalised system (1.59 queries) than with the non-personalised system (2.26 queries); a trend that is consistent with individual tasks. This difference however was not statistical significant (z=-0.752, p=.452, r=.09). It is possible that these two measures of search effort would gain strength from a larger sample of participants. Note that task time and the query numbers were not effected by system orders. The next subsection takes a more detailed look into participants' search behaviour.

6.6.2 Search Behaviour

As already described in the previous section, the amount of queries that participants submitted to solve the experiment tasks with the two systems (non-personalised (NP) and personalised (P)) did differ but without statistical significance. This subsection takes a further look into participants' general search behaviour. In particular, descriptives statistics about query numbers and query term numbers are provided and a closer look into the query formulation process is taken.

[20] Small differences due to rounding.

Amount of Queries

On average, participants submitted 1.9 queries per task (2.26 (NP), 1.59(P)). Figure 6.12 shows the distribution of the amount of queries that were submitted. On the

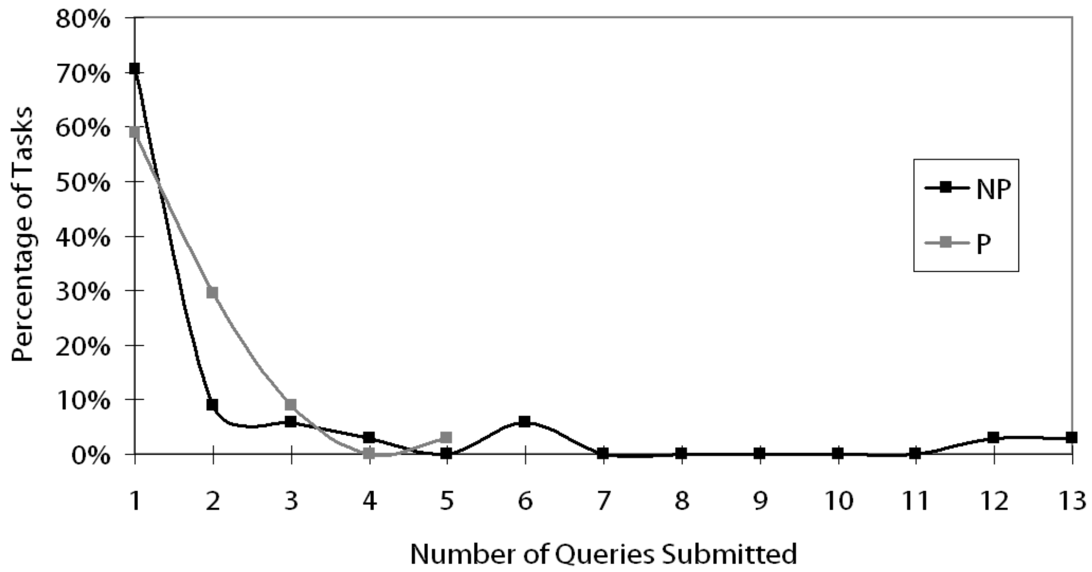


Figure 6.12: *Distribution of the number of queries per participant and task on personalised (P) and non-personalised (NP) system based on logs from the mobile application.*

non-personalised system 70.6% of the tasks were solved with only a single query; 8.8% with two and 5.9% with three queries. More than three queries were used in 14.7% of the cases. In comparison, users of the personalised system were less likely to use a single query (58.8%) and more likely two queries (29.4%). Three queries occurred in 8.8% of the cases and very few tasks were completed with more than three queries (2.9%). No user of the personalised system submitted more than five queries, an effect that is potentially caused by the small amount of participants.

This result relates to the findings from the search engine query log studies. The AltaVista study [Silverstein et al., 1999] reported an mean query amount of 2.02 whereas [Spink et al., 2001] found a mean of 2.52 unique queries per session. The AlltheWeb study [Jansen and Spink, 2005] reported a mean amount of 2.3 queries for the most recent 2002 data set in comparison to an equally large 2001 data set with an average of 3.0 queries.

Distributions also largely correspond to the findings. The AltaVista study reports 70.6% single query and 13.5% two query sessions. In [Spink et al., 2001] it was found that 48.8 % of web session are solved with single queries followed by 20.8% with two queries. The more recent AlltheWeb study shows a similar trend of most sessions having one query (59%) followed by very few session with two (23%) and more than two queries (25%).

Amount of Query Terms

The mean number of query terms was 2.8 with only very little variation between the two systems. The amount of query terms follows the shape of a normal distribution as shown in figure 6.13. Most queries where composed of two terms (28.8%), followed by three

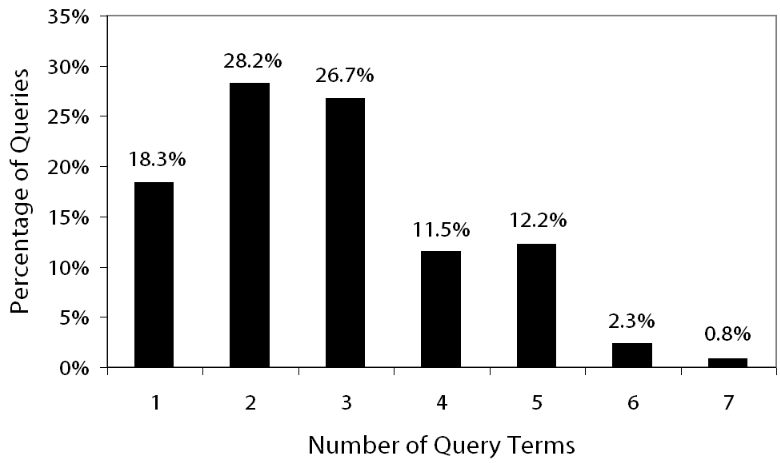


Figure 6.13: *Distribution of the number of query terms used on both systems based on logs from the mobile application.*

terms (26.7%) followed by a single term (18.3%). This results is similar to the search log findings; the AltaVista study reported a mean of 2.35 terms, Spink's Excite study a mean of 2.4 terms, followed with very similar results for the AlltheWeb study with 2.3 query terms for the 2001 data and 2.4 query terms for the 2002 data set. The query term distribution from the experiment is generally very similar to the findings from the search engine studies. The AlltheWeb study however shows more averaged trends on the 2002 data set with one third of one term queries (33%), one third of two term queries (33%) and one third (34%) of queries with more than two terms. This is a bit surprising given that users were performing tasks on small mobile devices. It was expected that users in mobile contexts submit shorter and more focused queries in comparison to those submitted on personal computers in more stationary environments. However, it is possible that the query box as part of the user interface caused some of that behaviour. The query box as part of the search view (see (2) in figure 6.9 on page 144) might have influenced people to adapt their search behaviour closer to normal web search. The size of the query box might also have influenced the length of their search queries. This was not the focus of this study, however, would be worthwhile for future investigations.

Query Formulation

Table 6.5 provides a list of the most frequent query terms. The list is divided by task

task 1 **term**	**(musical events)** **frequency**	**task 2** **term**	**(dance events)** **frequency**
musical	31	dance	53
aberdeen	27	aberdeen	31
musicals	12	festival	20
today	10	today	15
music	8	events	12
festival	8	scottish	8
tonight	6	musical	7
event	6	salsa	6
concert	5	folk	6
events	4	theatre	5
center	4	event	5
city	4	class	3
comedy	4	dancing	3
dance	4	portuguese	2
20:00	3	traditional	2
ballet	2	ceildish	2
play	2	ball	2
summer	2	african	2

Table 6.5: *Query term frequencies of valid submitted queries for task 1 (searching musical events) and task 2 (searching dance events) based on logs from the mobile application. Single term occurrences and stopwords have been removed.*

since participants most likely targeted their queries on the tasks that were given to them. It is not necessary to further investigate into precise query term distributions as the total amount of queries was limited due to the small number of participants. It is however worth looking into the influence of context on the query construction. When reviewing the most frequent query terms for each of the two tasks, it is evident that contextual information entered the query. The most frequent terms are in fact those that relate to the three investigated attributes; terms describing the location (e.g."aberdeen"), the categorical interest (e.g. "musical" and "dance") and temporal aspects (e.g. "today"). This indicates that participants perceived the experiment conditions which guided and focused their search effort during the experiment. Despite this, queries that participants submitted during the mobile experiment turned out to be very general and unspecific. This resulted in generally very low information retrieval scores. On average, information retrieval scores were strongly skewed to the low end with an average score of 0.03 for all retrieved events during the experiment. The average context score however was about nine times higher with an average of 0.28. This means, even though the personalisation model was balanced as described in chapter 5, information retrieval was much less influential on the total score as originally expected. This is not an effect that should be linked to the personalisation model but instead to the particular circumstances of the experiment and its conduction. Potential reasons for this effect will be discussed in section 6.7.

6.6.3 Event Rating Behaviour

As described in subsection 6.6.1, the mean amount of usefulness assigned to personalised events was about 20% higher than to non-personalised events. This subsection provides further descriptive results that highlight the differences between the personalised and the non-personalised system mode with respect to participants' rating behaviour.

The median number of rated events was 8 which means, on average, participants viewed 8 different detailed event descriptions and later rated them[21] on a scale between 1 (not useful) and 6 (highly useful). Unlike in the previous section, the results from the research on search engines do not directly compare. This is due to differences in structuring and visualising search results. Search engines normally present results using a sequence of result pages of fixed length. The mobile application however divided search results per location and distributed them on a two-dimensional map. Also, search engines measure the number of viewed result pages based on the query and not in relation to the task. Figure 6.14 shows the distribution of ratings on the provided 6-point rating scale for both system modes. The figure reveals a more detailed view on participants event rating

[21] As mentioned earlier, the research prototype was designed in a way that required participants to rate an event as soon as it was viewed. This ensured that they provided enough data for the study.

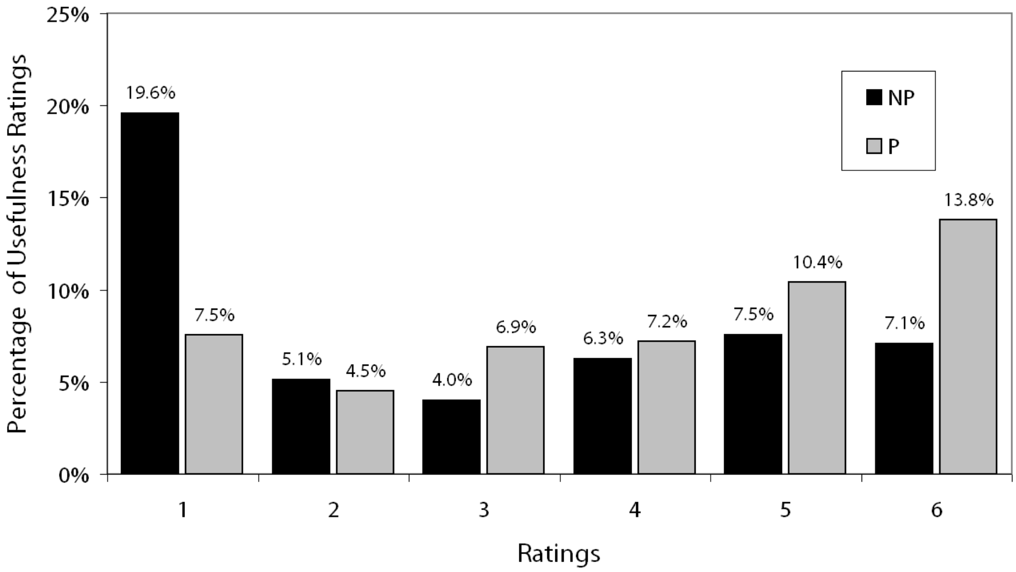


Figure 6.14: *Distribution of usefulness ratings between 1 (not useful) and 6 (highly useful) for the non-personalised system (NP) and the personalised system (P) based on logs from the mobile application.*

behaviour and shows the differences between the two distributions. The non-personalised system caused participants to view and rate a large quantity of results that turned out to be of no use for the given task. The personalised system, on the other hand, is clearly skewed towards higher ratings indicating that participants generally tended to find more useful events when they where personalised. Participants were not instructed how to browse and how to select events for closer inspection. This offered high freedom of choice and a degree of realism that approximates operational conditions. For this reason, it is interesting to know how likely events at different rank position were rated. Figure 6.15 shows the cumulative ratings of events at different rank positions based on all rated search results in the mobile experiment. The figure shows data for the non-personalised system mode (NP), the personalised system mode (P) and a combined result (Both). It depicts that events at top ranks were generally viewed and rated more often than events at lower ranks. In other words, people naturally targeted high up ranks when selecting content for inspection. This is demonstrated in the steep rise at the beginning of the curve (top rank positions) and its reduction to a lesser extend later on (low rank positions). The personalised system made participants more likely rate top ranks than

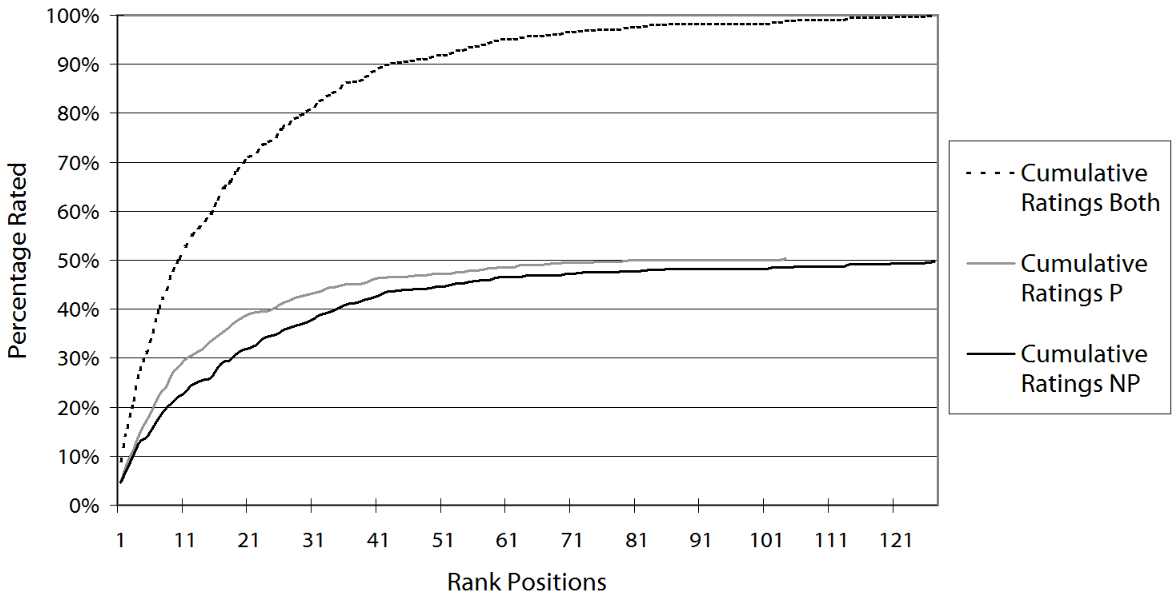


Figure 6.15: *Cumulative ratings of content in absolute rank positions for the non-personalised system (NP), the personalised system (P) and both systems (Both).*

the non-personalised system. At rank position one, 9% of all rated events were rated (4.5% (NP) and 4.5% (P)). More than half of the events (50.4%) were rated up to rank position 10 (22.2% (NP) and 28.3 (P)[22]).

6.6.4 Post Questionnaire

Besides the information logged during the mobile experiment, participants also completed a questionnaire after the experiment procedure where they expressed their opinion about the whole experience. Results showed high levels of agreement of participants with the suitability of the task situations, the suitability of the experiment locations, the level of interest / number / ease of the tasks, the level of interest in the events and the overall usability of the software (see figure 6.16). Participants also judged the performance of the two systems[23] by answering three additional questions as shown in figure 6.17. Most people agreed with the statement that the personalised system (P) outperformed the non-personalised (NP). The second strongest opinion was that both system where equal. The

[22]Numbers do not add up because of rounding.

[23]Since the experiment was blinded, participants answered these questions based on the system code name that was used during the experiment.

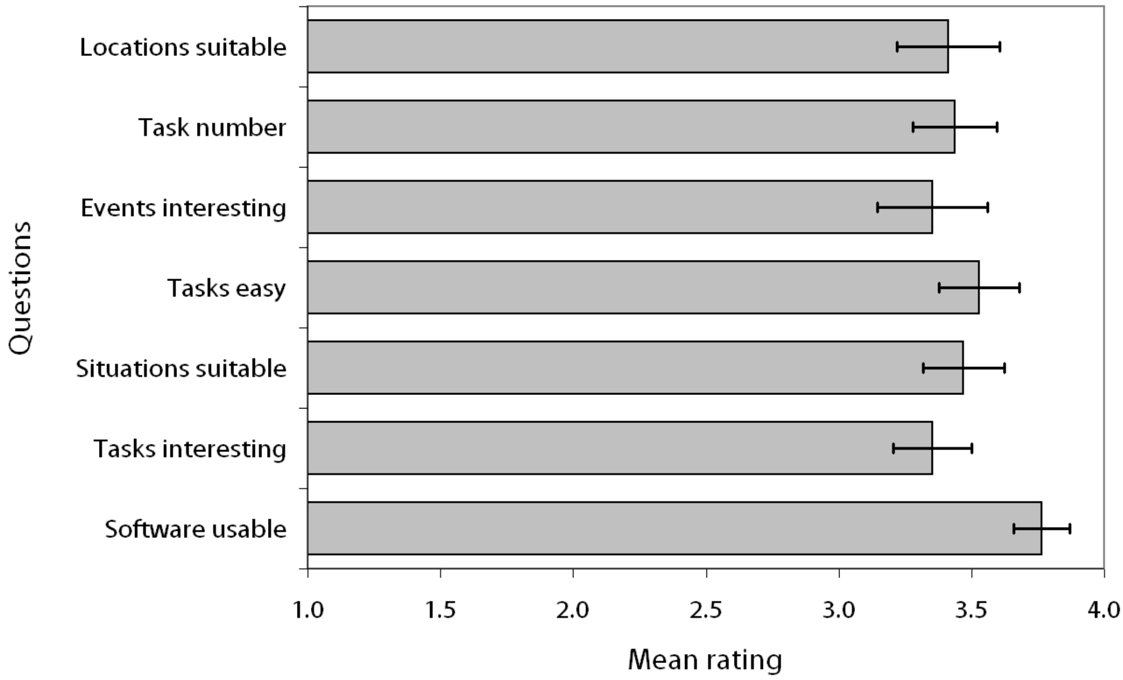


Figure 6.16: *Mobile experiment post-questionnaire results (error bars indicate standard errors).*

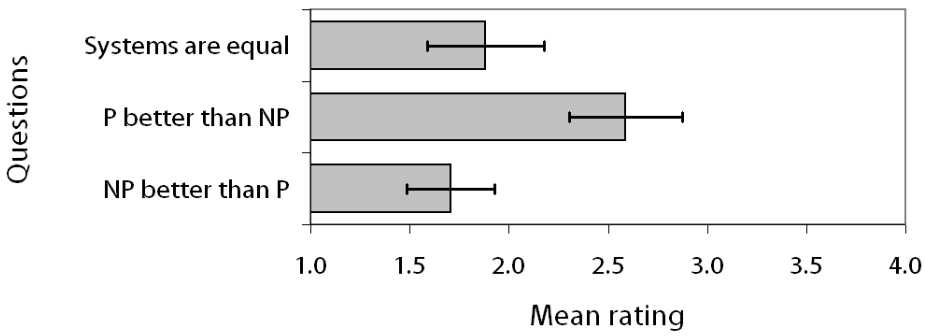


Figure 6.17: *Mobile experiment post-questionnaire results about the system (error bars indicate standard errors).*

weakest opinion was that the non-personalised system outperformed the personalised.

6.7 Discussion

The mobile experiment measured the overall effect of content personalisation with context. This was done using the personalisation model described in chapter 5 as part of a mobile application for users in mobile situations. The study was conducted based on the following three hypothesis as proposed earlier in section 6.2:

- People find more useful event content with a system that provides personalised event search results based on context compared with a system that provides non-personalised results. (H_5).
- People solve search tasks faster with a system that provides personalised event search results based on context compared with a system that provides non-personalised results. (H_6).
- People solve search tasks with fewer queries with a system that provides contextually personalised event search results based on context compared with a system that provides non-personalised results. (H_7).

Whereas usefulness represents a measure for the content quality that users receive, task time and the amount of submitted queries represent measures of users' effort to fulfil an information need. The remainder of this section discusses the whole of the results based on these two aspects representing the three hypothesis – *usefulness* as a measure of the content quality as well as task time and query amount as a measure of users' *search effort*. Furthermore, each of the following two subsections describes and discusses the limitations of this study with respect to the experiment design and its conduction.

6.7.1 Usefulness

Participants rated the usefulness of personalised events significantly higher compared with non-personalised results. Ratings of personalised events were on average 19.8 % more useful in comparison to non-personalised results. One limitation in the design of this study is that each task was performed twice by each participant – one time on each system. This is considered as a disadvantage and caused usefulness ratings being effected by the order in which the two systems were used by participants. People who first applied the non-personalised and later the personalised system produced higher ratings than users who performed their task in the opposite order. Nevertheless, on average all participants of both system orders still found personalised events about 20% more useful. This means that participants' experience of content quality has improved by 1/5 when using personalisation with the given model. It shows evidence for the ability of contextual information to be used for the personalisation of event content (H5). The result validates the context model as a whole in its ability to contribute both statistically and practically significant to the value of a mobile information system. This result is consistent with the findings from the post questionnaire where participants' largely agreed that the personalised system performed better than the non-personalised system. In the following, the measure of usefulness is further discussed and reviewed regarding possible limitations:

- *Mobile Application:* As previously stated, the user interface of the mobile application served as the main interface for the experiment procedure providing a range of interactive possibilities for participants to solve their search tasks. Specific choices implemented in the user interface might have affected participants' ratings and therefore limited the validity of the results obtained from this study. Recalling from section 6.5, the user interface visualised events in partial ranked lists presented when clicking on an event location in the map. Events on these ranked lists were displayed with the event title and a coloured representation of the score. Each of these partial result lists only contained a set of ranked events for this particular event location. This allowed people to browse events based on location but also to generally divert from top ranked events. Furthermore, it is possible that the event titles from these result lists did sometimes not indicate enough details about the content of the event. Therefore, selecting useful events might have been difficult for participants in cases with short or otherwise non-explanatory event titles. In other words, participants might have missed useful events or, on the contrary, might have selected events of low use for closer inspection. This might result in more events being rated as non-useful. However, the 20% difference of ratings between the personalised and the non-personalised system mode should be unaffected by the decisions taken in the user interface. Ratings from the non-personalised system mode represent the control group that were collected with the same user interface following the same presentation and rules. The only difference that could have been caused by the user interface is a generally lower average rating of usefulness for both systems. This limitation could be faced by investigating alternative user interfaces that provide people with different ways to access personalised mobile content.

- *Stimulus:* Participants' judgement of usefulness was based on the provided stimulus handout consisting of a background scenario, a situation and a short task statement which also defines the limits of the results of this study. Task statements were provided in a very general and open form basically motivating participants to find one or more suitable events; a task described by [White, 2004, p. 151] as the "search for a number of items". This presumably caused participants to issue only very general search queries resulting in low information retrieval scores. Although the personalisation model was balanced between context and information retrieval, the experiment scores showed that the context score was on average 9 times larger than the average information retrieval score. This means the experiment, although in its outset balanced, was largely operating on context. This effect can be explained with the close-to-operational experiment conditions were largely uncontrolled participants decided to leave queries relatively generic. It would be worthwhile to investigate the personalisation model based on more specific search tasks as well as different search task types such as decision search tasks (e.g. deciding between two musicals based on a given situation) and background search tasks (e.g. getting an overview on salsa dance events in Aberdeen based on a situation). It would also be interesting to repeat this experiment with a range of different weights for the personalisation model (information retrieval and context scores) as indicated in section 5.4 in the previous chapter.

6.7.2 Search Effort

Although participants on average spent less time completing the tasks in the personalised system mode, the effect was not statistically significant (H6). Furthermore, the mean amount of queries submitted with the personalised system was smaller than with the non-personalised system but also not statistically significant (H7). This means, the results do not support the two hypothesis that suggest that users search effort is reduced when using a personalised search based on context. The following points discuss the two measures in more detail and review potential limitations:

- *Realism:* It is important to consider that results appeared under conditions of very little control with respect to how the tasks were completed. Every participant was allowed to decide freely on how many queries to submit, how much to navigate and to explore and which and how many event locations and events to choose and to rate. Also, the experiment procedure allowed participants to use the two systems differently when completing their tasks. When participants performed the same task on each of the two systems, they were able to pursue each task with a different depth and intensity if desired. Whether to solve the task simply with the first possible result or to search and browse the entire event content collection was entirely up to the participant. As part of this freedom, time was not restricted. Participants conducted the experiment in their own time similar to operational conditions. This also caused large variations between participants. One participant may have solved a task with only a single query whereas another may have issued dozens. Whereas one participant might have finished a task after viewing and rating a single event, another might have viewed and rated a large number of events from different searches and different event locations. This is the potential cause for the large variations for task time and the amount of queries per task. This created a more natural setting for information seeking supporting realism to a very high standard. This however could also be seen as a limitation that might have caused the indifference for the task time and the number of queries per task. As an alternative, it would have been possible to restrict the amount of available time artificially. This however would have changed the experiment setting into more laboratory boundaries and away from the intended open and holistic evaluation that was pursued with the mobile field experiment.

- *Search Behaviour:* One interesting result from this experiment was that participants' search behaviour showed close similarities with the search behaviour of users of web search engines based on a number of recent and large-scale studies. The average amount of queries for a task turned out to be similar to the average amount of queries used to complete a web search session. The average amount of query terms also closely followed the distribution of typical web search. This might indicate similarities between web search and searching on a mobile device with respect to search behaviour particularly if users are very familiar with search engines. It is also possible that these findings were influenced and limited by the choices that were made for the user interface and the hardware that was handed to participants. It is not unreasonable to assume that the query search box (see search view (2) in figure 6.9 on page 144) might have caused people to adapt their behaviour to normal web search. Also, the number of submitted query terms might have been additionally influenced by the size of this query box. Furthermore, the keyboard of the Sharp Zaurus provided people with a tool to type and construct queries relatively fast and direct. Although the data of this study cannot answer these questions, future studies could review the findings in this work by investigating different interfaces and different forms of interaction with mobile event information.

6.8 Summary

In this chapter, the personalisation model was evaluated with three hypotheses in a mobile experiment – usefulness as a measure of the event content quality, and task time and query number as a measure of users' search effort. Participants of the experiment performed situational search tasks and rated content about entertainment events. Results showed that context-aware personalisation was able to deliver about 20% more useful content to the mobile user compared with standard search. The study also found indications of search tasks being solved faster and with less queries when using personalisation, however without statistical significance. This was achieved despite the fact that the context model consisted of only 3 attributes and in almost natural experiment conditions. The results of this chapter show some initial evidence for the benefits of contextual personalisation in a simulated and largely uncontrolled experiment environment. The next chapter summarises the main contributions of this thesis followed by an account of potential limitations and suggestions for future work.

7 Conclusions

> The future's uncertain and the end is always near.
>
> — Jim Morrison, The Doors

This thesis presented an investigation into the effectiveness of context as a means to personalise content for users in mobile environments. Specifically, this research aimed to understand the role of personalisation and context, evaluate the effectiveness of context for content personalisation and investigate the event and map content domain for mobile usage.

During the development of this thesis, research on personalisation has changed considerably. The exclusive focus of personalisation on adaptive hypermedia and the web has been extended into mobile computing. This is an area where content personalisation is not only important but crucial, since mobile devices offer only limited ways to present content and their users have only a limited attention span. Like personalisation and other areas related to information science, context-aware computing also advanced significantly during the time of this research. Five years ago, context-aware computing was focused on more general discussions about context, the development of frameworks and prototypes and the establishment of initial studies. Now, the field has advanced and context modelling has emerged as a major direction in context-aware research. The past few years have also shown an increasing effort to apply context in other research fields such as information retrieval [Cool and Spink, 2002, Ingwersen et al., 2005], geographic information systems [Baus et al., 2005] and adaptive hypermedia (based on user modelling) [Jameson and Krueger, 2005] actively supported by a rising interest in mobile computing.

The conclusions of this thesis and its studies are provided as follows: Research contributions and their relation to the initial research questions are presented in the following section. This section also highlights the significance of each contribution within the wider field with implications for researchers and practitioners. A critical account of potential limitations is given in section 7.2 together with potential areas for future work.

7.1 Contributions

This thesis contributes in four areas based on the research questions presented in section 1.5 on page 7. Firstly, contributions are made towards a deeper understanding of personalisation from different perspectives that involve a range of research fields. Secondly, this work contributes with an example of a methodology for developing and evaluating context in the application of context-aware personalisation for mobile environments. Thirdly, this thesis provides a critical investigation of time, location and interest as context attributes commonly used in context models. Fourthly, causal attribution theory is linked with context modelling in an attempt to comprehend and model context from a psychological perspective.

Recalling from chapter 1, this thesis addresses three main research questions with several sub-questions –here, explicit answers are provided before presenting the contributions.

1. *Role of personalisation and context:* What is the role of content personalisation and user context?

   (a) *How does content personalisation relate to relevant research fields?*

   Personalisation is not yet a distinct research field but a set of techniques and methods that is used across different research areas. This thesis reviewed personalisation in adaptive hypermedia (including user modelling), context-aware computing, information retrieval and geographic information systems. Personalisation is applied in all these areas but hardly addressed with the same far-reaching scope. This thesis reviews Kobsa's and Brusilovsky's framework on personalisation and contributes a multidisciplinary and integrative extension for this framework for the benefit of both researchers and practitioners.

(b) *How do user context and personalisation relate to each other?*

Personalisation is traditionally based on user modelling but has recently been extended with contextual aspects. This thesis provided a review of both types of models and contributed an example for an application of a combined user and context model throughout this thesis.

2. *Effectiveness of context:* How effective is user context for content personalisation in the mobile event and map content domain?

(a) *How do selected context attributes - time, location and interest - influence users' perception of usefulness?*

All three attributes revealed significant effects, appear to have priorities and are highly interactive with respect to usefulness. Most importantly, however, it was shown that context has a large potential to dynamically influence users perception of usefulness.

(b) *How can user context be applied for a personalised information system?*

In this thesis, a context model was developed by connecting data about contextual effects on usefulness to attribution theory. Attribution theory provides a possibility for explaining contextual reasoning in humans and relating it to basic statistical models. A personalisation model was created by combining a standard information retrieval score with a context score and using it in a personalised information system on a mobile device.

(c) *How effective is user context in a personalised information system for providing useful content?*

The context-aware personalisation model was integrated in a mobile information system and evaluated with a mobile experiment. On average, the model provided people with event content that was about 20% more useful than content delivered with a standard information retrieval system.

3. *Domain investigation:* What are the possibilities and limitations of the event and map content domain with respect to mobile use?

   (a) *What are the specific characteristics of mobile use?*

   (b) *What are the specific characteristics of the event and map content domain?*

   Both questions (a and b above) were largely explored with data that was provided by the AmbieSense project based on active participation and involvement (see chapter 3). Although results from this data do not represent a contribution of this thesis, it nevertheless shaped and influenced this work. Results from the AmbieSense data revealed that mobile users generally want information in context. It showed that users in mobile situations have preferences for particular types of content. Mobile users have a demand for event and map content (among other types) and have distinct preferences about content delivery and consumption. Mobile users are willing to provide personal information (especially about personal interests and hobbies) in order to benefit from personalised information services.

The first contribution (see section 7.1.1) relates to the research question about the role of personalisation and context. The remaining contributions (see section 7.1.2 - 7.1.4) connect to the second research question about the effectiveness of context; in particular how selected context attributes influence people, how context can be applied in a personalised information system and how effective such a system can become in providing useful content to its users. The contribution of a strategy for the development and evaluation of a context model (see section 7.1.2) also relates to the third research question about the characteristics of mobile use as well as the event and map content domain; connecting AmbieSense results on mobile use (as presented in chapter 3) with the two user studies (as presented in chapter 4 and 6). Both user studies were focused on mobile use as well as the event and map content domain.

7.1.1 Understanding Personalisation from Different Angles

The first research question asks about the role of content personalisation in relation to relevant research fields as well as the relation between personalisation and context.

Throughout this thesis, it was discovered that personalisation is not yet a separate research field. Instead, it is mostly treated as a concept that is operationalised with a set of methods and tools, as shown in the review of related work in section 2.4 on page 35. These methods and tools relate to many different research areas. Some of these research areas have been associated with information science – in particular information retrieval [Saracevic, 1999], geographic information systems [Goodchild, 1992], and, more recently, adaptive hypermedia [Aroyo et al., 2004]. Whereas the field of context-awareness is usually associated with computer science, context in general is a central topic in information science [Ingwersen and Järvelin, 2005] and all mentioned research areas share the most central focus of information science, that of providing people with effective methods and tools to manage increasing amounts of information. This multidisciplinary background is the main reason why personalisation is generally difficult to grasp. It was identified that more research is needed to better understand and overview the basic concepts of personalisation and analyse methods and tools from this multidisciplinary viewpoint. This thesis has addressed the following two issues:

- *Multidisciplinary personalisation:* This contribution provides a multidisciplinary overview on personalisation for both researchers and practitioners (e.g. mobile application developers) in chapter 2. It reconfirms the framework on personalisation produced by Kobsa (e.g. [Kobsa et al., 2001]) and Brusilovsky (e.g. [Brusilovsky, 1996, Brusilovsky, 2001]) which were both very focused on adaptive hypermedia and user modelling. The overview created in this thesis enhances their views to a much wider range of research fields, in particular
  - Adaptive hypermedia (including user modelling)
  - Context-aware computing
  - Information retrieval
  - Geographic information systems

and adjusts and enriches them with various new aspects (e.g. the consideration of context and mobile usage). All these research areas connect to information science as they share its central challenge - to effectively manage and increase the accessibility of the growing volume of available information. Furthermore, the work promotes an integrative view of personalisation spanning across these fields and delivering a potential structure to organise research from these areas with respect to personalisation.

- *User Models and Context:* This integrative view is also applied for the modelling aspect of personalisation. Both user and context models are compared and their communalities and differences are highlighted (see section 2.2 on page 16) based on a more general recognition that both types of models are important for the delivery of personalised services; this is for example highlighted in a special issue on user modeling in ubiquitous computing [Jameson and Krueger, 2005] and expressed in the fact that context models in context-aware computing now include user modelling aspects [Göker and Myrhaug, 2002, Tazari et al., 2004]. The review of related work in chapter 2 in particular highlighted the developments in personalisation research in the light of user and context modelling for the acquisition, the modelling and the creation of personalised output.

This overview allows researchers and practitioners to adopt a more interdisciplinary view on personalisation that includes an array of research fields and application areas. Researchers can apply this overview to position their own work and relate their research with others across different fields. Practitioners (e.g. developers of mobile applications and services) can use it to gain a better oversight about the various stages and processes involved in personalisation from a multitude of different application areas supported by various examples. This allows them to position their own products within a wider area and review their applications with the various aspects of personalisation that have been highlighted. This may help them to identify important aspects of personalisation they would not consider otherwise and inspire future product improvements to the benefit of both businesses and end-users.

7.1.2 Strategy for the Development and Evaluation of a Context Model

This thesis presents an example of the step-by-step development and evaluation of a context model that is used for content personalisation by personalised search through the re-ranking of search results. This is based on the research question on the effectiveness of context; in particular the question of how user context can be applied for a personalised information system and how effective it can provide useful content to users. The model is focused on entertainment events in a mobile application environment that encourages the use of geographic maps. The model therefore also highlights the characteristics of mobile use as well as the event and map content domain expressed in the third research question. Specifically, it is demonstrated how a content and usage domain is initially explored based on relevant results from related studies, how this results are then used to specify a context model that is then analysed in a laboratory study, formalised into a personalisation model and then verified in a simulated mobile field experiment. In the following, each of this steps is highlighted in more detail:

- *Exploration of relevant AmbieSense results:* In a first step, results from a range of relevant AmbieSense user studies provided some useful insights about travellers and tourists as one type of mobile user. Results from these studies were selected with respect to the question of what types of content mobile users expect, how people acquire, access and use content and how willing (mobile) users are in providing personal information.

- *Context model specification and analysis:* As a second step, a more generic context model, described in [Myrhaug and Göker, 2003], was further specialised by consideration of content and usage based on the AmbieSense results reported in chapter 3. The generic context model divides context into the five groups of environment context, personal context, task context, social context and spatio/temporal context. The more specialised context model used in this thesis however focused on three attributes selected from two of these categories; time and location (from spatio/temporal context) and interest (from personal context). A laboratory experiment measured the influence of the three chosen context attributes on individual's perception of usefulness (understood as situational relevance) based on a simulated mobile scenario. The findings provide empirical evidence for the complexity of context, the strength of individual attributes, their interactive strength and the priorities of importance that emerged from that.

- *Formalisation of context-aware personalisation:* The next step applied the results from the laboratory experiment. Findings where related to some of the theory of causal attribution and then used to create a predictive model of context. Causal attribution theory describes the human process of explanation finding and investigates how people relate effects to their potential (contextual) causes. This theory was highlighted based on Kelley's covariation principle, that links human explanation with factorial ANOVA, and causal schematas that connect attribution with a basic form of regression modelling. Regression was applied to develop a predictive context model where a score expresses the amount of usefulness (situational relevance) based on time, location and interest. Within a personalisation model, this context score was combined with an information retrieval score that allows the model to be used for personalised search.

- *Verification of context-aware personalisation:* In the final step, a mobile field experiment investigated the effect of the personalisation model that was described above. A mobile application was designed and equipped with the model to allow for personalised search. Users applied the system to search for event content based on simulated work tasks. The mobile experiment evaluated the effect of context on usefulness holistically in combination with an information retrieval system; a refined vector space model equivalent that has been described in section 2.4. Results showed that context-aware personalisation was able to deliver about 20% more useful content to mobile users. Furthermore, results also indicated that search tasks were solved faster with less queries, however without statistical significance.

This contribution provides a detailed example of a possible process for the creation of a personalised information system that applies contextual information within a context model. It delivers a strategy for researchers that demonstrates how user studies can be combined from both laboratory and field environments. As described in [Kjeldskov and Graham, 2003] and [Scholtz, 2006], evaluation of mobile information systems are still sparse and largely limited in realism. Combined studies are still far too few in many areas of information science (e.g. information retrieval) given that its intended focus is on the users of information systems. Researchers may apply and adapt this strategy as a template for investigating and evaluating context for other areas of information behaviour (e.g. based on specific user groups and tasks). Practitioners may use and adapt this template for investigating and evaluating known or hypothetical contextual conditions that are targeted by current and future products.

7.1.3 Time, Location and Interest - Evaluation of Contextual Relationships

Another important research question is how selected context attributes influence users' perception of usefulness. The experiment presented in chapter 4 investigated a context model for personalised information retrieval of entertainment events in a mobile application environment. Literature to date does not show evidence of such a context model being established and evaluated in this depth for the event content domain. The context model consisted of three context attributes (time, location and user's interest); attributes frequently used for context-aware systems which have not been subject to closer investigation. It is argued that more knowledge needs to be gathered about the actual effect of contextual attributes upon which a system is built. Also, it can be highly valuable to know how such attributes interact in conjunction. This thesis provides such an investigation for these three attributes. Results are summarised as follows:

- *General context validation:* All three context attributes are important for context-aware personalisation of event content based on how they affected users' perception of usefulness.
- *Context priority:* A priority between the three context attributes was discovered. It appears that interest is the most influential contextual attribute, followed by time and then location. This is particularly interesting since location is generally regarded as a very strong attribute, particularly in location-based systems research.
- *Non-linearity of time:* It was discovered that the time context attribute had a non-linear effect on peoples' perception of usefulness. This means, people responded stronger to temporal information the closer the proximity to a relevant time (i.e. the performance time of an event). Specifically, the structure of the time effect indicated the existence of a power law relationship.

- *Context interactions:* Time, location and interest revealed some strong and complex interactions. This indicates that even small context models with few attributes have complex interrelationships that should be considered when building context-aware information systems.

This contribution more generally aims to make researchers and practitioners aware of the dynamics and interconnected effects that particular contextual attributes have on people. Practitioners, more specifically, can use the results from this experiment as an initially (learned) context model for an application, assuming it fits the intended type of user and content (i.e. mobile users of entertainment information[1]). The application could then refine this default model as the user provides more explicit or implicit feedback to the system. This would allow an application to deliver basic personalised search results even if no feedback has been collected from the user while allowing to adjust the performance later on.

Furthermore, researchers can apply and adapt the methodology that was used for the laboratory experiment to construct similar studies to investigate other context models. This addresses the claim in [Kjeldskov and Graham, 2003] and more generally what was described in section 6.2.2 on page 129 that highlights the lack of evaluation for mobile applications by providing one methodology for testing context model in advance[2]. Similarly, practitioners can use the methodology to investigate particular contextual effects on users in order to apply them more effectively in their context-aware products and services. Initial investigations on hypothetical context models can be performed before any system is even designed and implemented. Findings from such investigations can be used to shape requirements and inform design and implementation of new systems.

7.1.4 Connection between Attribution Theory and Context

This thesis revealed links between the theory of causal attribution and the process of context modelling. In particular, it was demonstrated that causal attribution as a theoretical framework allows to view context modelling as a human process of finding

[1] It is certainly possible that results may also apply to more general types of content and user populations.

[2] Note that section 6.2.2 also makes a strong case for extending evaluation to the field which is addressed in the contribution described in section 7.1.2.

explanations. This relates to the second research question on the effectiveness of context; in particular to the question of how selected context attributes influence people's perception on usefulness. This was addressed by providing a possible theory that can explain how people model context themselves. The theory of attribution was highlighted from the viewpoint of Harold H. Kelley [Kelley, 1973], who significantly enriched the theory with key models and tools. The process of linking causal attribution with context was developed in several steps in chapter 5:

- *Overview:* The two key elements of Kelley's causal attribution were introduced and summarized - the covariation principle (expressed as Kelley's cube) and the causal schema as a more specialised template based on the covariation principle. These two concepts were described and visualised.

- *Relating context with attribution theory:* The process of context modelling was related to attribution theory based on four different arguments:

  - A causal schema is a person's mental model that describes how that person relates an effect with possible situational (i.e. contextual) attributes. In this respect, Kelley's cube was discussed as one example of a context model. As such, the cube also addresses attributes that are common to context models (e.g. temporal attributes).
  - Both attribution theory and context modelling distinguish between external and internal attributes. For context, this division has been reviewed comprehensively in chapter 2. Internal attributions map to user models and external attributions relate more closely to context models.
  - Both causal schemata and context models tend to be focused on only a few attributes. Both types of research are focused on researching these attributes.
  - Attribution theory discovered a range of simple causal patterns in humans with respect to how inferences are developed. Similarly, research in context modelling has also started to put forward considerable effort for a deeper understanding of contextual effects.

- *Context as causal schema:* The context model, as developed and evaluated in chapter 4, was expressed and visualised as a causal schema. The three attributes of the context model were presented as three dimensions of a causal schema along their various levels. The model was not only related to the causal schema but also to Kelley's cube and the covariation principle.

- *Causal schematas and regression:* A link suggesting a connection between causal schemata and multiple regression was discovered from attribution literature. This link was used to describe the context model using multiple regression based on the data set collected during the laboratory experiment reported in chapter 4. The regression model expresses the quantitative strength of the dimensions in the causal schema (impact of time, location and interest plus their interactions) in relation to the effect (users' perception of usefulness). The functional representation of the regression model is then linked with the cube diagram to show this relationship.
- *Coexisting causal schemata:* The possibility of data representing different causal schemata is discussed. This means the regression data might represent different kinds of causal schemata, each representing one mental model for assigning usefulness for contextual situations. It is suggested that future work can look more comprehensively into the effect of coexisting causal schemata with respect to context modelling.

This contribution provides an initial stepping stone for a context theory that is potentially useful to all research fields that apply contextual information. This is particularly the case in the area of context-aware computing. After many years of close focus on the creation and evaluation of context-aware systems and frameworks, as reviewed in [Baldauf et al., 2007], context-aware computing has only recently started extending its efforts towards context modelling [Indulska and Roure, 2004]. Information science, on the other hand, has already developed an initial tradition by integrating contextual information to enable more user-centred forms for information interaction and developing essential theory for explaining and modelling context. This is summarised in [Ingwersen and Järvelin, 2005] and conferences such as CoLIS [Crestani and Ruthven, 2005] and IIiX [Ruthven et al., 2006] represent continued efforts in this direction. Despite all these efforts, there is still a significant gap between the amount of empirical research that investigates context and the amount of solid theory that is able to explain contextual effects that is discovered during such investigations. This contribution offers one possible direction for explaining contextual reasoning in humans from a psychological perspective thus connecting context closer with one of the more profound and more established causal theories. This has the potential to contribute toward creating stronger theoretical constructs for context for a better and more solid understanding of context but also, more practically, for creating more effective personalised applications.

7.2 Limitations and Future Work

The research presented in this thesis has a number of limitations. These are discussed in the following points with suggestions for future work:

- *User Population:* As a limitation of the AmbieSense results presented in chapter 3, gender distributions were not always completely balanced and some gender effect might exist in these data sets [Myrhaug et al., 2004b]. The studies conducted in chapter 4 and chapter 6 were balanced. Both the AmbieSense studies and the studies conducted for this thesis mostly recruited people between 18 and 29 years old. This means that results presented in this thesis are potentially more expressive with respect to these age groups and less expressive for others. Future studies could further investigate other age groups to verify if results generalise. Furthermore, the AmbieSense user population were travellers and tourists whereas the two user studies in this thesis where collected mostly from students and university staff. Although this could be viewed as a limitation, it is not unreasonable to assume that travellers and tourists share certain personal characteristics with general mobile users and therefore provide a reasonable population for initial investigations. Consistency between the two studies of this thesis is maintained since both studies recruited the same type of participants. Note that external validity of the results from the two experiments is naturally limited due to the relatively small numbers of participants (32 for the experiment on contextual usefulness reported in chapter 4 and 17 for the mobile experiment described in chapter 6). It would be worthwhile to repeat the experiments presented in this thesis to collect more data from an even wider and larger sample of people. This would allow to verify and strengthen results to be generalised with greater confidence beyond the basic threshold of statistical significance.

- *Methodology and Stimulus:* Both the interest and the location attribute in chapter 4 might be underestimated based on the simulated nature of the experiment. The interest attribute might have had an even stronger effect on usefulness if participants would have been sampled from particular interest groups (e.g. members of a comedy or jazz club). Likewise, the location attribute might have revealed stronger effects if people would have performed their task with a real map; this approach was taken in the mobile user experiment in chapter 6. One limitation in the design of the mobile user study is that each task was performed twice. This caused a slight variation in usefulness ratings between participants depending in which order they used the two systems, however, the overall positive effect of the personalised system remained stable. Furthermore, participants from the mobile experiment often submitted only very general search queries which caused the search system not being used to its full potential. This can be explained with the close-to-operational experiment conditions with open and unrestricted tasks that caused participants to leave queries relatively generic. For this reason, it would be interesting to repeat the experiment with more specific and different types of tasks in order to obtain a wider spectrum of search behaviours. It is also not certain how well the stimulus actually modelled realistic task situations for participants – despite the positive feedback that was provided in the post-questionnaire of the mobile experiment (see 6.6.4 on page 157). Future work should repeat experiments in a real festival scenario using real events based on users real information needs and current contextual situations.

- *Mobile Application:* The application that was used in the mobile experiment provided participants search results through a geographic interface rather than a ranked result list and results were only presented with their titles. This might have effected participants' overall performance in retrieving useful events. However, the difference in system performance should be unaffected since both systems used the same interface. Future work could nevertheless investigate different user interfaces that provide people with alternative ways to access personalised mobile content.

- *Context and Personalisation Model:* The following limitations apply to the personalisation model and its embedded context model:

- *Focused Context Modelling:* The context model that was used for personalisation was based on three contextual attributes expressed by a small set of distinct attribute levels; a necessary limitation to allow for a full investigation of the model based on user judgements. Too large a set of attributes and attribute levels would easily become overly complex and would have compromised the expressiveness of the results in chapter 4. Nevertheless, it would be worthwhile to evaluate a wider range of attribute levels for time, location and interest and also evaluate different context attributes in future studies. This would be helpful for generally validating the results of this thesis and for extending this research to other contextual aspects.
- *Static Context Modelling:* The context model that was developed and verified in this thesis is static and was established on data that has been gathered by a user study within a short period of time (see chapter 4). In real and operational conditions, a personalised information system would update its context model continuously and dynamically based on new data. The existing context model nevertheless could still be applied as a default and updated based on explicit (e.g. ratings on contextual results) or implicit (e.g. selection or viewing times of events) feedback. This feedback could, for example, be used to adjust the regression model over time and adapt it further to the individual user.
- *Personalisation Model Weighting:* The personalisation model that was presented in chapter 5 and evaluated in chapter 6 implemented a balanced relationship (i.e. equal weights) for a context score and an information retrieval score. Since the precise effect for different weights is unknown, it was a sensible first approximation that consequently resulted in about 20% improvement in the usefulness of retrieved event content based on users' judgement. It would be interesting to explore further the parameter space by using different weights between the two scores and measure their effects in a series of additional comparative user studies.

• *Economic Validity:* The experiments presented in this thesis are based on the collection of a considerable amount of data from participants upon which the model was created and later evaluated. It could be argued that this strategy might not be economic when repeated, for example, for a mobile application that targets another content domain (e.g. a personalised news or shopping guide). However,different techniques can be used to both ease and accelerate data collection. The use of explicit relevance feedback, for example, can be used either with initial questionnaires or with ongoing ratings for viewed content. Implicit relevance feedback techniques can be used to collect data based on user behaviour (e.g. when selecting or viewing content). Collaborative methods can help to collect data from distributed mobile devices, creating the model centrally and distributing it back to each mobile application for individual use. Section 2.4 on page 35 discusses implicit and explicit personalisation techniques as well as collaborative methods for potential solutions for an easier and faster data collection.

Although bound to a number of limitations, the studies presented in this thesis provide evidence for the promising potential of context to facilitate personalised information delivery for mobile users – an area that will experience much development in the years to come. Overall, the work serves as an example of an investigation into context from multiple angles, using multiple experiment methods and statistical techniques. Also, the thesis carefully links to some of the theoretical aspects of psychology as a potential source for a deeper understanding of contextual processes in humans.

Bibliography

The secret to creativity is knowing
how to hide your sources.

ALBERT EINSTEIN

[Abowd et al., 1997] Abowd, G. D., Atkeson, C. G., Hong, J., Long, S., Kooper, R., and Pinkerton, M. (1997). Cyberguide: A mobile context-aware tour guide. *Wireless Networks*, 3(5):421–423.

[Aroyo et al., 2004] Aroyo, L., Bra, P. D., Houben, G.-J., and Vdovjak, R. (2004). Embedding information retrieval in adaptive hypermedia: IR meets AHA! *Hypermedia*, 10(1):53–76.

[Bahl and Padmanabhan, 2000] Bahl, P. and Padmanabhan, V. N. (2000). RADAR: An In-Building RF-based User Location and Tracking System. In *19th Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom)*, pages 775–784, Tel Aviv, Israel. IEEE Press.

[Balabanovic and Shoham, 1997] Balabanovic, M. and Shoham, Y. (1997). Fab: Content-Based, Collaborative Recommendation. *Communications of the ACM*, 40(3):66–72.

[Baldauf et al., 2007] Baldauf, M., Dustdar, S., and Rosenberg, F. (2007). A Survey on Context-Aware Systems. *International Journal of Ad Hoc and Ubiquitous Computing*, 2(4):263–277.

[Baus et al., 2005] Baus, J., Cheverst, K., and Kray, C. (2005). A Survey of Map-Based Mobile Guides. In Meng, L., Zipf, A., and Reichenbacher, T., editors, *Map-Based Mobile Services: Theories, Methods and Implementations*, pages 197–216. Springer Verlag, Berlin.

[Beigl, 2002] Beigl, M. (2002). Special Issue on Location Modeling in Ubiquitous Computing. *Personal and Ubiquitous Computing*, 6(5-6):311–357.

[Belkin, 1996] Belkin, N. (1996). Intelligent information retrieval: Whose intelligence? In *Fifth International Symposium for Information Science (ISI)*, pages 25–31, Konstanz, Germany. Universtaetsverlag Konstanz.

[Belkin et al., 1995] Belkin, N., Cool, C., Stein, A., and Thiel, U. (1995). Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems. *Expert Systems with Applications*, 9(3):379–395.

[Belkin and Croft, 1992] Belkin, N. and Croft, B. (1992). Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(2):29–38.

[Belkin et al., 1982] Belkin, N., Oddy, R. N., and Brooks, H. M. (1982). Ask for Information Retrieval: Part 1. Background and Theory. *Journal of Documentation*, 38(2):61–71.

[Bierig and Göker, 2006] Bierig, R. and Göker, A. (2006). Time, Location, and Interest: An Empirical and User-Centred Study. In *First Symposium on Information Interaction in Context (IIiX)*, pages 79–87, Copenhagen, Denmark. ACM-Press.

[Billsus and Pazzani, 2000] Billsus, D. and Pazzani, M. (2000). User Modeling for Adapative News Access. *User Modeling and User-Adapted Interaction*, 10(2-3):147–180.

[Borlund, 2000] Borlund, P. (2000). Experimental Components for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 56(1):71–90.

[Borlund, 2003a] Borlund, P. (2003a). The Concept of Relevance in IR. *Journal of the American Society for Information Science*, 54(10):913–925.

[Borlund, 2003b] Borlund, P. (2003b). The IIR Evaluation Model: A Framework for Evaluation of Interactive Information Retrieval Systems. In *Information Research*, volume 8. number 3, paper no. 152. [Available at: http://informationr.net/ir/8-3/paper152.html].

[Borlund and Ingwersen, 1997] Borlund, P. and Ingwersen, P. (1997). The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 53(3):225–250.

[Borlund et al., 2008] Borlund, P., Schneider, J. W., Lalmas, M., Tombros, A., Feather, J., Kelly, D., de Vries, A., and Azzopardi, L. (2008). Second Symposium on Information Interaction in Context (IIiX). London, UK. ACM Press.

[Bradley et al., 2000] Bradley, K., Rafter, R., and Smyth, B. (2000). Case-Based User Profiling for Content Personalisation. In Brusilovsky, P., Stock, O., and Strapparava, C., editors, *First International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH)*, pages 62–72, Trento, Italy. Springer Verlag.

[Bradley and Smyth, 2002] Bradley, K. and Smyth, B. (2002). Information Ordering vs Information Filtering. In *6th European Conference on Case Based Reasoning (ECCBR) Workshops*, pages 13–14, Aberdeen, UK.

[Bradley and Dunlop, 2004] Bradley, N. A. and Dunlop, M. D. (2004). Towards a User-Centric and Multidisciplinary Framework for Designing Context-aware Applications. In *6th International Conference on Ubiquitous Computing (UbiComp), 1st International Workshop on Advanced Context Modelling, Reasoning And Management*, Nottingham, UK.

[Brusilovsky, 1996] Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, 6(2-3):87–129.

[Brusilovsky, 2001] Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction (Ten Year Anniversary Issue)*, 11(1-2):87–110.

[Brusilovsky et al., 2005] Brusilovsky, P., Callaway, C., and (Eds.), A. N. (2005). 1st International Workshop on New Technologies for Personalized Information Access. In *10th International Conference on User Modeling (UM)*, Edinburgh, UK.

[Brusilovsky et al., 2003] Brusilovsky, P., Corbett, A., and de Ross, F. (2003). Preface of Proceedings. In *9th International Conference on User Modeling (UM)*, Johnstown, PA, USA. Springer Verlag.

[Buckley, 1993] Buckley, C. (1993). The Importance of Proper Weighting Methods. In *Workshop on Human Language Technology (HLT)*, pages 349–352, Princeton, NJ, USA. Association of Computational Linguistics.

[Bush, 1945] Bush, V. (1945). As We May Think. *Atlantic Monthly*, 176(1):101–108.

[Byun and Cheverst, 2001] Byun, H. E. and Cheverst, K. (2001). Exploiting User Models and Context-Awareness to Support Personal Daily Activities. In *8th International Conference on User Modeling (UM), Workshop on Workshop on User Modeling for Context-Aware Applications*, Sonthofen, Germany.

[Cassel and Wolz, 2001] Cassel, L. N. and Wolz, U. (2001). Client Side Personalization. In *Joint DELOS-NSF Workshop on Personalisation and Recommender Systems in Digital Libraries*, Dublin City University, Ireland. European Research Consortium for Informatics and Mathematics (ERCIM).

[Chalmers, 2002] Chalmers, D. (2002). *Contextual Mediation to Support Ubiquitous Computing*. Phd thesis, Imperial College of Science, Technology and Medicine, University of London.

[Chalmers and Sloman, 1999] Chalmers, D. and Sloman, M. (1999). QoS and Context Awareness for Mobile Computing. In *1st International Symposium on Handheld and Ubiquitous Computing*, pages 380–382, Karlsruhe, Germany. Springer Verlag.

[Chen and Finin, 2003] Chen, H. and Finin, T. (2003). An Ontology for Context-Aware Pervasive Computing Environments. In *18th International Joint Conference on Artificial Intelligence (IJCAI), Workshop on Ontologies and Distributed Systems*, Apaculco, Mexico.

[Chen and Sycara, 1998] Chen, L. and Sycara, K. (1998). WebMate: A Personal Agent for Browsing and Searching. In *2nd International Conference on Autonomous Agents*, pages 132–139. ACM Press.

[Cheverst et al., 2000] Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efstratiou, C. (2000). Developing a Context-Aware Electronic Tour Guide: Some Issues and Experiences. In *Conference on Human Factors in Computing Systems (CHI)*, pages 17–24, The Hague, The Netherlands. ACM Press.

[Chisholm and Kolda, 1999] Chisholm, E. and Kolda, T. G. (1999). New Term Weighting Formulas for the Vector Space Method in Information Retrieval. Technical Report ORNL-TM-13756, Oak Ridge National Laboratory.

[Cingil et al., 2000] Cingil, I., Dogac, A., and Azgin, A. (2000). A broader approach to personalization. *Communications of the ACM*, 43(8):136–141.

[Cleverdon and Keen, 1966] Cleverdon, C. W. and Keen, E. M. (1966). Factors Determining the Performance of Indexing Systems. Technical report, Aslib Cranfield Research Project.

[Cohen and Cohen, 1975] Cohen, J. and Cohen, P. (1975). *Applied Multiple Regression/-Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum, Hillsdale, NJ, USA.

[Cool and Spink, 2002] Cool, C. and Spink, A. (2002). Special Issue on context in information retrieval. *Information Processing and Management*, 38(5):605–611.

[Crestani and Ruthven, 2005] Crestani, F. and Ruthven, I. (2005). Context: Nature, Impact, and Role. In *5th International Conference on Conceptions of Library and Information Sciences (CoLIS 2005)*, page 250 p., Glasgow, UK. Springer Verlag.

[Cunningham and Kelley, 1975] Cunningham, J. D. and Kelley, H. H. (1975). Causal Attributions for Interpersonal Events of Varying Magnitude. *Journal of Personality*, 43(1):74–93.

[Dey, 2001] Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5(1):4–7.

[Dunlop et al., 2004] Dunlop, M. D., Ptaskinski, P., Morrison, A., McCallum, S., Risbey, C., and Stewart, F. (2004). Design and development of Taeneb City Guide - from paper maps and guidebooks to electronic guides. In *IFITT's Global Travel & Tourism Technology and eBusiness Forum (ENTER)*, Cairo, Egypt. Springer Verlag.

[Fahy and Clarke, 2004] Fahy, P. and Clarke, S. (2004). CASS - Middleware for Mobile Context-Aware Applications. In *2nd International Conference on Mobile Systems, Applications, and Services (MobiSys 2004), Workshop on Context-Awareness*, Bosten, MA, USA.

[Faloutsos and Oard, 1995] Faloutsos, C. and Oard, D. W. (1995). A Survey of Information Retrieval and Filtering Methods. Technical Report CS-TR-3514, University of Maryland, College Park, US.

[Ferro and Potorti, 2005] Ferro, E. and Potorti, F. (2005). Bluetooth and Wi-Fi Wireless Protocols: A Survey and a Comparison. *IEEE Wireless Communications*, 12(1):12–26.

[Field, 2005] Field, A. (2005). *Discovering Statistics Using SPSS.* Sage Publications, London, 2nd edition.

[Fink et al., 1998] Fink, J., Kobsa, A., and Nill, A. (1998). Adaptable and Adaptive Information Provision for All Users, Including Disabled and Eldery People. *The New Review of Hypermedia and Multimedia*, 4:163–188.

[Frieze and Weiner, 1971] Frieze, I. and Weiner, B. (1971). Cue Utilization and Attributional Judgements for Success and Failure. *Journal of Personality*, 39(4):591–605.

[Gellersen et al., 2004] Gellersen, H., Kortuem, G., Schmidt, A., and Beigl, M. (2004). Physical Prototyping with Smart-Its. *IEEE Pervasive Computing*, 3(3):74–82.

[Göker, 1994] Göker, A. (1994). *An Investigation into the Application of Machine Learning in Information Retrieval.* PhD thesis, City University.

[Göker et al., 2004] Göker, A., Cumming, H., and Myrhaug, H. I. (2004). Content Retrieval and Mobile Users: An Outdoor Investigation of an Ambient Travel Guide. In *6th International Conference on Human Computer Interaction with Mobile Devices and Services (Mobile HCI), 2nd International Workshop on Mobile and Ubiquitous Information Access*, Glasgow, UK.

[Göker and Myrhaug, 2002] Göker, A. and Myrhaug, H. (2002). User Context and Personalisation. In *6th European Conference on Case Based Reasoning (ECCBR) Workshops*, Aberdeen, UK.

[Göker and Myrhaug, 2007] Göker, A. and Myrhaug, H. I. (2007). Evaluation of a Mobile Information System in Context. *Information Processing and Management*, 44(1):39–65.

[Goldberg et al., 1992] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70.

[Goodchild, 1992] Goodchild, M. F. (1992). Geographic Information Science. *International Journal of Geographical Information Systems*, 6(1):31–45.

[Griffith and O'Riordan, 2000] Griffith, J. and O'Riordan, C. (2000). Collaborative Filtering. Technical report NUIG-IT-160900, Department of Information Technology, National University of Ireland.

[Griffith and O'Riordan, 2002] Griffith, J. and O'Riordan, C. (2002). Non-Traditional Collaborative Filtering Techniques. Technical Report NUIG-IT-121002, Department of Information Technology, National University of Ireland.

[Grossman and Frieder, 2004] Grossman, D. A. and Frieder, O. (2004). *Information Retrieval - Algorithms and Heuristics*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Boston, 2nd edition.

[Gu et al., 2004] Gu, T., Pung, H. K., and Zhang, D. Q. (2004). A Middleware for Building Context-Aware Mobile Services. In *59th IEEE Semiannual Vehicular Technology Conference (VTC)*, Milan, Italy. IEEE Computer Society Press.

[Hallberg et al., 2002] Hallberg, J., Nilsson, M., and Synnes, K. (2002). Bluetooth Positioning. In *3rd Annual Conference on Computer Science and Electrical Engineering (CSEE 2002)*, Luleå, Sweden.

[Harmon and Anderson, 2003] Harmon, J. E. and Anderson, S. J. (2003). *The Design and Implementation of Geographic Information Systems*. Wiley, Hoboken, New Jersey, USA, hardcover edition.

[Harrison and List, 2004] Harrison, G. W. and List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42(4):1009–1055.

[Hatcher and Gospodnetic', 2004] Hatcher, E. and Gospodnetic', O. (2004). *Lucene in Action*. Manning Publications, Greenwich, CT, USA.

[Heider, 1958] Heider, F. (1958). *The Psychology of Interpersonal Relations*. Wiley, New York.

[Hesketh, 1984] Hesketh, B. (1984). Attribution Theory and Unemployment: Kelley's Covariation Model, Self-Esteem, and Locus of Control. *Journal of Vocational Behavior*, 24(1):94–109.

[Hewstone, 1989] Hewstone, M. (1989). *Causal Attribution: From Cognitive Processes to Collective Beliefs*. Blackwell Publishers Ltd., Oxford, UK.

[Hofer et al., 2002] Hofer, T., Schwinger, W., Pichler, M., Leonhartsberger, G., Altmann, J., and Retschitzegger, W. (2002). Context-Awareness on Mobile Devices - The Hydrogen Approach. In *36th Annual Hawaii International Conference on System Sciences (HICSS)*, pages 292–302, Hawaii, USA. IEEE Computer Society Press.

[Hull et al., 1997] Hull, R., Neaves, P., and Bedford-Roberts, J. (1997). Towards Situated Computing. In *1st International Symposium on Wearable Computers*, pages 146–153, Cambridge, MA, USA. IEEE Computer Society Press.

[Indulska and Roure, 2004] Indulska, J. and Roure, D. D. (2004). Workshop on Advanced Context Modelling, Reasoning And Management. In *6th International Conference on Ubiquitous Computing (UbiComp)*, Nottingham, UK.

[Ingwersen, 1992] Ingwersen, P. (1992). *Information Retrieval Interaction*. Taylor Graham, London.

[Ingwersen, 1996] Ingwersen, P. (1996). Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation*, 52(1):3–50.

[Ingwersen et al., 2005] Ingwersen, P., Jaervelin, K., and Belkin, N. (2005). Workshop on Information Retrieval in Context. In *28th Annual International ACM SIGIR Conference* *on Research and Development in Information Retrieval*, Salvador, Brazil. Royal School of Library and Information Science, Copenhagen, Denmark.

[Ingwersen and Järvelin, 2005] Ingwersen, P. and Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. The Information Retrieval Series. Springer Verlag.

[Ingwersen et al., 2004] Ingwersen, P., van Rijsbergen, K., and Belkin, N. (2004). Workshop on Information Retrieval in Context (IRiX). In *27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Sheffield, UK.

[Jameson, 2001] Jameson, A. (2001). Modelling both the Context and the User. *Personal and Ubiquitous Computing*, 5(1):29–33.

[Jameson and Krueger, 2005] Jameson, A. and Krueger, A. (2005). Special Issue on User Modeling in Ubiquitous Computing. *User Modeling and User-Adapted Interaction*, pages 193–338.

[Jansen and Spink, 2005] Jansen, B. J. and Spink, A. (2005). An analysis of Web searching by European AlltheWeb.com users. *Information Processing and Management*, 41(2):361–381.

[Jansen et al., 2000] Jansen, B. J., Spink, A., and Saracevic, T. (2000). Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. *Information Processing and Management*, 36(2):207–227.

[Jiang and Yao, 2007] Jiang, B. and Yao, X. (2007). Location Based Services and GIS in Perspective. In Gartner, G., Cartwright, W., and Peterson, M. P., editors, *Lecture Notes in Geoinformation and Cartography*, pages 27–45. Springer Verlag, Berlin.

[Joachims et al., 1997] Joachims, T., Freitag, D., and Mitchell, T. M. (1997). WebWatcher: A Tour Guide for the World Wide Web. In *Fifteenth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 770–777, Nagoya, Aichi, Japan.

[Kaenampornpan and ONeill, 2004] Kaenampornpan, M. and ONeill, E. (2004). An Integrated Context Model: Bringing Activity to Context. In *6th International*

Conference on Ubiquitous Computing (UbiComp), 1st International Workshop on Advanced Context Modelling, Reasoning And Management, Nottingham, UK.

[Kay, 1995] Kay, J. (1995). The um Toolkit for Reuseable, Long Term User Models. *User Modeling and User-Adapted Interaction*, 4(3):149–196.

[Keenoy and Levene, 2005] Keenoy, K. and Levene, M. (2005). Personalisation of Web Search. In Mobasher, B. and Anand, S., editors, *Intelligent Techniques for Web Personalisation*, pages 1–36. Springer Verlag.

[Kelley, 1967] Kelley, H. H. (1967). Attribution Theory in Social Psychology. In Levine, D., editor, *Nebraska Symposium on Motivation*, Lincoln. University of Nebraska Press.

[Kelley, 1973] Kelley, H. H. (1973). The Processs of Causal Attribution. *American Psychologist*, 28(2):107–128.

[Kelly and Teevan, 2003] Kelly, D. and Teevan, J. (2003). Implicit Feedback for Inferring User Preference: A Bibliography. *Special Interest Group on Information Retrieval (SIGIR) Forum*, 37(2):18–28.

[Khopkar et al., 2003] Khopkar, Y., Spink, A., Giles, C. L., Shah, P., and Debnath, S. (2003). Search Engine Personalisation: An Exploratory Study. *First Monday*, 8(7).

[Kizza, 2003] Kizza, J. M. (2003). *Ethical and Social Issues in the Information Age*. Texts in Computer Science. Springer Verlag, 2nd edition.

[Kjeldskov and Graham, 2003] Kjeldskov, J. and Graham, C. (2003). A Review of Mobile HCI Research Methods. In Chittaro, L., editor, *5th International Symposium on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI)*, pages 317–335, Udine, Italy. Springer Verlag.

[Kjeldskov et al., 2005] Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S., and Davies, J. (2005). Evaluating the Useability of a Mobile Guide: The Influence of Location, Participants and Resources. *Behaviour and Information Technology*, 24(1):51–65.

[Kjeldskov et al., 2004] Kjeldskov, J., Skov, M. B., Als, B. S., and Hoegh, R. T. (2004). Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of

Context-Aware Mobile Systems in the Field. In *6th International Symposium on Human Computer Interaction with Mobile Devices and Services (Mobile HCI)*, Glasgow, UK. Springer Verlag.

[Kobsa, 1990] Kobsa, A. (1990). Modeling the User's Conceptual Knowledge in BGP-MS: A User Modeling Shell System. *Computational Intelligence*, 6(4):193–208.

[Kobsa, 1995] Kobsa, A. (1995). Editorial. *User Modeling and User-Adapted Interaction*, 4(2):iii–v.

[Kobsa, 2001a] Kobsa, A. (2001a). Generic User Modeling Systems. *User Modeling and User-Adapted Interaction*, 11(1-2):49–63.

[Kobsa, 2001b] Kobsa, A. (2001b). Tailoring Privacy to Users' Needs. In *8th International Conference on User Modeling (UM)*, Sonthofen, Germany. Springer Verlag.

[Kobsa, 2002] Kobsa, A. (2002). Personalized Hypermedia and International Privacy. *Communications of the ACM (Special Issue on Adaptive Web-based Systems and Adaptive Hypermedia)*, 45(5):64–67.

[Kobsa et al., 2001] Kobsa, A., Koenemann, J., and Pohl, W. (2001). Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships. *Knowledge Engineering Review*, 16(2):111–155.

[Kobsa and Pohl, 1995] Kobsa, A. and Pohl, W. (1995). The BGP-MS User Modeling System. *User Modeling and User-Adapted Interaction*, 4(2):59–106.

[Koenemann and Belkin, 1996] Koenemann, J. and Belkin, N. (1996). A Case for Interaction: A Study of Interactive Information Retrieval Behaviour and Effectiveness. In *Conference on Human Factors in Computing Systems (CHI)*, Vancouver, BC, Canada. ACM Press.

[Koike et al., 1999] Koike, Y., Kamba, T., and Langheinrich, M. (1999). PIDL-Personalized Information Description Language. W3C Note http://www.w3.org/TR/NOTE-PIDL, W3C.

[Konstan et al., 1997] Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. (1997). GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, 40(3):77–87.

[Korpipää et al., 2003] Korpipää, P., Mäntyjärvi, J., Kela, J., Keränen, H., and Malm, E.-J. (2003). Managing Context Information in Mobile Devices. *IEEE Pervasive Computing*, 2(3):42–51.

[Krüger et al., 2004] Krüger, A., Butz, A., Müller, C., Stahl, C., Wasinger, R., Steinberg, K.-E., and Dirschl, A. (2004). The Connected User Interface: Realizing a Personal Situated Navigation Service. In *International Conference on Intelligent User Interfaces (IUI)*, Island of Madeira, Portugal. ACM Press.

[Kun and Weiner, 1973] Kun, A. and Weiner, B. (1973). Necessary versus Sufficient Causal Schemata for Success and Failure. *Journal of Research in Personality*, 7(3):197–207.

[Lee et al., 1997] Lee, D. L., Chuang, H., and Seamons, K. (1997). Document Ranking and the Vector-Space Model. *IEEE Software*, 14(2):67–75.

[Lieberman, 1995] Lieberman, H. (1995). Letizia: An Agent That Assists Web Browsing. In *14th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 924–929. Morgan Kaufmann Publishers Inc., Montreal, Quebec, Canada.

[Lieberman et al., 2001] Lieberman, H., Fry, C., and Weitzman, L. (2001). Exploring the Web with Reconnaissance Agents. *Communications of the ACM*, 44(8):69–75.

[Lieberman and Selker, 2000] Lieberman, H. and Selker, T. (2000). Out of Context: Computer Systems That Adapt To, and Learn From, Context. *IBM Systems Journal*, 39(3/4):617–632.

[Lucas, 2001] Lucas, P. (2001). Mobile Devices and Mobile Data - Issues of Identity and Reference. *Human-Computer Interaction*, 16(2, 3 & 4):323–336.

[Lueg, 1997] Lueg, C. (1997). Social Filtering and Social Reality. In *5th DELOS Workshop on Filtering and Collaborative Filtering*, Budapest, Hungary. European Research Consortium for Informatics and Mathematics (ERCIM).

[Manber et al., 2000] Manber, U., Patel, A., and Robison, J. (2000). Experience with personalization of Yahoo! *Communications of the ACM*, 43(8):35–39.

[Mantyjarvi et al., 2006] Mantyjarvi, J., Paterno, F., Salvador, Z., and Santoro, C. (2006). Scan and Tilt Towards Natural Interaction for Mobile Museum Guides. In *8th Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 191–194, Helsinki, Finland. ACM Press.

[Marmasse and Schmandt, 2000] Marmasse, N. and Schmandt, C. (2000). Location-aware information delivery with comMotion. In *2nd International Symposium on Handheld and Ubiquitous Computing (HUC)*, pages 157–171, Bristol, UK. Springer Verlag.

[Maron and Kuhns, 1960] Maron, M. E. and Kuhns, J. L. (1960). On Relevance, Probabilistic Indexing, and Information Retrieval. *Journal of the ACM*, 7(3):216–244.

[McArthur, 1972] McArthur, L. A. (1972). The How and What of Why: Some Determinants and Consequences of Causal Attribution. *Personality and Social Psychology*, 22(2):171–193.

[Mladenic, 1996] Mladenic, D. (1996). Personal WebWatcher: Design and Implementation. Technical Report IJS-DP-7472, Department of Intelligent Systems, J. Stefan Institute.

[Mooney and Roy, 2000] Mooney, R. J. and Roy, L. (2000). Content-based book recommending using learning for text categorization. In *5th International Conference on Digital Libraries*, pages 195–204, San Antonio, TX, USA. ACM Press.

[Morse et al., 2000] Morse, D. R., Armstrong, S., and Dey, A. K. (2000). The what, who, where, when, why and how of context-awareness. In *Conference on Human Factors in Computing Systems (CHI)*, pages 371–371, The Hague, The Netherlands. ACM Press.

[Mountain and MacFarlane, 2007] Mountain, D. M. and MacFarlane, A. (2007). Geographic Information Retrieval in a Mobile Environment: Evaluating the Needs of Mobile Individuals. *Journal of Information Science*, 33(5):515–530.

[Mulligan et al., 2000] Mulligan, D., Schwartz, A., Cavoukian, A., and Gurski, M. (2000). P3P and Privacy: An Update for the Privacy Community. Technical report, Center for Democracy & Technology.

[Murphy and Myors, 2003] Murphy, K. R. and Myors, B. (2003). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests*. Erlbaum, Mahwah, NJ, USA, 2nd edition.

[Myrhaug et al., 2004a] Myrhaug, H., Whitehead, N., Göker, A., Faegri, T. E., and Lech, T. C. (2004a). AmbieSense: A System and Reference Architecture for Personalised Context-Sensitive Information Services for Mobile Users. In *European Symposium on Ambient Intelligence (EUSAI)*, Eindhoven, The Netherlands. Springer Verlag.

[Myrhaug and Göker, 2003] Myrhaug, H. I. and Göker, A. (2003). AmbieSense - Interactive Information Channels in the Surroundings of the Mobile User. In *2nd International Conference on Universal Access in Human-Computer Interaction*, volume 4, pages 1158–1162, Crete, Greece. Lawrence Erlbaum Associates.

[Myrhaug and Göker, 2004] Myrhaug, H. I. and Göker, A. (2004). AmbieSense Test and Evaluation Report. Technical report, AmbieSense Project, The Robert Gordon University.

[Myrhaug et al., 2004b] Myrhaug, H. I., Göker, A., Pollich, J., and Watt, S. (2004b). AmbieSense Final Report. Technical report, AmbieSense Project, The Robert Gordon University.

[Ni et al., 2004] Ni, L. M., Liu, Y., Lau, Y. C., and Patil, A. P. (2004). LANDMARC: Indoor Location Sensing Using Active RFID. *Wireless Networks*, 10(6):701–710.

[Nichols, 1997] Nichols, D. M. (1997). Implicit rating and filtering. In *5th DELOS Workshop on Filtering and Collaborative Filtering*, pages 31–36, Budapest, Hungary. European Research Consortium for Informatics and Mathematics (ERCIM).

[Nielsen, 1998] Nielsen, J. (1998). Personalization is Over-Rated, http://www.useit.com/alertbox/981004.html.

[Nielsen, 2005] Nielsen, J. (2005). Search: Visible and Simple, http://www.useit.com/alertbox/20010513.html.

[Nieminen and Röykkee, 2006] Nieminen, M. and Röykkee, M. (2006). 8th International Conference on Human Computer Interaction with Mobile Devices and Services (Mobile HCI). Helsinki, Finland. ACM Press.

[Oard and Kim, 1998] Oard, D. and Kim, J. (1998). Implicit Feedback for Recommender Systems. In *AAAI Technical Report WS-98-08: Workshop on Recommender Systems*, Madison, WI, USA. AAAI Press.

[Oard and Kim, 2001] Oard, D. W. and Kim, J. (2001). Modeling Information Content Using Observable Behaviour. In *64th Annual Meeting of the American Society for Information Science and Technology (ASIST)*, Washington DC, USA.

[Oppermann and Specht, 2000] Oppermann, R. and Specht, M. (2000). A Context-Sensitive Nomadic Information System as an Exhibition Guide. In Thomas, P. and Gellersen, H.-W., editors, *2nd International Symposium on Handheld and Ubiquitous Computing (HUC)*, pages 127–142, Bristol, UK. Springer Verlag.

[O'Riordan and Sorensen, 1999] O'Riordan, C. and Sorensen, H. (1999). Information Retrieval and Filtering - An Overview. Technical report, National University of Ireland, Department of Information Technology.

[Orwant, 1995] Orwant, J. (1995). Heterogenous Learning in the Doppelgänger User Modeling System. *User Modeling and User-Adapted Interaction*, 4(2):107–130.

[Oulasvirta et al., 2005] Oulasvirta, A., Tamminen, S., Roto, V., and Kuorelahti, J. (2005). Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. In *Conference on Human Factors in Computing Systems (CHI)*, pages 919–928, Portland, OR, USA. ACM Press.

[P3P, 2002] P3P (2002). Platform for Privacy Preferences (P3P) Project, http://www.w3.org/P3P.

[Pazzani, 2002] Pazzani, M. J. (2002). Commercial Application of Machine Learning for Personalized Wireless Portals. In *Pacific Rim Conference on Artificial Intelligence*, pages 1–5, Tokyo, Japan. Springer Verlag.

[Pazzani et al., 1998] Pazzani, M. J., Muramatsu, J., and Billsus, D. (1998). Syskill & Webert: Identifying Interesting Web Sites. In *13th National Conference on Artificial Intelligence (AAAI)*, pages 54–61, Portland, OR, USA. AAAI Press.

[Picard, 1997] Picard, R. (1997). *Affective Computing*. MIT Press, Cambridge, MA.

[Pitkow et al., 2002] Pitkow, J., Schuetze, H., Cass, T., Cooley, R., Turnbull, D., Edmonds, A., Adar, E., and Breuel, T. (2002). Personalized Search. *Communications of the ACM*, 45(9):50–55.

[Pretschner and Gauch, 1999] Pretschner, A. and Gauch, S. (1999). Ontology Based Personalized Search. In *International Conference on Tools with Artificial Intelligentce (ICTAI)*, pages 391–398, Chicago, USA.

[Priyantha et al., 2000] Priyantha, N., Chakraborty, A., and Balakrishnan, H. (2000). The Cricket Location-Support System. In *6th International Conference on Mobile Computing and Networking (MOBICOM)*, Boston, MA, USA. ACM Press.

[Rantakokko and Plomp, 2003] Rantakokko, T. and Plomp, J. (2003). An Adaptive Map-Based Interface for Situated Services. In *Smart Objects Conference 2003*, Grenoble, France.

[Reichenbacher, 2003] Reichenbacher, T. (2003). Adaptive Methods for Mobile Cartography. In *21st International Cartographic Conference*, Durban, South Africa.

[Reichenbacher, 2007] Reichenbacher, T. (2007). Adapation in Mobile and Ubiquitous Cartography. In Cartwright, W., Peterson, M. P., and Gartner, G., editors, *Multimedia Cartography*, pages 383–397. Springer Verlag, Berlin.

[Reid, 2000] Reid, J. (2000). A Task-Oriented Non-Interactive Evaluation Methodology for Information Retrieval Systems. *Information Retrieval*, 2(1):115–129.

[Resnick et al., 1994] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *5th Conference on Computer Supported Cooperative Work*, pages 175–186, Chapel Hill, NC, USA. ACM Press.

[Resnick and Varian, 1997] Resnick, P. and Varian, H. R. (1997). Special Issue on Recommender Systems. *Communications of the ACM*, 40(3).

[Rich, 1979] Rich, E. (1979). User Modeling via Stereotypes. *Cognitive Science*, 3(4):329–354.

[Riecken, 2000] Riecken, D. (2000). Special Issue on Personalisation. *Communications of the ACM*, 43(8).

[Robertson and Hancock-Beaulieu, 1992] Robertson, S. E. and Hancock-Beaulieu, M. M. (1992). On the Evaluation of IR Systems. *Information Processing and Management*, 28(4):457–466.

[Ruthven, 2001] Ruthven, I. (2001). *Abduction, Explanation and Relevance Feedback*. Phd thesis, University of Glasgow.

[Ruthven, 2005] Ruthven, I. (2005). Integrating Approaches to Relevance. In Spink, A. and Cole, C., editors, *New Directions in Cognitive Information Retrieval*, pages 61–80. Springer Verlag, Dordrecht, The Netherlands.

[Ruthven et al., 2006] Ruthven, I., Borlund, P., Ingwersen, P., Belkin, N., Tombros, A., and Vakkari, P., editors (2006). *Information Interaction in Context. 1st International Symposium on Information Interaction in Context, IIiX 2006*. ACM Press, Copenhagen, Denmark.

[Sakagami and Kamba, 1997] Sakagami, H. and Kamba, T. (1997). Learning personal preferences on online newspaper articles from user behaviors. *Computer Networks and ISDN Systems*, 29(8-13):1447–1455.

[Salber et al., 1999] Salber, D., Dey, A. K., and Adowd, G. D. (1999). The Context Toolkit: Aiding the Development of Context-Enabled Applications. In *Conference on Human Factors in Computing Systems (CHI)*, pages 434–441, Pittsburgh, PA, USA. ACM Press.

[Salton, 1971] Salton, G. (1971). *The SMART Retrieval System - Experiments in Automatic Document Processing*. Prentice-Hall, Upper Saddle River, NJ, USA.

[Salton and Buckley, 1988] Salton, G. and Buckley, C. (1988). Term-Weighting Approaches in Automatic Text Retrieval. *Information Processing and Management*, 24(5):513–523.

[Saracevic, 1995] Saracevic, T. (1995). Evaluation of Evaluation in Information Retrieval. In Fox, E. A., Ingwersen, P., and Fidel, R., editors, *18th Annual International ACM*

SIGIR Conference on Research and Development in Information Retrieval, pages 138–146, Seattle, WA, USA. ACM Press.

[Saracevic, 1996] Saracevic, T. (1996). Modeling Interaction in Information Retrieval (IR): A Review and Proposal. In *Proceedings of the American Society for Information Science (ASIS) Annual Meeting*, volume 33, pages 3–9.

[Saracevic, 1997] Saracevic, T. (1997). The Stratified Model of Information Retrieval Interaction: Extensions and Applications. In *Proceedings of the American Society for Information Science (ASIS) Annual Meeting*, volume 34, pages 313–327.

[Saracevic, 1999] Saracevic, T. (1999). Information Science. *Journal of the American Society for Information Science*, 50(12):1051–1063.

[Schamber et al., 1990] Schamber, L., Eisenberg, M. B., and Nilan, M. S. (1990). A Re-Examination of Relevance: Toward a Dynamic, Situational Definition. *Information Processing and Management*, 26(6):755–776.

[Schilit et al., 1994] Schilit, B., Adams, N., and Want, R. (1994). Context-Aware Computing Applications. In *IEEE Workshop on Mobile Computing Systems and Applications*, pages 85–90, Santa Cruz, CA, USA. IEEE Computer Society Press.

[Schmidt, 2002] Schmidt, A. (2002). *Ubiquitous Computing - Computing in Context*. Phd thesis, Lancaster University.

[Schmidt et al., 1999a] Schmidt, A., Aidoo, K. A., Takaluomai, A., Tuomelai, U., Laerhoven, K. V., and de Velde, W. V. (1999a). Advanced Interaction in Context. In *1st international Symposium on Handheld and Ubiquitous Computing (HUC)*, pages 89–101, Karlsruhe, Germany. Springer Verlag.

[Schmidt et al., 1999b] Schmidt, A., Beigl, M., and Gelersen, H.-W. (1999b). There is More to Context than Location. *Computers & Graphics Journal*, 23(6):893–902.

[Schmidt and Laerhoven, 2001] Schmidt, A. and Laerhoven, K. V. (2001). How to Build Smart Appliances? *IEEE Personal Communications*, 8(4):66–71.

[Schmidt-Belz et al., 2003] Schmidt-Belz, B., Laamanen, H., Poslad, S., and Zipf, A. (2003). Location-Based Mobile Tourist Services - First User Experiences. In *IFITT's*

Global Travel & Tourism Technology and eBusiness Forum (ENTER), Helsinki, Finland. Springer Verlag.

[Scholtz, 2006] Scholtz, J. (2006). Metrics for Evaluating Human Information Interaction Systems. *Interacting with Computers*, 18(4):507–527.

[Schroeder et al., 1986] Schroeder, L. D., Sjoquist, D. L., and Stephan, P. E. (1986). *Understanding Regression Analysis: An Introductory Guide*. Sage Publications.

[Shardanand and Maes, 1995] Shardanand, U. and Maes, P. (1995). Social Information Filtering: Algorithms for Automating "Word of Mouth". In *Conference on Human Factors in Computing Systems (CHI)*, pages 210–217, Denver, CO, USA.

[Silverstein et al., 1999] Silverstein, C., Marais, H., Henzinger, M., and Moricz, M. (1999). Analysis of a very large web search engine query log. *Special Interest Group on Information Retrieval (SIGIR) Forum*, 33(1):6–12.

[Smeaton and Callan, 2001] Smeaton, A. and Callan, J. (2001). Joint DELOS-NSF Workshop on Personalisation and Recommender Systems in Digital Libraries. *Special Interest Group on Information Retrieval (SIGIR) Forum*, 35(1):7–11.

[Spärck Jones, 1972] Spärck Jones, K. (1972). A Statistical Interpretation of Term Specificity and its Application in Retrieval. *Journal of Documentation*, 28(1):11–21.

[Spärck Jones, 2003] Spärck Jones, K. (2003). Privacy: What's different now? *Interdisciplinary Science Reviews*, 28(4):287–292.

[Spink and Losee, 1996] Spink, A. and Losee, R. M. (1996). Feedback in Information Retrieval. *Annual Review of Information Science and Technology*, 31:33–78.

[Spink et al., 2001] Spink, A., Wolfram, D., Jansen, M. B. J., and Saracevic, T. (2001). Searching the Web: The Public and Their Queries. *Journal of the American Society for Information Science and Technology*, 52(3):226–234.

[Strang and Linnhoff-Popien, 2004] Strang, T. and Linnhoff-Popien, C. (2004). A Context Modeling Survey. In *6th International Conference on Ubiquitous Computing (UbiComp), 1st International Workshop on Advanced Context Modelling, Reasoning And Management*, Nottingham, UK.

[Surber, 1981] Surber, C. F. (1981). Necessary versus Sufficient Causal Schemata: Attributions for Achievement in Difficult and Easy Tasks. *Journal of Experimental Social Psychology*, 17(6):569–586.

[Tamminen et al., 2004] Tamminen, S., Oulasvirta, A., Toiskallio, K., and Kankainen, A. (2004). Understanding mobile contexts. *Personal and Ubiquitous Computing*, 8(2):135–143.

[Tazari et al., 2004] Tazari, M.-R., Grimm, M., and Finke, M. (2004). Modelling User Context. In *6th International Conference on Ubiquitous Computing (UbiComp), 1st International Workshop on Advanced Context Modelling, Reasoning And Management*, Nottingham, UK.

[Teevan, 2007] Teevan, J. (2007). The Re:Search Engine: Simultaneous Support for Finding and Re-Finding. In *20th Annual ACM Symposium on User Interface Software and Technology (UIST)*, Newport, RI, USA. ACM Press.

[Vakkari, 2001] Vakkari, P. (2001). A Theory of the Task-Based Information Retrieval Process: A Summary and Generalisation of a Longitudinal Study. *Journal of Documentation*, 57(1):44–60.

[van Elzakker, 2000] van Elzakker, C. P. J. M. (2000). Use and Users of Maps on the Web. *Cartographic Perspectives*, 37:34–50.

[van Rijsbergen, 1979] van Rijsbergen, C. (1979). *Information Retireval*. Butterworths, London.

[Voorhees and Harman, 2005] Voorhees, E. M. and Harman, D. K. (2005). *TREC: Experiment and Evaluation in Information Retrieval*. MIT Press.

[Want et al., 1992] Want, R., Hopper, A., Falcao, V., and Gibbons, J. (1992). The Active Badge Location System. *ACM Transactions on Information Systems (TOIS)*, 10(1):91–102.

[Ward et al., 1997] Ward, A., Jones, A., and Hopper, A. (1997). A New Location Technique for the Active Office. *IEEE Personal Communications*, 4(5):42–47.

[White, 2004] White, R. (2004). *Implicit Feedback for Interactive Information Retrieval.* Phd thesis, University of Glasgow.

[Wilcoxon, 1945] Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics*, 1(6):80–83.

[Wilson, 1999] Wilson, T. D. (1999). Models in Information Retrieval Research. *Journal of Documentation*, 55(3):249–270.

[Wirth, 1998] Wirth, U. (1998). What is Abductive Inference? In Bouissac, P., editor, *Encyclopaedia of Semiotics*. Oxford University Press, Oxford.

[WTO, 2007] WTO (2007). UNWTO World Tourism Barometer. 5(2).

[Xiang et al., 2004] Xiang, Z., Song, S., Chen, J., Wang, H., Huang, J., and Gao, X. (2004). A Wireless LAN-Based Inddor Positioning Technology. *IBM Journal of Research and Development*, 48(5/6):617–626.

[Zipf, 2002] Zipf, A. (2002). User-Adaptive Maps for Location-Based Services (LBS) for Tourism. In *9th Int. Conf. for Information and Communication Technologies in Tourism (ENTER 2002)*, pages 329–337, Innsbruck, Austria. Springer Verlag.

[Zukerman and Albrecht, 2001] Zukerman, I. and Albrecht, D. W. (2001). Predictive Statistical Models for User Modeling. *User Modeling and User-Adapted Interaction*, 11(1-2):5–18.

Appendices

A

Privacy and Usability Issues in Personalisation

Personalisation is frequently related to other issues that are not of major concern for this thesis, but should nevertheless be covered in its basics. Two issues are briefly reviewed here - the topic of privacy as a concern that is frequently expressed in relation to personalisation and the issue of usability that faces new challenges within a personalised system.

A.1 Personalisation and Privacy

According to [Kizza, 2003, p. 108], privacy is control over personal information and external influences. This includes, based on the work by Jerry Durlak, the right to be alone, the right to remain anonymous, the right not to be monitored and the right to have control over both the personal information itself and the methods for its dissemination. Karen Spärck Jones describes privacy more fundamentally as the ability "not having things known about you that you don't choose to have known, or at least you know that they are known, and by whom" [Spärck Jones, 2003].

Privacy is an important topic as personalised systems collect information about their users - a necessity for system and web developers and a concern for users [Kobsa, 2001b]. A recent paper by Kobsa found that Internet users are more likely to provide data when they feel sure that they can remain anonymous [Kobsa, 2002]. This is important, as the success or failure of personalised information systems strongly depends on the willingness of individuals to provide data. Therefore, personalised information system should take an initiative to ensure that users privacy is secured. A number of potential solutions for this are available and discussed in the literature:

- **Client-sided personalisation** offers a solution for avoiding the uncontrolled distribution of personal data through local storage [Kobsa, 2002]. Examples for client sided personalisation include the WHAT system [Cassel and Wolz, 2001], CASPER (Case-Based Profiling for Electronic Recruitment) [Bradley et al., 2000], PResTo! [Keenoy and Levene, 2005] and Pitkow's Outride system [Pitkow et al., 2002] just to name a few.

- **Privacy policies** allow to regulate and formalise the treatment of user data for building a trust relationship between users and businesses. One examples of such a policy is the Yahoo! Privacy policy that is described in [Manber et al., 2000] or the Lycos privacy policy [1]. However, such private policies do usually not cover for unexpected changes in the structure of companies (i.e. change in ownership or bankruptcy). Based on that, there are increasing efforts in establishing privacy and trust standards that reach beyond the boundaries of individual enterprises. The Platform for Privacy Preferences [P3P, 2002] attempts to standardise privacy for the web. With P3P, users setup their privacy preferences in a standard and machine-readable form. Information providers (i.e. a website offering personalised services) provide a machine-readable privacy policy. Applications can then automatically evaluate the policy of a web resource [Cingil et al., 2000]. When implemented, P3P can help to establish trust between users and web resources but will only work in an environment, where the users jurisdiction is supported by sufficient data privacy laws. This, of cause, is beyond the powers of the standard [Mulligan et al., 2000]. Since November 2006, the standard is in a final state and put on hold, as current web browsers do not yet provide the necessary support. A similar aim in global privacy standardisation is followed by the Policy Aware Web [2]. The EU-IST project PRIME [3] currently develops a prototype system for managing privacy and evaluating it in real-world scenarios some of which are relevant for mobile applications (i.e. internet communication and location-based services).

[1] `http://www.lycos.com/privacy`, accessed April 14, 2008
[2] `http://www.policyawareweb.org`, accessed April 14, 2008
[3] `https://www.prime-project.eu`, accessed April 14, 2008

In [Kobsa, 2001b] and [Kobsa, 2002] a number of design guidelines for personalised hypermedia systems are proposed to improve user privacy and the trust between the user and the enterprise. These guidelines also apply for any system that applies personalisation:

1. The application should inform users clearly and comprehensively about potentially sensitive data and whenever it is processed by the service. It should be done in a way so that the purpose of personal information and its use within the system is clear to the user.

2. The personalised application should allow users inspecting all data that is stored about them. This enables the user to resolve the potential uncertainty about what information the system is using and processing.

A.2 Personalisation and Useability

A personalised information system has a different underlying philosophy with respect to the meaning of information and its use which may affect traditional patterns of usage. Some of the most important issues are:

1. **Predictability:** The article on MyYahoo! [Manber et al., 2000] states that personalisation is better when it is straight forward allowing the user to predict its actions. MyYahoo! uses for example location information to highlight content that is associated with that location (e.g. weather or sport news). Location is a very straight-forward attribute that allows people to make this connection between cause and effect easily. However, this cannot be said for almost any other attribute like the user's interests, the user's current task, role, etc. As soon as personalisation is based on a more indirect attribute, the process becomes ultimately much less straight forward and potentially mysterious for the user. However, this effect can be compensated by an informative system that allows users to investigate the information that is used, how it is used and what the personalised service infers from this information [Kobsa, 2001b, Kobsa, 2002]. This is also relevant with respect to privacy as discussed in the previous section.

2. **Content Dynamics:** A publishing house usually keeps records of previous publications such as newspapers. One major purpose of libraries is the record

keeping of past publications (i.e. scientific journals). Dynamic and fast changing content however comes with the problem that record keeping becomes difficult or even impossible. With personalisation this becomes even more challenging as a user might have a unique compilation of content (i.e. news). Since content might change quickly and perhaps infrequent, it can be challenging to recover a previous state (i.e. a content item that was recommended the day before). This also relates to the problem of re-finding information [Teevan, 2007], an important daily search activity where users put preference on information that was previously discovered rather than any new content. It also relates to the predictability issue since the user might not be completely aware of how that personalised content item was produced and what can be done to reproduce that output. For this reason, it can be valuable for a personalised information system to provide navigation aids and other means of control that help users to reproduce past states.

3. **Sharing:** The previous point about content dynamics also bears the issue of content sharing - a concept that is commonly understood by users in a rather static way. This means, users expect content to stay and keep being accessible over a longer period of time. Evidence for that can be found in the success of bookmarks that are built on this assumption. Also, it is quite common for users to communicate web hyperlinks (e.g. through email and messengers). When sharing or storing hyperlinks, users assume that the content will reappear equally for themselves or others. However, this concept becomes weak and potentially invalid when using a personalised information system. The personalisation process might produce alternative content, content collections and/or visualisations for each of its users. A weaker form of this effect can be observed with dynamic content. For example, a dynamic website might produce output and its hyperlink might be expired only minutes later. This means, the hyperlink has lost its purpose for collaborative use. To cope with that, the end user has to understand the individual nature of the information that is provided. On the other hand, system designers should provide tools that allow content to be communicated despite of its personalised nature.

B

AmbieSense Questionnaires

B.1 Overview

This appendix provides relevant AmbieSense questionnaires based on the data presented in chapter 3 that highlighted some AmbieSense findings relevant for this thesis. Note that most of the data in chapter 3 was produced by AmbieSense and is owned by the AmbieSense Consortium[1]. It was presented in this thesis based on its relevance and its strong connection with this research.

In particular, data from three AmbieSense studies were presented in chapter 3 – a large-scale market survey, a mobile study conducted in Seville in June 2004, and another mobile study also conducted in Seville in September 2004. The questionnaires are listed in this order in the following three subsections[2]

B.2 AmbieSense Market Survey Questionnaire

This is the AmbieSense questionnaire that was used to collect data for a large-scale market survey. Data was collected from four different locations - Seville, Oslo, the Oslo Airport website [3], and the Lonely Planet website[4]. The gender question (Question A1a on the following page) has been removed for the Lonely Planet web version based on direct request

[1]For further information, please visit the project website at `http://www.ambiesense.net/`, accessed April 14, 2008

[2]The AmbieSense title page and one page with a general description about AmbieSense has been omitted for reasons of focus.

[3]http://www.osl.no, accessed April 14, 2008

[4]http://www.lonelyplanet.com, accessed April 14, 2008

by Lonely Planet. All other questions were presented equally to participants in all four locations.

Demographic Information

A1. Please answer a few general questions about yourself

a) Gender: ☐ Male ☐ Female
b) Age : ☐Under 20 ☐20 – 29 ☐30 – 39 ☐40 – 49 ☐50 - 59 ☐60 and over

Computing Experience

A2. How often do you use the following devices, please answer those you have personally used and are familiar with?

1)	Mobile Phone	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
2)	PDA	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
3)	Smartphones (e.g. P900, Nokia 6600)	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
4)	Laptop	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
5)	Wireless communication	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily

A3. Where do you use handheld device(s)? (Please tick all that apply)

1) ☐ At home
2) ☐ At work
3) ☐ While on the move (e.g. plane, train, bus)
4) ☐ At other places (e.g. hotel, restaurant, bar, shop, museum)

Travel information

A4. In which state of your travel are you in? (Please tick only one)
1) ☐ Departure (the first flight)
2) ☐ Transit (between two flights)
3) ☐ Arrival (the end of flight(s))
4) ☐ At the destination (e.g. in the tourist city, meeting location)

A5. How many times do you travel (between city or country) per year? (Please tick only one)
1) ☐ 0-1 time
2) ☐ 2-5 times
3) ☐ 5-10 times
4) ☐ above 10 times

A6. For what purposes do you travel? (Please tick only one)
1) ☐ Business
2) ☐ Private
3) ☐ Business and private

Figure B.1: *Printed with permission of AmbieSense^{TM}.*

Before Travelling

A7. What **type** of information do you gather **before** you travel? (Please tick all that apply)

☐ Accommodation (e.g. hotels, B&B) ☐ Restaurants ☐ Sites/ Attractions ☐ Transport ☐ Maps	☐ Events (e.g. concert, festival) ☐ Shops ☐ Weather ☐ Other ___________________

A8. Where do you find the information **before** you travel? (Please tick all that apply)

☐ Internet ☐ Digital travel guide (e.g. CD ROM) ☐ Brochures ☐ Tourist Information Centre / Office ☐ Travel Agent	☐ Word of Mouth (e.g. friends) ☐ Bookstore ☐ Library ☐ Hotel reception ☐ Other ___________________

While Travelling

A9. What **type** of information do you gather **while** travelling? (Please tick all that apply)

☐ Accommodation (e.g. hotels, B&B) ☐ Restaurants ☐ Sites/ Attractions ☐ Transport ☐ Maps	☐ Events (e.g. concert, festival) ☐ Shops ☐ Weather ☐ Other ___________________

A10. Where do you find the information **while** travelling? (Please tick all that apply)

☐ Internet ☐ Digital travel guide (e.g. CD ROM) ☐ Brochures ☐ Tourist Information Centre / Office ☐ Travel Agent	☐ Word of Mouth (e.g. friends) ☐ Bookstore ☐ Library ☐ Hotel reception ☐ Other ___________________

AmbieSense Evaluation Questionnaire 3

Figure B.2: *Printed with permission of AmbieSense^{TM}.*

IST 2001- 34244 AmbieSense.

The future digital airport

A11. What kind of information would you like to receive while you are in the airport?
(Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
1) My flight	1	2	3	4	5	6	7
2) Other flights	1	2	3	4	5	6	7
3) Airline companies	1	2	3	4	5	6	7
4) Transportation	1	2	3	4	5	6	7
5) Airline lounges	1	2	3	4	5	6	7
6) Lost and found	1	2	3	4	5	6	7
7) Map over the airport	1	2	3	4	5	6	7
8) Airport information	1	2	3	4	5	6	7
9) Electronic news	1	2	3	4	5	6	7
10) Electronic magazines	1	2	3	4	5	6	7
11) Advertisements	1	2	3	4	5	6	7
12) Airport shopping	1	2	3	4	5	6	7
13) Personalised offers	1	2	3	4	5	6	7
14) Last minute tickets	1	2	3	4	5	6	7
15) Restaurants and food	1	2	3	4	5	6	7
16) Digital radio/ TV programs	1	2	3	4	5	6	7
17) Information about your destination	1	2	3	4	5	6	7

AmbieSense Evaluation Questionnaire 4

Figure B.3: *Printed with permission of AmbieSense^{TM}.*

A12. What kind of digital information services would be most useful/ fun for you to use if you have time there? (Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
	1	2	3	4	5	6	7
1) Internet access	1	2	3	4	5	6	7
2) My flight	1	2	3	4	5	6	7
3) Digital travel guides	1	2	3	4	5	6	7
4) Airport shopping services	1	2	3	4	5	6	7
5) Digital books	1	2	3	4	5	6	7
6) Weekly and monthly magazines	1	2	3	4	5	6	7
7) Financial news services	1	2	3	4	5	6	7
8) Information services based on my interests	1	2	3	4	5	6	7
9) Infotainment for children	1	2	3	4	5	6	7
10) Digital radio/ TV channels	1	2	3	4	5	6	7
11) Airport information	1	2	3	4	5	6	7
12) Airline information	1	2	3	4	5	6	7
13) Online booking services	1	2	3	4	5	6	7
14) Other transportation services	1	2	3	4	5	6	7

Figure B.4: *Printed with permission of AmbieSense^{TM}.*

A13. What kind of equipment would you prefer to use in the airport for accessing digital information services - including personal information? (Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
1) My own mobile phone	1	2	3	4	5	6	7
2) My own PDA	1	2	3	4	5	6	7
3) My own laptop	1	2	3	4	5	6	7
4) A borrowed computer or screen	1	2	3	4	5	6	7
5) A borrowed laptop	1	2	3	4	5	6	7
6) A borrowed PDA	1	2	3	4	5	6	7
7) Wired communication equipment	1	2	3	4	5	6	7
8) Wireless communication equipment	1	2	3	4	5	6	7
9) Large screens in public areas	1	2	3	4	5	6	7
10) Small screens in public areas	1	2	3	4	5	6	7
11) Interactive screens in the airplane seats	1	2	3	4	5	6	7

Figure B.5: *Printed with permission of AmbieSense^{TM}.*

A14. Where would you like to access digital information services in conjunction with the airport? (Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
1) In the airport train	1	2	3	4	5	6	7
2) In the airport bus	1	2	3	4	5	6	7
3) In the arrivals hall	1	2	3	4	5	6	7
4) In the departure hall	1	2	3	4	5	6	7
5) In the lounge	1	2	3	4	5	6	7
6) In the restaurant/bar	1	2	3	4	5	6	7
7) At the security point	1	2	3	4	5	6	7
8) All over the airport	1	2	3	4	5	6	7
9) While walking between gates	1	2	3	4	5	6	7
10) While sitting down	1	2	3	4	5	6	7
11) While shopping	1	2	3	4	5	6	7
12) While waiting by the gate	1	2	3	4	5	6	7
13) Specific information in specific areas of the airport	1	2	3	4	5	6	7

Figure B.6: *Printed with permission of AmbieSense^{TM}.*

The future digital tourist city

A15. When you visit a region, or city, how important are the following aspects for you? (Please circle only on of the numbers on the scale)

	Not relevant						Highly relevant
Food/ Restaurants	1	2	3	4	5	6	7
Art/ Sculpture/ Crafts	1	2	3	4	5	6	7
People and Culture	1	2	3	4	5	6	7
Transportation (e.g. type, schedule)	1	2	3	4	5	6	7
Accommodation (e.g. hotels, B&B)	1	2	3	4	5	6	7
History	1	2	3	4	5	6	7
Nature/ Countryside	1	2	3	4	5	6	7
Event (e.g. concert, festivals)	1	2	3	4	5	6	7
Nightlife	1	2	3	4	5	6	7
Exchange rate and economics	1	2	3	4	5	6	7
Research/Universities	1	2	3	4	5	6	7
Architecture	1	2	3	4	5	6	7
Political stability	1	2	3	4	5	6	7
News	1	2	3	4	5	6	7
Sites / Attractions	1	2	3	4	5	6	7
Shopping	1	2	3	4	5	6	7
Health / Sport Facilities	1	2	3	4	5	6	7

Figure B.7: *Printed with permission of AmbieSense^{TM}.*

A16. What kind of digital information services would be most useful/ fun for you to use as a tourist in a city? (Please circle only on of the numbers in each row)

	Not relevant						**Highly relevant**
1) Internet access	1	2	3	4	5	6	7
2) My travel service	1	2	3	4	5	6	7
3) Digital travel guides	1	2	3	4	5	6	7
4) Local shopping services	1	2	3	4	5	6	7
5) Digital books	1	2	3	4	5	6	7
6) Weekly and monthly magazines	1	2	3	4	5	6	7
7) Financial news services	1	2	3	4	5	6	7
8) Information services based on my interests	1	2	3	4	5	6	7
9) Infotainment for children	1	2	3	4	5	6	7
10) Digital radio/ TV channels	1	2	3	4	5	6	7
11) Airport information	1	2	3	4	5	6	7
12) Airline information	1	2	3	4	5	6	7
13) Online booking services	1	2	3	4	5	6	7
14) Other transportation services	1	2	3	4	5	6	7
15) News service	1	2	3	4	5	6	7
17) Local/regional events service	1	2	3	4	5	6	7
18) City map service	1	2	3	4	5	6	7

Figure B.8: *Printed with permission of AmbieSense^{TM}.*

A17. What kind of equipment would you prefer to use in the tourist city for accessing digital information services - including personal information? (Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
1) My own mobile phone	1	2	3	4	5	6	7
2) My own PDA	1	2	3	4	5	6	7
3) My own laptop	1	2	3	4	5	6	7
4) A borrowed computer or screen	1	2	3	4	5	6	7
5) A borrowed laptop	1	2	3	4	5	6	7
6) A borrowed PDA	1	2	3	4	5	6	7
7) Wired communication equipment	1	2	3	4	5	6	7
8) Wireless communication equipment	1	2	3	4	5	6	7
9) Large screens in public areas	1	2	3	4	5	6	7
10) Small screens in public areas	1	2	3	4	5	6	7
11) Interactive screens in the airplane seats	1	2	3	4	5	6	7

Figure B.9: *Printed with permission of AmbieSense^{TM}.*

A18. Where would you like to access digital information services in conjunction with the tourist city? (Please circle only on of the numbers in each row)

	Not relevant						Highly relevant
1) In the city train	1	2	3	4	5	6	7
2) In the city bus	1	2	3	4	5	6	7
3) In the train/ bus station/ airport	1	2	3	4	5	6	7
4) In the shopping centre	1	2	3	4	5	6	7
5) In the hotel	1	2	3	4	5	6	7
6) In the restaurant/bar	1	2	3	4	5	6	7
7) In the pedestrian area	1	2	3	4	5	6	7
8) While driving car	1	2	3	4	5	6	7
9) While walking between shops	1	2	3	4	5	6	7
10) While sitting down	1	2	3	4	5	6	7
11) While shopping	1	2	3	4	5	6	7
12) While at the museum	1	2	3	4	5	6	7
13) Specific information in specific areas of the tourist city	1	2	3	4	5	6	7

End of Questionnaire

Thank you for taking part in this evaluation about future digital airports, your time and effort are greatly appreciated.

Figure B.10: *Printed with permission of AmbieSense^{TM}.*

B.3 AmbieSense Seville June 2004 Questionnaire

This is the AmbieSense questionnaire that was used to collect data during a mobile user study in Seville in June 2004. Only the pre- and post-questionnaire are shown since no data from the search tasks has been presented in this thesis. For more information on the entire mobile study, please refer to [Göker and Myrhaug, 2007].

IST 2001- 34244 AmbieSense.

Interviewer: _______**Date:** ______**Time:**______**Location:** _______**Questionnaire No:**___

Part 1: Before the Test

Demographic Information

A1. Please answer a few general questions about yourself

Gender : ☐ Male ☐ Female

Age : ☐ Under 18 ☐ 19 – 39 ☐ 40 - 60 ☐ Over 60

Spoken Languages: ____

Computing Experience

A2. How often do you use one of the following devices? (please tick one box for each device)

1. Mobile Phone
☐ not used so far ☐ A few times a year ☐ Weekly ☐ Daily

2. PDA
☐ not used so far ☐ A few times a year ☐ Weekly ☐ Daily

3. Smartphone
☐ not used so far ☐ A few times a year ☐ Weekly ☐ Daily

4. Laptop
☐ not used so far ☐ A few times a year ☐ Weekly ☐ Daily

5. Desktop
☐ not used so far ☐ A few times a year ☐ Weekly ☐ Daily

If you have *not* used a Mobile Phone, a PDA or a Smartphone, please go to question A5.

A3. Where have you used them? (please tick all that apply)
☐ At home
☐ At work
☐ While on the move (e.g. plane, train, bus)
☐ At other places (e.g. hotel, restaurant, bar, shop, museum)

A4. Which of the following (wireless) communication methods have you used?
(please tick all that apply)
☐ GSM / GPRS ☐ Infrared ☐ Bluetooth ☐ Wireless LAN (WLAN, WiFi)
☐ Don't know

Travel Information

A5. How many times do you travel (per year)?
☐ 0-1 time ☐ 2-5 times ☐ 6-10 times ☐ Above 10 times

A6. For what purposes do you travel?
☐ Business ☐ Private ☐ Both

AmbieSense Evaluation Questionnaire 3

Figure B.11: *Printed with permission of AmbieSense*^{TM}.

Before Travelling

A7. What information do you gather **before** travelling? (please tick all that apply)

- ☐ Food/ Restaurants
- ☐ Art/ Sculpture/ Crafts
- ☐ People and Culture
- ☐ Transportation (e.g. type, schedule)
- ☐ Accommodation (e.g. hotels, B&B)
- ☐ History
- ☐ Nature/ Countryside
- ☐ Event (e.g. concert, festivals)
- ☐ Nightlife
- ☐ Exchange rate and economics
- ☐ Research/Universities
- ☐ Architecture
- ☐ Political stability
- ☐ News
- ☐ Sites / Attractions
- ☐ Shopping
- ☐ Health / Sport Facilities
- ☐ Maps
- ☐ Weather
- ☐ Other ___________________

A8. How important are these kinds of information to you, **before** you travel?
(Please circle one number for each type.)

		Unimportant				**Important**
1	Food/ Restaurants	1	2	3	4	5
2	Art/ Sculpture/ Crafts	1	2	3	4	5
3	People and Culture	1	2	3	4	5
4	Transportation (e.g. type, schedule)	1	2	3	4	5
5	Accommodation (e.g. hotels, B&B)	1	2	3	4	5
6	History	1	2	3	4	5
7	Nature/ Countryside	1	2	3	4	5
8	Event (e.g. concert, festivals)	1	2	3	4	5
9	Nightlife	1	2	3	4	5
10	Exchange rate and economics	1	2	3	4	5
11	Research/Universities	1	2	3	4	5
12	Architecture	1	2	3	4	5
13	Political stability	1	2	3	4	5
14	News	1	2	3	4	5
15	Sites / Attractions	1	2	3	4	5
16	Shopping	1	2	3	4	5
17	Health / Sport Facilities	1	2	3	4	5
18	Maps	1	2	3	4	5
19	Weather	1	2	3	4	5
20	Other (specify if necessary)________________	1	2	3	4	5

A9. Where do you get the information from **before** you travel? (please tick all that apply)

- ☐ Internet
- ☐ Digital travel guide (e.g. CD ROM)
- ☐ Brochures
- ☐ Tourist Information Centre / Office
- ☐ Travel Agent
- ☐ TV
- ☐ Word of Mouth (e.g. friends)
- ☐ Bookstore
- ☐ Library
- ☐ Hotel reception
- ☐ Language guide
- ☐ Other ___________________

AmbieSense Evaluation Questionnaire 4

Figure B.12: *Printed with permission of AmbieSense^{TM}.*

While Travelling

A10. What information do you gather **while** travelling? (please tick all that apply)

- ☐ Food/ Restaurants
- ☐ Art/ Sculpture/ Crafts
- ☐ People and Culture
- ☐ Transportation (e.g. type, schedule)
- ☐ Accommodation (e.g. hotels, B&B)
- ☐ History
- ☐ Nature/ Countryside
- ☐ Event (e.g. concert, festivals)
- ☐ Nightlife
- ☐ Exchange rate and economics
- ☐ Research/Universities
- ☐ Architecture
- ☐ Political stability
- ☐ News
- ☐ Sites / Attractions
- ☐ Shopping
- ☐ Health / Sport Facilities
- ☐ Maps
- ☐ Weather
- ☐ Other ___________________

A11. How important are these kinds of information to you, **while** you travel?
(Please circle one number for each type.)

		Unimportant				Important
1	Food/ Restaurants	1	2	3	4	5
2	Art/ Sculpture/ Crafts	1	2	3	4	5
3	People and Culture	1	2	3	4	5
4	Transportation (e.g. type, schedule)	1	2	3	4	5
5	Accommodation (e.g. hotels, B&B)	1	2	3	4	5
6	History	1	2	3	4	5
7	Nature/ Countryside	1	2	3	4	5
8	Event (e.g. concert, festivals)	1	2	3	4	5
9	Nightlife	1	2	3	4	5
10	Exchange rate and economics	1	2	3	4	5
11	Research/Universities	1	2	3	4	5
12	Architecture	1	2	3	4	5
13	Political stability	1	2	3	4	5
14	News	1	2	3	4	5
15	Sites / Attractions	1	2	3	4	5
16	Shopping	1	2	3	4	5
17	Health / Sport Facilities	1	2	3	4	5
18	Maps	1	2	3	4	5
19	Weather	1	2	3	4	5
20	Other (specify if necessary)________________	1	2	3	4	5

A12. Where do you get the information from **while** travelling? (please tick all that apply)

- ☐ Internet
- ☐ Digital travel guide (e.g. CD ROM)
- ☐ Brochures
- ☐ Tourist Information Centre / Office
- ☐ Travel Agent
- ☐ TV
- ☐ Word of Mouth (e.g. friends)
- ☐ Bookstore
- ☐ Library
- ☐ Hotel reception
- ☐ Language guide
- ☐ Other ___________________

Figure B.13: *Printed with permission of AmbieSense^{TM}.*

A13. What information would you like to get when using a handheld device **while** travelling? (Please all that apply.)

- ☐ Food/ Restaurants
- ☐ Art/ Sculpture/ Crafts
- ☐ People and Culture
- ☐ Transportation (e.g. type, schedule)
- ☐ Accommodation (e.g. hotels, B&B)
- ☐ History
- ☐ Nature/ Countryside
- ☐ Event (e.g. concert, festivals)
- ☐ Nightlife
- ☐ Exchange rate and economics
- ☐ Research/Universities
- ☐ Architecture
- ☐ Political stability
- ☐ News
- ☐ Sites / Attractions
- ☐ Shopping
- ☐ Health / Sport Facilities
- ☐ Maps
- ☐ Weather
- ☐ Other ___________________

A15. What kind of transportation do you use while visiting an area? (please tick all that apply)

- ☐ Taxi
- ☐ Public transport (bus, train, underground)
- ☐ Car
- ☐ Other: ___________________

Figure B.14: *Printed with permission of AmbieSense^{TM}.*

IST 2001- 34244 AmbieSense.

Interviewer: _______**Date:** ______**Time:**______**Location:** _______**Questionnaire No:**___

Part 3: After Test

AmbieSense Mobile User Perspective

C1. Usage of the mobile device	**Difficult**				**Easy**
	1	2	3	4	5
C2. Usage of the system compared to my normal method for finding information	**Harder**				**Easier**
	1	2	3	4	5
C3. Coverage of my information needs by the system	**Inadequate**				**Adequate**
	1	2	3	4	5
C4. Amount of time spent getting familiar with AmbieSense	**Unsatisfying**				**Satisfying**
	1	2	3	4	5

C5. Would you be willing to pay for a mobile tourism system such as this?
(Please tick one box.)
☐ No ☐ Yes ☐Unsure

C6. If payment was required what payment methods would you prefer?
(Please tick all that apply)
☐ A time base ☐ A usage base ☐One off payment
☐ Other: (please specify)

C7. Would you like to rent or own the mobile device? (Please tick one box.)
☐Rent ☐ Buy ☐ Either

C8. If you choose to **rent** the mobile device where would you like to rent it from?
(Please tick all that apply.)
☐ Hotel ☐ Car Rental Station ☐ Airport
☐ Tourist Information ☐ Travel Agency ☐ Public Transport Office

Other: (please specify)

C9. If you choose to **buy** a mobile device or already own one, using this service would require a small program to be installed on your mobile / handheld device. How would you like it installed?
☐Preinstalled on mobile device ☐ Download from Internet ☐ Install on first use

C10. Which type of mobile device(s) do you have? (Please state the manufacturer and model.)

AmbieSense Evaluation Questionnaire 7

Figure B.15: *Printed with permission of AmbieSense*^{TM}.

AmbieSense Overall and Comments

C11. How much was the quality of your stay, in Sevilla, improved by using the system?	**Not at all**				**Very**
	1	2	3	4	5
C12. To what extent were your information needs fulfilled by the system?	**Not at all**				**Very**
	1	2	3	4	5
C13. How valuable do you think the system might be for other purposes other than your current one?	**Not at all**				**Very**
	1	2	3	4	5
C14. To what extent do you think you have benefited from the system?	**Not at all**				**Very**
	1	2	3	4	5

C15. Please rate the following features of AmbieSense.
(Please circle one number for each feature).

Travel Guide Application (Please answer if you did this application).

	Unimportant				Important	Not Seen
➢ Recommendation of sites and attractions	1	2	3	4	5	6
➢ Recommendation of restaurants	1	2	3	4	5	6
➢ Recommendation of hotels	1	2	3	4	5	6
➢ Recommendation of events	1	2	3	4	5	6
➢ Recommendation of shops	1	2	3	4	5	6

Map Application (Please answer if you did this application).

	Unimportant				Important	Not Seen
➢ Pictures of places	1	2	3	4	5	6
➢ Maps in general	1	2	3	4	5	6
➢ Maps with current location shown	1	2	3	4	5	6
➢ Directions	1	2	3	4	5	6
➢ Maps keeping detailed information	1	2	3	4	5	6

Figure B.16: *Printed with permission of AmbieSense^{TM}.*

C15. Would you use the AmbieSense system again?
☐Yes ☐No

C16. Would you recommend AmbieSense to a friend or relative?
☐Yes ☐No

C17. What did you like about the system?
(Please specify)______________________________

C18. What didn't you like about the system?
(Please specify)______________________________

C19. Overall, in summary, what are your general impressions of the system?
(You may wish to comment on performance, usability, information content for example.)

C20. What personal information would you provide to get content results from our system which matches you better as an individual? (e.g. your age group or topics you are interested in) (Please tick all those that apply)

- ☐ Name
- ☐ Address
- ☐ Age group
- ☐ Educational background
- ☐ Job category
- ☐ Interests and Hobbies
- ☐ Social network (e.g. friends)

End of Questionnaire

Thank you for taking part in this evaluation of the AmbieSense Technology, your time and effort are greatly appreciated.

AmbieSense Evaluation Questionnaire 9

Figure B.17: *Printed with permission of AmbieSense^{TM}.*

IST 2001- 34244 AmbieSense.

Your questionnaire will be kept anonymous. The information on this page will be kept separately from the previous pages. It will be retained for purposes of updating you with information if you wish.

C21. Would you like to stay in contact with AmbieSense in order to get information about further developments, tests and evaluations?

☐ No ☐ Yes -- If yes, please give us your name and postal address.
[The information will be handled confidentially and will not be disclosed to third parties]

Name: ______________________________
Address:
____
____
____

Email: ______________________________

End of Questionnaire

Thank you for taking part in this evaluation of the AmbieSense Technology, your time and effort are greatly appreciated.

AmbieSense Evaluation Questionnaire 10

Figure B.18: *Printed with permission of AmbieSense*^{TM}.

B.4 AmbieSense Seville September 2004 Questionnaire

This is the AmbieSense questionnaire that was used to collect data during a mobile user study in Seville in September 2004. Only the pre- and post-questionnaire are shown since no data from the search tasks has been presented in this thesis. For more information on the entire mobile study, please refer to [Göker and Myrhaug, 2007].

IST 2001- 34244 AmbieSense.

Part A: Before Tasks

Demographic Information

A1. Please answer a few general questions about yourself

Gender : ☐ Male ☐ Female

Age : ☐Under 20 ☐20 – 29 ☐30 – 39 ☐40 – 49 ☐50 - 59 ☐60 and over

Spoken Languages: ____

Computing Experience

A2. How often do you use the following devices, please answer those you have personally used and are familiar with?

1. Mobile Phone	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
2. PDA	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
3. Smart Phones (e.g. P900, Nokia 6600)	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily
4. Laptop	☐ Rarely	☐ Monthly	☐ Weekly	☐ Daily

A3. Where do you use handheld device(s)? (please tick all that apply)

☐ At home
☐ At work
☐ While on the move (e.g. plane, train, bus)
☐ At other places (e.g. hotel, restaurant, bar, shop, museum)

A4: Which of the following (wireless) communication methods have you used?

☐ GSM / GPRS ☐ Infrared ☐ Bluetooth ☐ Wireless LAN (WLAN, WiFi)
☐ Don't know

Travel Information

A5. How many times do you travel (between city or country) per year?
☐ 0-1 time ☐ 2-5 times ☐ 5-10 times ☐ Above 10 times

A6. For what purposes do you travel? (Please tick all that apply)
☐ Business ☐ Private

Before Travelling

Figure B.19: *Printed with permission of AmbieSense^{TM}.*

A7. What type of information do you gather **before** travelling? (please tick all that apply)

☐ Accommodation (e.g. hotels, B&B) ☐ Restaurants ☐ Sites/ Attractions ☐ Transport ☐ Maps	☐ Events (e.g. concert, festival) ☐ Shops ☐ Weather ☐ Other ___________________

A8. Where do you get the information from **before** you travel? (please tick all that apply)

☐ Internet ☐ Digital travel guide (e.g. CD ROM) ☐ Brochures ☐ Tourist Information Centre / Office ☐ Travel Agent	☐ Word of Mouth (e.g. friends) ☐ Bookstore ☐ Library ☐ Hotel reception ☐ Other ___________________

While Travelling

A9. What information do you gather **while** travelling? (please tick all that apply)

☐ Accommodation (e.g. hotels, B&B) ☐ Restaurants ☐ Sites/ Attractions ☐ Transport ☐ Maps	☐ Events (e.g. concert, festival) ☐ Shops ☐ Weather ☐ Other ___________________

A10. Where do you get the information from **while** you travel? (please tick all that apply)

☐ Internet ☐ Digital travel guide (e.g. CD ROM) ☐ Brochures ☐ Tourist Information Centre / Office ☐ Travel Agent	☐ Word of Mouth (e.g. friends) ☐ Bookstore ☐ Library ☐ Hotel reception ☐ Other ___________________

Figure B.20: *Printed with permission of AmbieSense*^{TM}.

A11. When you visit a region, how important are the following aspects for you? (Please circle the numbers on the scale)

	Unimportant				Important
• Food/ Restaurants	1	2	3	4	5
• Art/ Sculpture/ Crafts	1	2	3	4	5
• People and Culture	1	2	3	4	5
• Transportation (e.g. type, schedule)	1	2	3	4	5
• Accommodation (e.g. hotels, B&B)	1	2	3	4	5
• History	1	2	3	4	5
• Nature/ Countryside	1	2	3	4	5
• Event (e.g. concert, festivals)	1	2	3	4	5
• Nightlife	1	2	3	4	5
• Exchange rate and economics	1	2	3	4	5
• Research/Universities	1	2	3	4	5
• Architecture	1	2	3	4	5
• Political stability	1	2	3	4	5
• News	1	2	3	4	5
• Sites / Attractions	1	2	3	4	5
• Shopping	1	2	3	4	5
• Health / Sport Facilities	1	2	3	4	5

A12. What do you wish you could get when using a handheld device, **while** travelling?

____

____

Figure B.21: *Printed with permission of AmbieSense^{TM}.*

Part C: After Tasks

AmbieSense Mobile User Perspective

C1. Using the mobile device was	**Difficult** **Easy** 1 2 3 4 5
C2. Using this system compared to my normal method for finding information was	**Harder** **Easier** 1 2 3 4 5
C3. Using this system my information needs were covered	**Inadequately** **Adequately** 1 2 3 4 5
C4. Exploring the city with this system I felt	**Worse Prepared** **Better Prepared** 1 2 3 4 5

C5. Would you be willing to pay for a mobile tourism system such as this?
☐ No ☐ Yes

C6. If payment was required what payment methods would you prefer? (please tick all that apply)
☐ A time base ☐ A usage base ☐One off payment
☐ Sponsored by Advertisements
☐ Other: (please specify)
____

C7. Using this service would require a small program to be installed on your mobile / handheld device. How would you like it installed?

☐Preinstalled on mobile device ☐ Download from Internet ☐ Install on first use

C8. Which type of mobile device would you prefer?

☐ Mobile Phone ☐ PDA ☐ Smart Phones (e.g. P900, Nokia 6600) ☐ Laptop

C9. Are there any particular brands that you would prefer for your mobile device?
(tick all that apply)

☐ Apple	☐ IBM	☐ Psion
☐ Casio	☐ Ipaq	☐ Samsung
☐ Compaq	☐ Motorola	☐ Sharp
☐ Dell	☐ Nokia	☐ Siemens
☐ Fujitsu	☐ Packard Bell	☐ Sony
☐ HandSpring	☐ Palm	☐ Sony Ericsson
☐ Hewlett Packard	☐ Panasonic	☐ Toshiba

☐ Other: (please specify): ____________________________________

Figure B.22: *Printed with permission of AmbieSense^{TM}.*

AmbieSense Overall and Comments

C10. Using this system my stay in Seville would be	**No Different**				**Improved**
	1	2	3	4	5
C11. This system provided for my information needs	**Inadequately**				**Adequately**
	1	2	3	4	5
C12. If I had come to the city/ airport with a different purpose the system would be	**Of No Use**				**Useful**
	1	2	3	4	5
C13. From the short time I used the system I felt the system would	**Not Benefit Me**				**Benefit Me**
	1	2	3	4	5

C14. Please rate the following features of this system.

	Unimportant				**Important**
• Recommendation of sites and attractions	1	2	3	4	5
• Recommendation of restaurants	1	2	3	4	5
• Recommendation of hotels	1	2	3	4	5
• Recommendation of events	1	2	3	4	5
• Recommendation of shops	1	2	3	4	5

C15. I would use the AmbieSense system again?
☐Yes ☐No

C16. I would recommend AmbieSense to a friend or relative?
☐Yes ☐No

C17. What personal information would you provide to get better and more personal recommendations (e.g. where to go or in which hotel to stay)?
(Please check those that apply)

☐ Name	☐ Educational background
☐ Post Address	☐ Job category
☐ Email Address	☐ Interests and Hobbies
☐ Age group	☐ Social network (e.g. friends)

Figure B.23: *Printed with permission of AmbieSense^{TM}.*

IST 2001- 34244 AmbieSense.

C18. What did you like or not like about this system? (Comment on performance, usability, information content for example)

____

____

____

____

____

____

____

____

____

C19. Would you like to stay in contact with this system, AmbieSense, in order to get information about further developments, tests and evaluations?

☐ No ☐ Yes If yes, please give us your name and postal address.
[The information will be handled confidentially and will not be disclosed to third parties]

Address:
____

____

____

Email: ____

End of Questionnaire
Thank you for taking part in this evaluation of the AmbieSense Technology, your time and effort are greatly appreciated.

Figure B.24: *Printed with permission of AmbieSense*^{TM}.

C

Questionnaire of the Personalization Consortium

The following three pages show questionnaire and results from the web survey that was conducted by the Personalization Consortium in 2000. In particular, results from question 6 and 7 have been highlighted in section 3.3.3. These questions asked about peoples' willingness to provide personal information depending on a website providing personalised or non-personalised services. Results from these two questions have inspired this work to contribute similar questions to the AmbieSense questionnaires as described in section 3.3.3[1] Results from the two Personalization Consortium questions were confirmed by the results from the mobile studies conducted by AmbieSense.

[1]Questionnaires are provided in section B.3 and B.4 in the previous appendix.

Personalization & Privacy Survey

1.	How many hours a week, on average, do you surf the web (not including e-mail)?	Numbers	Percentage
	0-1	122	3%
	2-7	1548	34%
	8+	2850	63%
	Total	4520	100%

2.	How much have you spent online within the last 6 months?	Numbers	Percentage
	$0	662	15%
	$1 - $100	1657	37%
	$100+	2194	49%
	Total	4513	100%

3.	Are you currently a registered user with any web sites (providing name and email address)?	Numbers	Percentage
	Yes	4231	94%
	No	274	6%
	Total	4505	100%

4.	Have you provided personal information to any web sites (address, phone number, hobbies/interests)?	Numbers	Percentage
	Yes	4231	97%
	No	274	3%
	Total	4511	100%

5. Please indicate your level of agreement with the following statements:	1 Strongly disagree	2 Disagree	3 Neither agree or disagree	4 Agree	5 Strongly Agree
1. Before registering on web sites I always read the privacy statement.	9% (386)	18% (829)	22% (1008)	33% (1473)	18% (820)
2. Most privacy statements on web sites are simple and easy to understand.	12% (542)	24% (1070)	26% (1189)	30% (1364)	8% (343)
3. A privacy statement is necessary for me to share personal information.	5% (227)	11% (517)	24% (1096)	31% (1420)	27% (1231)
4. I find it helpful and convenient when a web site remembers basic information about me (e.g., my name and address).	4% (179)	5% (224)	17% (753)	42% (1891)	31% (1415)
5. I find it helpful and convenient when a web site remembers more personal information about me (e.g., my preferred colors, music or delivery options).	7% (324)	13% (588)	30% (1363)	31% (1389)	19% (838)
6. Banner ads and "pop ups" are an invasion of my privacy.	8% (367)	21% (950)	35% (1560)	16% (736)	19% (867)
7. Online solicitations (offers) from the web site hinder my shopping experience.	7% (318)	22% (1014)	40% (1790)	19% (850)	11% (514)
8. I am willing to give information about myself in order to receive an online experience truly personalized for me.	4% (198)	11% (507)	33% (1510)	41% (1837)	10% (435)
9. I have more control shopping online than I do offline.	9% (424)	20% (902)	38% (1718)	22% (993)	10% (441)
10. It bothers me when a web site requests personal information I've already provided (e.g., my mailing address).	5% (206)	9% (429)	23% (1042)	38% (1724)	24% (1102)

6.	What pieces of information would you provide a web-shopping site that DOES NOT provide any features personalized for you? (Please check all that apply.)		
	Name	3751	85%
	Address	2642	60%
	Credit Card Number	845	19%
	Income	855	19%
	Job Title	1416	32%
	Phone Number	1262	29%
	Hobbies/Interests	2222	51%
	Social Security Number	295	7%
	Mother's Maiden Name	607	14%
	E-mail Address	3856	88%

7.	What pieces of information would you provide a web site that used the information that you gave them to personalize/customize your experience? (Please check all that apply.)		
	Name	4266	96%
	Address	3600	81%
	Credit Card Number	973	22%
	Income	1508	34%
	Job Title	2235	50%
	Phone Number	1988	45%
	Hobbies/Interests	3426	76%
	Social Security Number	270	6%
	Mother's Maiden Name	1001	22%
	E-mail Address	4232	95%

8.	Do you know what a web browser cookie is?	Numbers	Percentage
	Yes	3463	77%
	No	1035	23%
	Total	4498	100%

9.	If yes, please indicate which of the following statements apply to you. (Please check all that apply.)	Numbers	Percentage
	I know how to reject cookies.	1532	43%
	I generally accept cookies.	2204	62%
	A warning pop up comes on before I am given cookies.	888	25%
	I erase my cookies from my hard drive periodically.	1881	53%
	Cookies are an invasion of my privacy.	669	19%

D

Lucene IR Model

The following formula represents the information retrieval scoring algorithm of the Lucene search library as it has been used in this thesis. As this appendix only describes only the high level elements of the formula, more information about the underlying retrieval model and Lucene's technical differences are provided in section 2.4. More detail about the application of the forumla can be found in chapter 5 and 6.

$$Score_{IRModel} = \sum_{t\epsilon q}\left(\frac{\sqrt{|t\epsilon q|} * (log(\frac{|d|}{|df+1|}) + 1)}{\sqrt{\sum_{t\epsilon q}(\sqrt{|t\epsilon q|} * (log(\frac{|d|}{df+1}) + 1))^2}} * \frac{\sqrt{|t\epsilon d|} * (log(\frac{|d|}{|df+1|}) + 1)}{\sqrt{|d|}}\right) * \frac{|t\epsilon d \wedge t\epsilon q|}{|t\epsilon q|} \tag{D.1}$$

$\lvert t\epsilon q\rvert$	Query term frequency
$\lvert t\epsilon d\rvert$	Document term frequency
$\sqrt{\lvert t\epsilon d\rvert}$	Normalised document term frequency
$\sqrt{\lvert t\epsilon q\rvert}$	Normalised query term frequency
$\lvert d\rvert$	Number of documents in the index

df	Document frequency that states in how many documents a term occurred
$log(\frac{\|d\|}{\|df+1\|})+1$	Inverse document frequency (IDF)
$\frac{1}{\sqrt{\sum_{t\epsilon q}(\sqrt{\|t\epsilon q\|}*(log(\frac{\|d\|}{df+1})+1))^2}}$	Standard cosine normalisation for query terms
$\frac{1}{\sqrt{\|t\epsilon d\|}}$	Document term normalisation
$\frac{\|t\epsilon d \wedge t\epsilon q\|}{\|t\epsilon q\|}$	Coordination level matching that boosts terms based on their level of co-occurrence in query and document

E

The Reuters KALENDS Event Collection

E.1 Overview

For this thesis, Reuters Group plc provided *KALENDS*, a content corpus consisting of a collection of 10500 entertainment event content items. Reuters is a global information and news company that is mainly aimed to provide information to professionals in finance, media and corporate markets with a major focus (>90%) on finance services[1]. The Reuters KALENDS Event content collection is one of Reuters information services. At the beginning of this study, one set of the content collection was provided to the student in support for this research. In the meantime, the KALENDS product has changed from an entertainment event content service into a service about future financial events [2]. The next section presents an overview to the types of data that a KALENDS event may contain and a number of typical KALENDS event examples.

E.2 KALENDS

This section first describes the formal XML schema that defines the data format of KALENDS followed by a list of selected KALENDS events in XML.

[1]Information obtained from `http://www.reuters.com`, accessed April 14, 2008

[2]More information about the current KALENDS service can be obtained from `http://www.kalends.com/dotcom/home.htm`, accessed April 14, 2008

E.2.1 XML schema

The XML Schema[3] in figure E.1 depicts the elements of the KALENDS event content collection. Not all elements are used for every single event content item. Note that KALENDS was not provided with an XML schema hence the schema was generated ad-hoc from the collection for the purpose of analysis and documentation. The figures in this section graphically highlight the elements of the schema. Due to its ad-hoc generation, data elements re-appear in different pattern. It is not the aim to focus on the precise arrangement of these content element types. Rather, this section describes the various content elements since they inspired context modelling from the viewpoint of the event content domain. Figure E.1 shows the main content element types of a KALENDS event. A KALENDS event content item consists of a *title*, a *description*, an unique identifier (*sourceeventid*), temporal information (*startdate*, *endddate*, *allday*, *uncertainperiod*), contact information (*phonenum*, *faxnum*, *emailaddress*), *category* information, *location* information and information about the *organizer* as well as people (*person*) that are involved in the event (e.g. performers or producers). The bottom part of figure E.1 shows a tree structure of the various ways after which content elements are combined in the data. Rather than to focus on this structure, it is preferred to focus more on its entities; the content elements of a KALENDS event. The figure does not contain all details. More detailed information about the elements for category (1), organization (2), location (3) and person (4) are factored out in separate diagrams depicted in the figures E.2 - E.5 that are further described below. As shown in figure E.2, an event category consists of an unique category identifier and an optional number of fields that further specifies the content category. The unique category identifier was used for the two experiments to select events from different categories. A very small selection of jazz and comedy events where selected for the user study on contextual usefulness (see chapter 4). For the mobile user experiment in chapter 6, a larger amount of dance and musical events where selected from KALENDS. A KALENDS event also models information about the associated organisation(s) (see figure E.3) of an entertainment event. These are modelled by a name and a number of roles. Similarly, the 'person' content element models the people that are involved in an event performance (such as artists/performers/musicians

[3]An XML Schema is a formal description of data that defines the data vocabulary and the rules after which this data is organised.

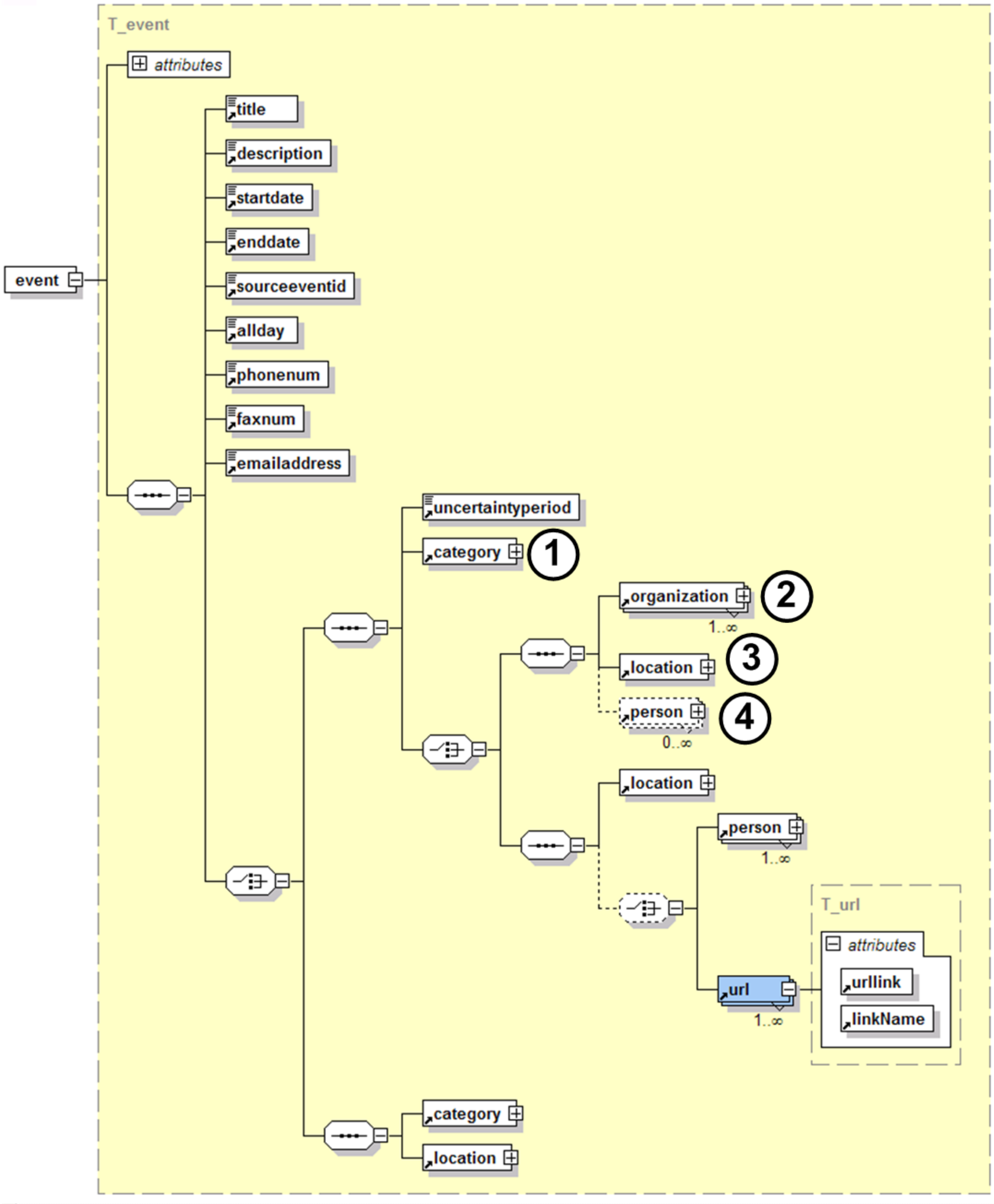


Figure E.1: *XML Schema of KALENDS event. Details are provided in separate diagrams; (1) category in figure E.2, (2) organization in figure E.3, (3) location in figure E.5, (4) person in figure E.4*

or directors/producers) with their nick and full name as well as a series of different roles. Information about organisations and persons does not always exist and represent more

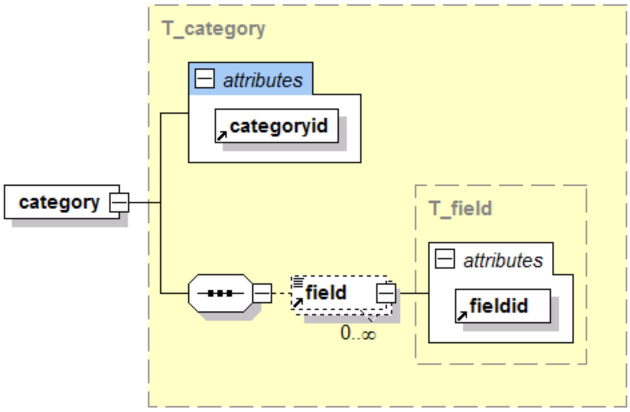


Figure E.2: *XML Schema of category part of KALENDS event*

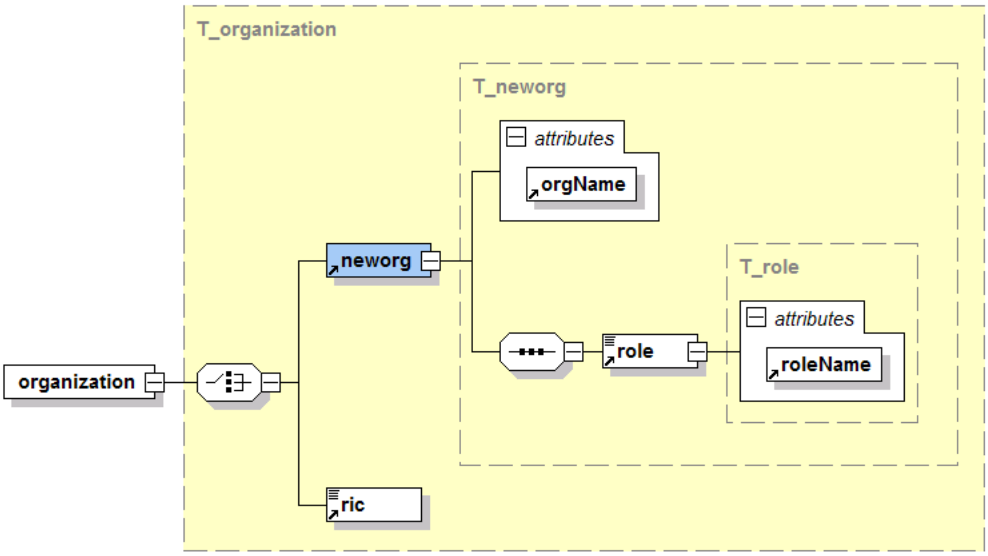


Figure E.3: *XML Schema of organization part of KALENDS event*

optional content elements. Every KALENDS event is associated with a location. Figure E.5 shows a number of different combinations of sub elements. Generally, a location is described by information about *street*, *postcode*, province (*provinceName*, *provinceCode*), country (*countryISO2code*), *city* and points of interest (*pointofinterest*). The point of interest models more detailed information about the place of performance with place name, URL, phone, fax, email, an image, additional description, travel information and available services (e.g. disabled access).

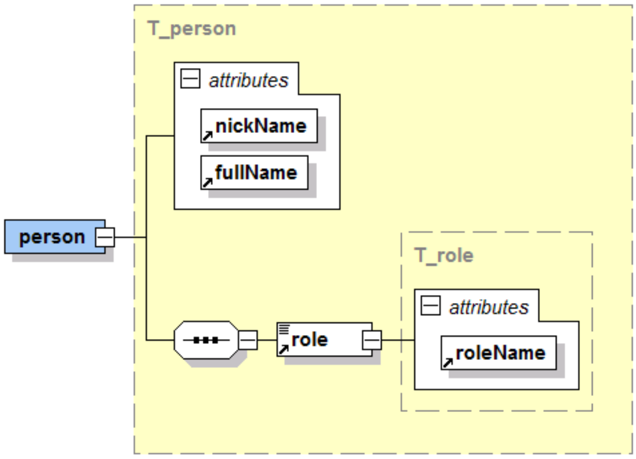


Figure E.4: *XML Schema of person part of KALENDS event*

After providing the formal structure of a KALENDS event in this subsection, the next subsection presents a selective number of specific examples from the collection.

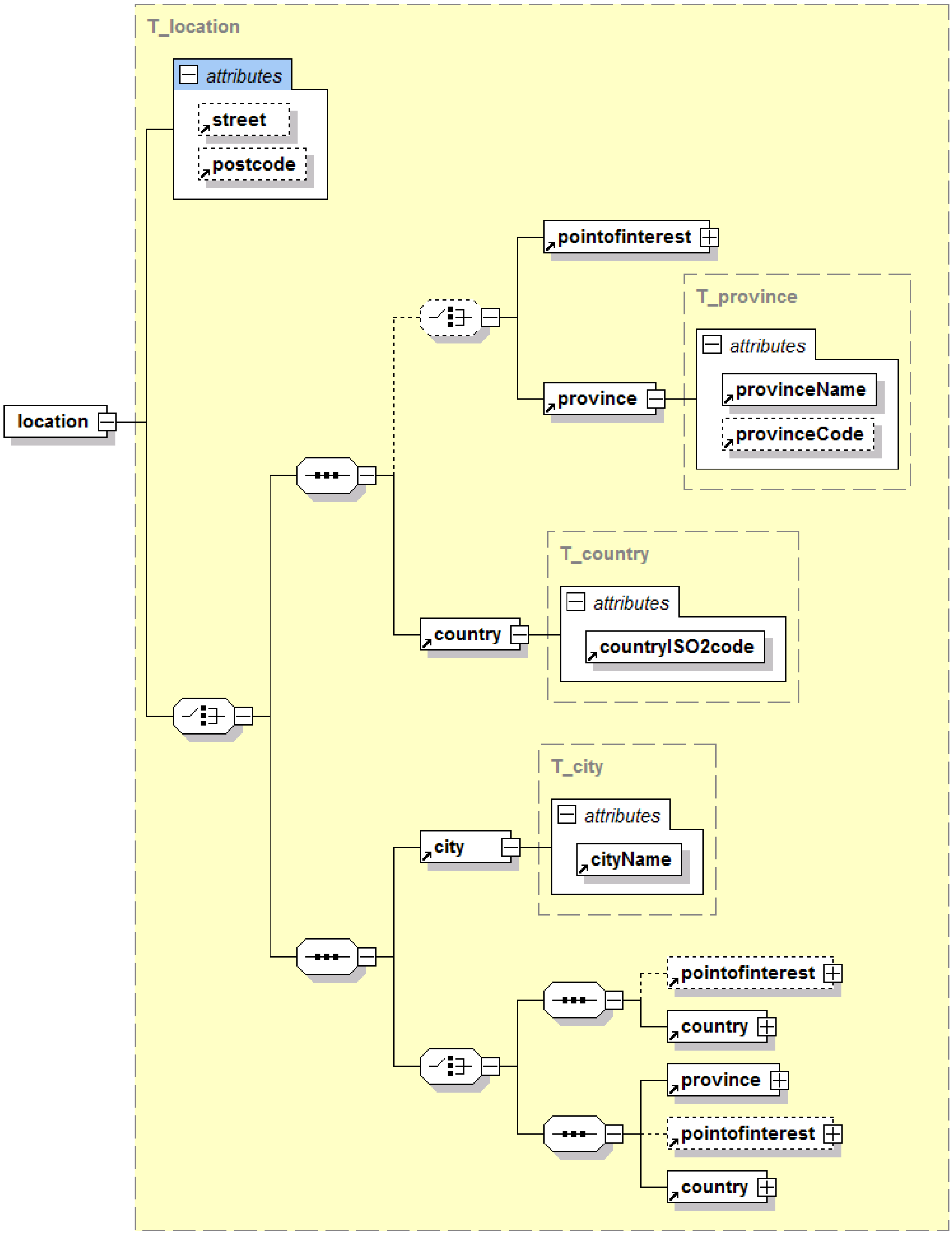


Figure E.5: *XML Schema of location part of KALENDS event*

E.2.2 Selective KALENDS Examples

The following three listings are examples of typical event content items from the KALENDS collection. These three content items have been selected since they have all been used in either one of the two user experiments - the study on contextual usefulness that is reported in chapter 4 and the mobile user study that is described in chapter 6.


```xml
<event workflowID="3">
    <title>Ken Dodd - The Happiness Show</title>
    <description>Ken Dodd is much more than a comedian. He is a comedy
        genius and showbiz legend whose humour has made him one of Britain's
        best-loved entertainers. For his Diddymen, jam-butty mines and black
        pudding plantations, the Professor of Giggleology and Master of
        Applied Tickleology has been awarded The British Comedy Awards
        highest accolade - the Lifetime Achievement Award. Come and join the
        King of Comedy for more quick-fire gags than you can shake a tickle-
        stick at!</description>
    <startdate>2001-11-04T00:00Z</startdate>
    <enddate>2001-11-04T22:59Z</enddate>
    <sourceeventid>L249820425</sourceeventid>
    <allday>false</allday>
    <phonenum>+44 (0)1908 606 090</phonenum>
    <faxnum />
    <emailaddress>info@mktgc.co.uk</emailaddress>
    <uncertaintyperiod>0</uncertaintyperiod>
    <category categoryid="ENTF" />
    <location street="900 Midsummer Boulevard" postcode="MK93NZ">
        <pointofinterest
            poiName="Milton Keynes Theatre"
            poiPhone="+44 (0)1908 606 090"
            poiURL="http://www.mktgc.co.uk"
            poiEmail="info@mktgc.co.uk">
            <description>A new apound;30 million theatre is being built -work
                 started in 1997 and it opened in 1999. The name was chosen
                as the result of an extensive local survey. Flexible seating
                capacity 950-1400. An ATG member.</description>
            <poiservices>Theatre Tokens, Infra-red system,
            Wheelchair access, Disabled toilets</poiservices>
            <image
                imageType="Exterior Photo"
                imageFile="http://www.dynamiclisting.com/uktw/venues/ex754.
                    jpg" />
        </pointofinterest>
        <country countryISO2code="GB" />
    </location>
</event>
```


Listing E.1: *KALENDS event example used for the user study on contextual usefulness*


```xml
<event workflowID="3">
    <title>Alice the Musical</title>
    <description> An exciting new version of the timeless Wonderland story
        for all the family. The Mad Hatter, the White Rabbit and many other
        magical characters come to life in this modern, musical version of
        the classic tale. With vibrant, catchy pop music, great humour, super
         costumes and a multi-talented cast, this is a dazzling journey into
        a world of pure fantasy.</description>
    <startdate>2001-10-25T00:00Z</startdate>
    <enddate>2001-10-25T22:59Z</enddate>
    <sourceeventid>L01952624222</sourceeventid>
    <allday>false</allday>
    <phonenum>+44 (0)1923 771 542</phonenum>
    <faxnum>+44 (0)1923 710 121</faxnum>
    <emailaddress>watersmeet.theatre@threerivers.gov.uk</emailaddress>
    <uncertaintyperiod>0</uncertaintyperiod>
    <category categoryid="ENTTM">
```

```
        <field fieldid="PERFORM">13:00 and 15:30</field>
        <field fieldid="PRICES">??6.5</field>
    </category>
    <location street="High Street" postcode="WD3 1HJ">
        <city cityName="Rickmansworth"/>
        <pointofinterest
            poiName="Watersmeet Theatre"
            poiFAX="+44 (0)1923 710 121"
            poiPhone="+44 (0)1923 771 542"
            poiEmail="watersmeet.theatre@threerivers.gov.uk">
            <description>Watersmeet Youth Theatre include classes for three
                junior age groups. Call for details.</description>
            <poiservices>Wheelchair access, Disabled toilets, Bar</
                poiservices>
        </pointofinterest>
        <country countryISO2code="GB"/>
    </location>
</event>
```

Listing E.2: *KALENDS event example used for mobile experiment*

```
<event workflowID="3">
    <title>The Snowman</title>
    <description>Musical adaptation of Raymond Briggs‘s magical story. When
         a little boy builds a snowman in his garden, little does he know
        what Christmas Eve has in store for him. That night, he can‘t sleep
        , so he opens the front door and amazingly the snowman has come to
        life. The boy shows him into a cozy home and, in return, the snowman
         introduces the boy to his wintry world. They join hands and fly
        up into the night. As they fly, other snowman from the surrounding
         gardens take off to join them as they travel to the Snowman‘s Ball
         where Father Christmas gives the boy a present of a scarf. Next
         morning, having returned home, the boy is saddened to find the
        snowman has melted. His tears turn to joy however when he discovers
        the scarf and he remembers his fantastic journey.</description>
    <startdate>2001-12-11T00:00Z</startdate>
    <enddate>2002-01-13T22:59Z</enddate>
    <sourceeventid>L01796820933</sourceeventid>
    <allday>false</allday>
    <phonenum>+44 (0)20 7314 8800</phonenum>
    <faxnum/>
    <emailaddress/>
    <uncertaintyperiod>0</uncertaintyperiod>
    <category categoryid="ENTTM">
        <field fieldid="OPENDATE">12 Dec 2001</field>
        <field fieldid="PERFORM">11:00 (Dec
             15,16,18,19,22,23,26,27,28,29,30, Jan 2002 2,3,5,6,8,9,10,11,12)
            ,14:30 (Dec 12,15,16,18,19,22,23,26,27,28,29,30, Jan
             2002 2,3,4,5,6,8,9,10,11,12,13),16:00 (Dec 21), 19:00 (Dec
             11,13,14,15,16,18,22,23,28,29, Jan 2002 4,5,6,11,12)</field>
        <field fieldid="PRICES">??8.5 to ??27.5</field>
    </category>
    <location street="Portugal Street" postcode="WC2A 2HT">
        <pointofinterest
            poiName="Peacock Theatre"
            poiPhone="+44 (0)20 7314 8800"
            poiURL="http://www.sadlers-wells.com">
            <alias>Royalty Theatre</alias>
            <description>Renamed the Peacock Theatre Autumn 1996 in honour
                of benfactor Michael Peacock, formerly the Royalty Theatre
                . To be a temporary home (1996 to autumn 1998) to Sadler‘s
                 Wells during their refurbishment. Leasholders: The London
                 School of Economics. This site was originally home to the
                 London Opera House, built in 1911 and seating over 2600.
                The theatre was renamed in 1916 as the Stoll Theatre (after
```

```
                    purchase by Oswald Stoll) which was demolished in 1957. As
                    part of the redevelopment of the area as an office block
                    the Royalty was built, it opened in 1960. Renovation during
                    1996/7. Now 1037 seats. Wheelchair access possible (not bars
                    or toilets at present). Society of London Theatre member.</
                    description>
            <poiservices>Theatre Tokens</poiservices>
            <image
                imageType="Seating Plan"
                imageFile="http://www.dynamiclisting.com/uktw/venues/sp222.
                    gif"/>
            <railinfo>Holborn (LT)</railinfo>
            <travelinfo>Bus: High Holborn 8,19,38,22B,25,188,501, Kingsway
                    1,68,91,168,171,188,501,505,521,X68, Aldwych/Strand
                    4,11,15,23,26,76,171A,341</travelinfo>
        </pointofinterest>
        <country countryISO2code="GB"/>
    </location>
    <person fullName="Raymond Briggs" nickName="Raymond Briggs">
        <role roleName="Book by"/>
    </person>
    <person fullName="Howard Blake" nickName="Howard Blake">
        <role roleName="Music"/>
    </person>
    <person fullName="Howard Blake" nickName="Howard Blake">
        <role roleName="Lyrics"/>
    </person>
    <person fullName="Robert North" nickName="Robert North">
        <role roleName="Choreographer"/>
    </person>
</event>
```

Listing E.3: *KALENDS event example used for mobile experiment*

F

Situations and Tasks for the User Experiment on Contextual Usefulness

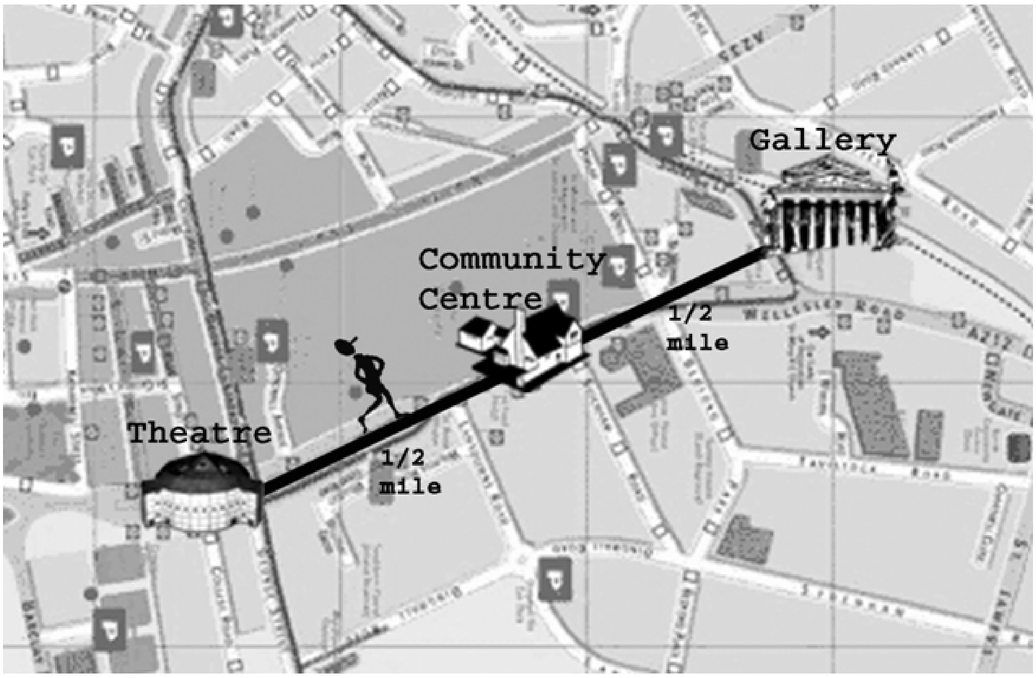
The following 6 pages contain the original handouts for the user experiment on contextual usefulness that is reported in chapter 4. During the experiment, these handouts were provided to participants and used for information and for recording feedback.

Festival in Town

This experiment investigates into how different situations influence your judgement of usefulness for event information. Your participation is very important and greatly appreciated. All answers are confidential. Thank you.

Background Scenario:

There is a festival in your town with a number of ***jazz and comedy events.*** The festival is held over ***3 days*** from ***Monday to Wednesday***. There are ***3 places*** where all the main events take place; the **Theatre**, the **Community Centre**, and the **Gallery** (see map below). The distance between each place is approximately ½ mile and they are arranged next to each other like in the following map.



During the festival, all roads between these places are closed for traffic because of numerous street performances and crowded roads. **The only way to get to each of these places is by walking**, in case you find one event useful and want to see the performance. The shortest route is indicated as a bold black line.

Your Task:

Baring in mind this background scenario, as a potential audience please rate the usefulness of 4 events from the provided event collection (page 2) imagining yourself in each of the following situations (page 3-6).

FESTIVAL EVENT CALENDAR

Event no:	Jazz Event 1
Title	Terry Lightfoot's Jazzmen - The Special Magic of Louis Armstrong
Description:	Terry Lightfoot first fronted a band in 1955, made his first major concert appearance at London's Royal Festival Hall in 1956 and quickly became one of the foremost leaders of 'Trad Jazz'. Today his band's repertoire has incorporated swing, blues and modern influences without ever losing his roots with traditional jazz.
Where?	Theatre
When?	Monday, 8 pm

Event no:	Jazz Event 2
Title	Piccadilly Dance Orchestra
Description:	Founded by Michael Law this 12 piece group have recreated as closely as possible the wonderful sounds and rhythms of the golden age of dance orchestras in the 20s, 30s and 40s. Music from the 20s to the 40s recreating the jazz age for today.
Where?	Community Centre
When?	Tuesday, 8 pm

Event no:	Comedy Event 1
Title	Johnny Vegas
Description:	Johnny Vegas is without doubt the hottest, most exciting and most talked about comedian on the live comedy circuit today. Johnny's unique abilities as a live performer are the stuff of legend. With his trade mark potter's wheel and Butlin's style singalongs, Johnny has created a sensation wherever he has played, winning the acclaimed Festival Critics' Award at the 1997 Edinburgh Festival, and being nominated for the Perrier Award.
Where?	Community Centre
When?	Monday, 8 pm

Event no:	Comedy Event 2
Title	Ken Dodd - The Happiness Show
Description:	Ken Dodd is much more than a comedian. He is a comedy genius and showbiz legend whose humour has made him one of Britain's best-loved entertainers. For his Diddymen, jam-butty mines and black pudding plantations, the Professor of Giggleology and Master of Applied Tickleology has been awarded The British Comedy Awards highest accolade - the Lifetime Achievement Award. Come and join the King of Comedy for more quick-fire gags than you can shake a tickle-stick at!
Where?	Gallery
When?	Wednesday, 8 pm

Situation 1: It is Monday morning. You are standing in front of the Theatre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 2: It is Wednesday morning. You are standing in front of the Gallery (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 3: It is Tuesday morning. You are standing in front of the Gallery (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 4: It is Tuesday morning. You are standing in front of the Theatre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 5: It is Wednesday morning. You are standing in front of the Community Centre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 6: It is Wednesday morning. You are standing in front of the Theatre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 7: It is Wednesday morning. You are standing in front of the Theatre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 8: It is Tuesday morning. You are standing in front of the Community Centre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 9: It is Monday morning. You are standing in front of the Theatre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 10: It is Tuesday morning. You are standing in front of the Community Centre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 11: It is Wednesday morning. You are standing in front of the Gallery (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 12: It is Wednesday morning. You are standing in front of the Community Centre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 13: It is Monday morning. You are standing in front of the Gallery (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 14: It is Monday morning. You are standing in front of the Community Centre (see map). You'd like to visit a jazz music event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 15: It is Tuesday morning. You are standing in front of the Gallery (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 16: It is Monday morning. You are standing in front of the Gallery (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 17: It is Monday morning. You are standing in front of the Community Centre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

Situation 18: It is Tuesday morning. You are standing in front of the Theatre (see map). You'd like to visit a comedy performance event. Please rate the following information about 4 events according to its usefulness in this particular situation.

Event Name	Information useful in this situation?					
	Not at all					Very much
Terry Lightfoot's Jazzmen	1	2	3	4	5	6
Piccadilly Dance Orchestra	1	2	3	4	5	6
Johnny Vegas	1	2	3	4	5	6
Ken Dodd - The Happiness Show	1	2	3	4	5	6

In order to classify the data, please answer the following questions about yourself:

Gender: ☐ Male ☐ Female

Age : ☐ Under 18 ☐ 18 – 29 ☐ 30 – 39 ☐ 40 – 49 ☐ Over 49

Thank you

G Situations and Tasks for the Mobile User Experiment

The handout consisted of three parts; the pre-questionnaire, the search tasks and the post-questionnaire. Training tasks and experiment tasks were presented to participants in counterbalanced order as shown in table G.1. Note that each training task was performed at the same location as the first experiment task. This was done for convenience to be able to perform the entire experiment with only a single change of location per participant. Training tasks did not reference any location in its description which means it was possible to use them interchangeably at both experiment locations.

Participant	**Training Task**	**1st Task**	**2nd Task**	**3rd Task**	**4th Task**
1	NP	T1 with NP	T1 with P	T2 with NP	T2 with P
2	P	T1 with P	T1 with NP	T2 with P	T2 with NP
3	NP	T2 with NP	T2 with P	T1 with NP	T1 with P
4	P	T2 with P	T2 with NP	T1 with P	T1 with NP
5	NP	T1 with NP	T1 with P	T2 with NP	T2 with P
6	P	T1 with P	T1 with NP	T2 with P	T2 with NP
7	NP	T2 with NP	T2 with P	T1 with NP	T1 with P
8	P	T2 with P	T2 with NP	T1 with P	T1 with NP
9	P	T1 with NP	T1 with P	T2 with NP	T2 with P
10	NP	T1 with P	T1 with NP	T2 with P	T2 with NP
11	P	T2 with NP	T2 with P	T1 with NP	T1 with P
12	NP	T2 with P	T2 with NP	T1 with P	T1 with NP
13	P	T1 with NP	T1 with P	T2 with NP	T2 with P
14	NP	T1 with P	T1 with NP	T2 with P	T2 with NP
15	P	T2 with NP	T2 with P	T1 with NP	T1 with P
16	NP	T2 with P	T2 with NP	T1 with P	T1 with NP
17	NP	T1 with NP	T1 with P	T2 with NP	T2 with P

Table G.1: *Order of training task and experiment task 1 (T1) and task 2 (T2) with personalised (P) and non-personalised (NP) system for each of the 17 participants.*

Table G.2 shows demographics (gender and age groups) from the 17 participants of the mobile user experiment. All participants are kept anonymous, however, their numbers (#) are used throughout the tables G.3, G.4, G.5 and G.6 to allow data to be brought into relation. Table G.3 presents individual results from the familiarity questions of the

#	Gender	Age group
1	male	18-29
2	female	18-29
3	male	18-29
4	male	30-39
5	male	30-39
6	female	30-39
7	male	18-29
8	female	18-29
9	female	18-29
10	male	18-29
11	female	30-39
12	male	30-39
13	male	18-29
14	female	Over 49
15	female	18-29
16	male	18-29
17	female	30-39

Table G.2: *Demographics (gender and age) for each of the 17 participants. Participant numbers (#) match with those from the next 4 tables below.*

#	PC	PDA	Mobile phone	Paper map	E-map	Search engine	Aberdeen
1	4	3	4	2	3	3	3
2	4	3	4	3	3	4	3
3	4	1	4	3	3	4	3
4	3	1	3	3	2	3	3
5	3	1	3	3	3	3	3
6	4	4	4	3	4	4	3
7	4	3	3	3	2	4	3
8	4	2	3	3	2	3	4
9	4	1	4	1	2	4	3
10	4	4	4	4	4	4	2
11	3	3	4	4	4	4	4
12	4	4	4	4	4	4	4
13	4	1	2	2	2	3	2
14	4	2	3	4	3	4	4
15	4	1	4	3	3	4	3
16	4	4	4	2	3	4	3
17	4	3	3	3	3	4	2

Table G.3: *Mobile experiment pre-questionnaire data about participants familiarity with PCs, PDA's, mobile phones, paper and electronic maps, search engines and the city centre of Aberdeen.*

pre-questionnaire. Questions were asked about how familiar participants were with PCs, Personal Digital Assistants (PDA's), mobile phones, maps (paper and electronic), search engines and the city centre of Aberdeen. Table G.4 below shows individual results from

#	Search printed events	Search electronic events	Attend events
1	3	3	4
2	2	3	2
3	3	4	2
4	2	3	2
5	3	2	3
6	2	4	3
7	2	4	3
8	2	3	2
9	2	2	2
10	2	2	2
11	4	4	3
12	3	4	2
13	1	2	1
14	3	4	4
15	1	4	2
16	2	4	2
17	3	3	2

Table G.4: *Mobile experiment pre-questionnaire data about participants event search behaviour and frequency of attendance.*

participants' event search behaviour and frequency of attendance. Results from these questions are summarised in chapter 6 in figure 6.5 and figure 6.6 on page 140. Table G.5 contains individual results from participants' post-questionnaire. They were asked about the suitability of the task situations, the suitability of the experiment locations, the level of interest / number / ease of the tasks, the level of interest in the events and the overall usability of the software. Furthermore, they gave ratings on which system they thought performed better – this data is provided in table G.6. Results from the post-questionnaires are shown in chapter 6 in figure 6.16 and figure 6.17 on page 158.

#	Software useable	Tasks interesting	Situations suitable	Tasks easy	Events interesting	Task number	Locations suitable
1	4	3	3	4	4	4	3
2	4	3	4	3	2	3	4
3	4	4	4	4	3	3	4
4	4	4	3	4	4	4	4
5	4	4	3	4	4	2	4
6	4	4	2	4	4	4	1
7	4	3	4	2	4	3	4
8	3	3	4	3	2	-	3
9	4	3	3	3	3	4	3
10	4	3	4	4	4	3	4
11	4	4	4	4	4	4	4
12	4	3	4	4	4	3	3
13	4	4	3	3	3	4	3
14	3	3	4	4	4	4	4
15	4	4	4	4	4	4	4
16	3	2	3	3	2	3	3
17	3	3	3	3	2	3	3

Table G.5: *Mobile experiment post-questionnaire data about participants event search behaviour and frequency of attendance.*

#	NP better than P	P better than NP	Systems are equal
1	2	4	2
2	4	3	2
3	2	4	1
4	1	4	1
5	1	1	3
6	1	1	4
7	1	4	1
8	3	2	1
9	1	3	1
10	1	2	1
11	2	2	3
12	1	4	1
13	2	3	1
14	3	2	1
15	1	1	4
16	2	3	1
17	1	1	4

Table G.6: *Mobile experiment post-questionnaire data about the system.*

On the following 9 pages, the original handouts of the mobile user experiment are shown.

Festival in Town

Part 2 - A mobile Study

1. Pre Questionnaire

This questionnaire contains a few questions which will help categorising experimental results. All information collected in this study remains anonymous.

1.1 In order to classify your data, please answer the following questions about yourself:

Gender: ☐ Male ☐ Female

Age : ☐ Under 18 ☐ 18 – 29 ☐ 30 – 39 ☐ 40 – 49 ☐ Over 49

1.2 Please indicate how familiar you are with the following:

How familiar are you with ...	Not at all 1	2	3	Very much 4
... the use of personal computers ?	☐	☐	☐	☐
... the use of PDAs (e.g. Palm) ?	☐	☐	☐	☐
... the use of mobile phones ?	☐	☐	☐	☐
... the use of paper maps ?	☐	☐	☐	☐
... the use of electronic maps (e.g. provided via websites) ?	☐	☐	☐	☐
... the use of search engines (e.g. Google) ?	☐	☐	☐	☐
... the city centre of Aberdeen?	☐	☐	☐	☐

1.3 Please provide some information about your general interest in events:

How often do you...	Never 1			Often 4
... search for events in printed press (e.g. newspaper or magazines) ?	☐	☐	☐	☐
... search for events electronically (e.g. on the web) ?	☐	☐	☐	☐
... attend events in your town (e.g. by going to theatre or music events)?	☐	☐	☐	☐

2. Mobile User Tasks

Background Scenario:

There is a **Musical and Dance Festival held in Aberdeen** offering a large amount of musical and dance events. The event venues are located in and around the city centre (see map – dark points indicate event locations). You also have a mobile device with an application that allows you to search for events.

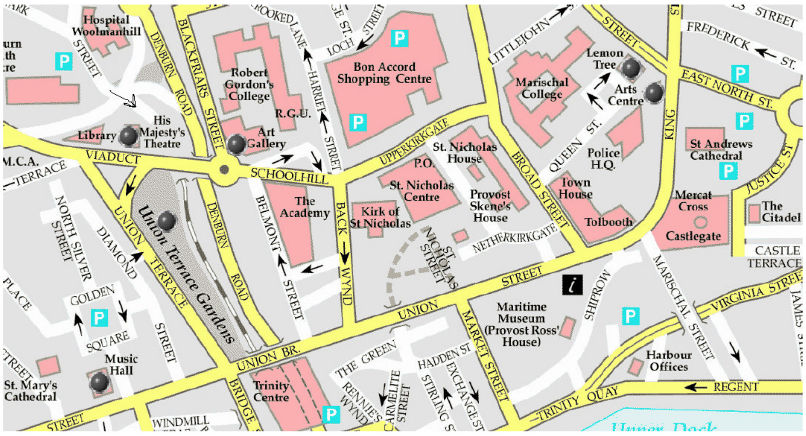


Figure 1: Map for Mobile Study

During the study, you will solve **4 information search tasks** on **2 different mobile software systems**. You will also perform one training task before in order to get used to the software system. Please find the tasks on the following pages.

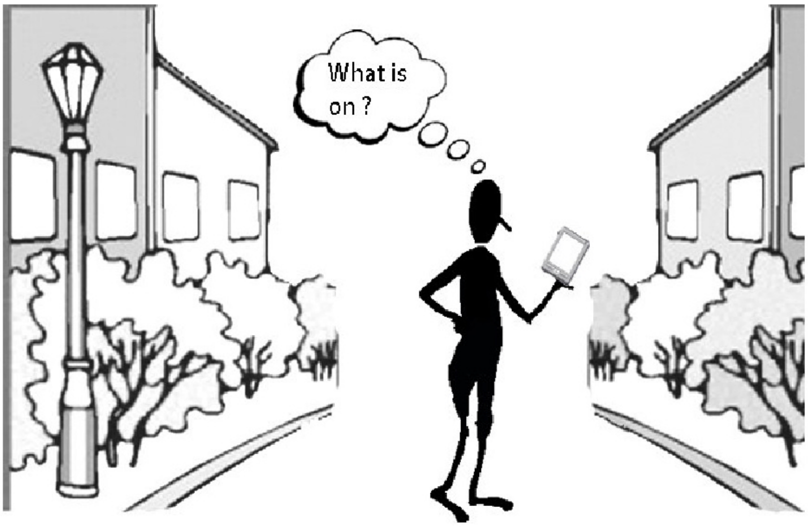
Based on relevance, determined by the system, the event will be plotted in the map with the following colour scheme:

"Red" events are regarded as most relevant by the system; "blue" events are regarded as least relevant. The other colours are in between. Based on this, the system assigns these coloured buttons to the events.

Your task is to rate events based on their usefulness, baring in mind the given task situation.

Training Task Blue

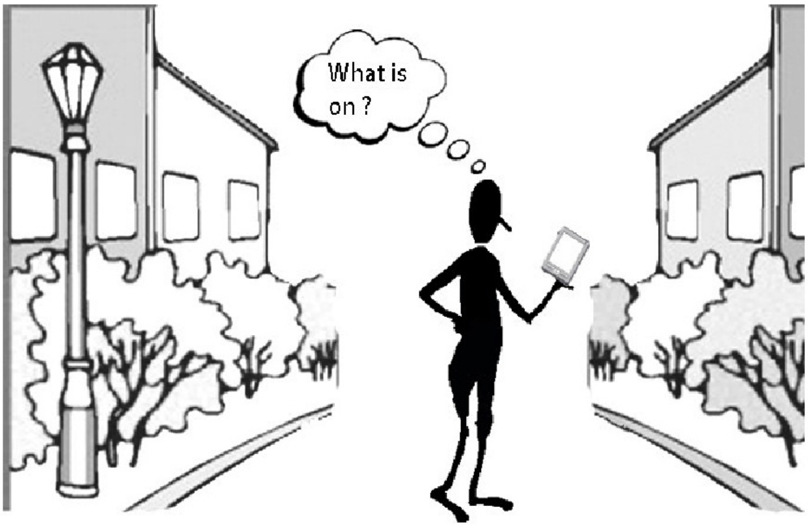
Situation: You are currently at the Art Gallery/ His Majesty's Theatre. You just found out that there is an ongoing Dance and Musical Event Festival here in Aberdeen. Not yet knowing what is offered, you want to make yourself a picture of what events are available for tonight.



Task: Baring in mind the given situation, please find one or more interesting events which you would consider.

Training Task Green

Situation: You are currently at the Art Gallery/ His Majesty's Theatre. You just found out that there is an ongoing Dance and Musical Event Festival here in Aberdeen. Not yet knowing what is offered, you want to make yourself a picture of what events are available for tonight.



Task: Baring in mind the given <u>situation</u>, please find one or more interesting events which you would consider.

Task 1 Blue

Situation: This morning, you and your friend arrived in Aberdeen. After you found your hotel, you and your friend are roaming around in the city centre – now you are at the Art Gallery. It has come to your attention that there is currently a musical and dance festival in town offering a large amount of events in and around the city centre. That is a fortunate incident as you both share a common interest for good musicals.

Task: Baring in mind the given <u>situation</u>, please find one or more suitable musical events which you would consider.

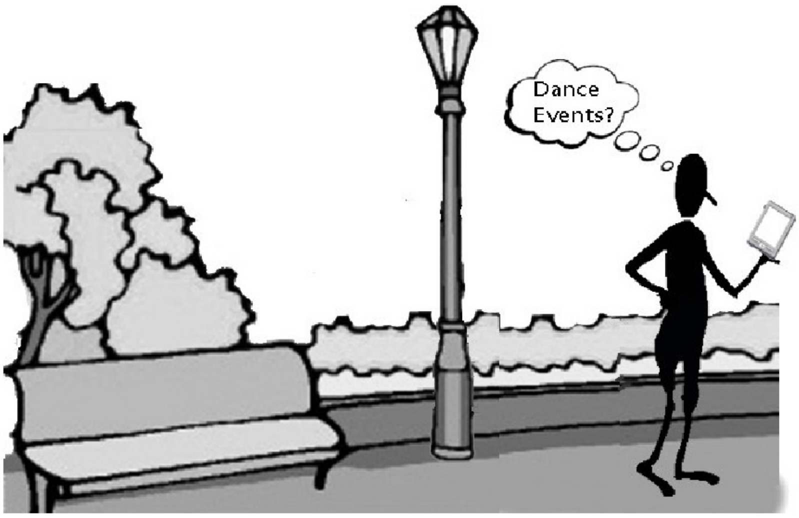
Task 1 Green

Situation: This morning, you and your friend arrived in Aberdeen. After you found your hotel, you and your friend are roaming around in the city centre – now you are at the Art Gallery. It has come to your attention that there is currently a musical and dance festival in town offering a large amount of events in and around the city centre. That is a fortunate incident as you both share a common interest for good musicals.

Task: Baring in mind the given situation, please find one or more suitable musical events which you would consider.

Task 2 Blue

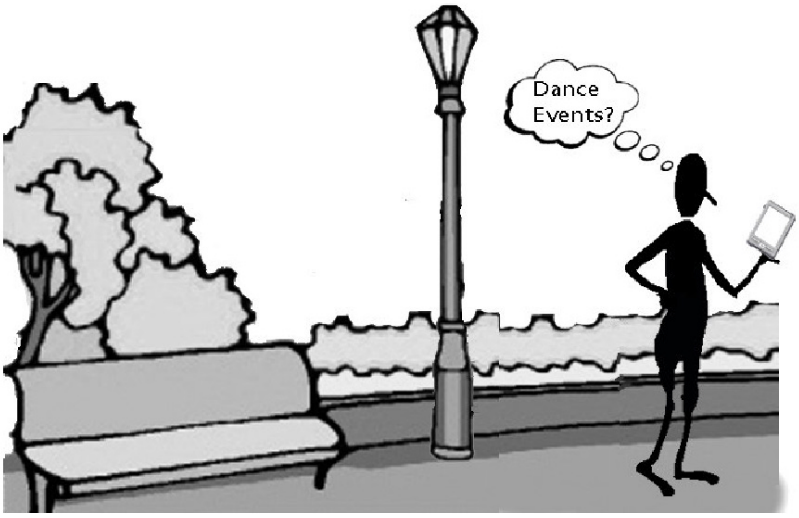
Situation: A friend has phoned you today telling you about the currently ongoing Aberdeen Musical and Dance Festival which offers a large amount of events in and around the city centre. Your friend asked you to find information about dance event performances in order to select something. Your busy working day did not allow you do that so far; now you are at His Majesty's Theatre where you suppose to meet your friend for a drink. You want to use the waiting time to search up some information before you meet.



Task: Baring in mind the given situation, please find one or more suitable dance events which you would consider.

Task 2 Green

Situation: A friend has phoned you today telling you about the currently ongoing Aberdeen Musical and Dance Festival which offers a large amount of events in and around the city centre. Your friend asked you to find information about dance event performances in order to select something. Your busy working day did not allow you do that so far; now you are at His Majesty's Theatre where you suppose to meet your friend for a drink. You want to use the waiting time to search up some information before you meet.



Task: Baring in mind the given situation, please find one or more suitable dance events which you would consider.

3. Post Questionnaire

This questionnaire contains a few questions that collect feedback about the overall experience. All information collected in this study remains anonymous.

3.1 Please tick the appropriate box based on how strong you agree with the following statements:

	Don't agree 1	2	3	Strongly agree 4
The software was easy to use.	☐	☐	☐	☐
The tasks where interesting.	☐	☐	☐	☐
I found it easy to imagine myself in the provided situations.	☐	☐	☐	☐
The tasks where easy.	☐	☐	☐	☐
The event content was interesting.	☐	☐	☐	☐
The amount of tasks was sufficient	☐	☐	☐	☐
The Blue System performed better than the Green System.	☐	☐	☐	☐
The Green System performed better than the Blue System.	☐	☐	☐	☐
Both systems (Blue and Green) performed equally.	☐	☐	☐	☐
The locations helped to imagine myself in the provided situations.	☐	☐	☐	☐

3.2 Do you have any other comments or suggestions?

H

The Mobile Information System

H.1 Design Overview

Users spent most of their time with the mobile application that was provided to them to solve the search tasks as part of the experiment procedure of the mobile experiment described in chapter 6. Figure H.1 presents a conceptual diagram of all main application components. The arrows in the diagram show the most important connections between components, and between components and content. The mobile application consists of a

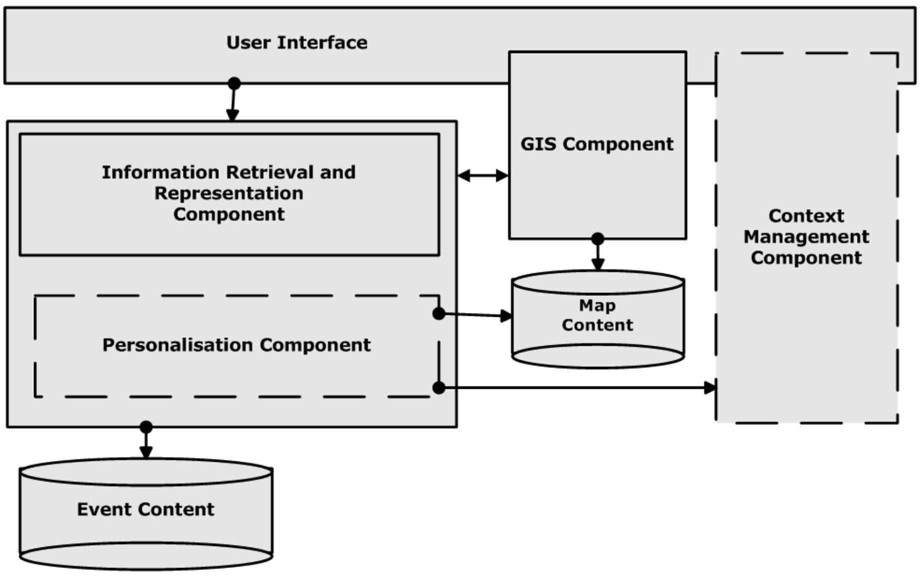


Figure H.1: *Main conceptual components of the mobile application and their connections*

number of components that logically separate its distinct system functionalities. The personalisation model as the key element of this study is part of the Personalisation

Component. In close collaboration, the Context Management Component stores and organises the context model that is only accessed by the Personalisation Component. An Information Retrieval and Representation Component provides the necessary search functionality for users and accesses the event content together with the Personalisation Component. The event content is represented as a Lucene inverted index that stores the event content and prepares it for retrieval. A minimalistic Geographic Information System (GIS) enables users to browse event search results using a small geographic map (see figure H.4 on page 277). Relevant parts of the system are accessed through the User Interface. This is represented in figure H.1 through intersections between User Interface and GIS as well as User Interface and Context Management. The mobile application operates in two different system modes based on the experiment design; a personalised system mode that uses the Personalisation Component and the Context Management Component and a non-personalised system mode that omits these two components[1]. Note that context is only accessed by personalisation as contextual information is directly set by the user interface. In a real application under operational conditions, context would be provided by separate components. These components would either sense contextual information from the environment (e.g. detecting location by GPS signal), explicitly obtain context from the user (e.g. setting status information (e.g. "in office") before starting the application) or implicitly reason context from user behaviour (e.g. detecting interest based on content browsing/viewing history). For the mobile experiment, context was automatically ingested by the user interface when the experimenter selected the task before returning the device back to the participant. The following subsections describe individual components in more detail.

H.2 Information Retrieval and Representation Component

The Information Retrieval and Representation Component uses the Lucene information retrieval library [Hatcher and Gospodnetic', 2004][2] that implements a variation of the vector space information retrieval model [Salton, 1971]. Besides the vector space model, probabilistic models are another very common type of information retrieval model.

[1]This is emphasised in the diagram by the use of dashed lines for these two components.

[2]The search library is generally highly recognised, supported and applied in over 140 applications and websites to date [3]

The question which model performs better is still debated and largely unconcluded [Grossman and Frieder, 2004]; one argument being that comparisons are generally performed on system level (including many adaptations and tools) but cannot easily be focused to the level of the IR model. The Lucene IR library is minimalistic enough to run on the device that was used for the mobile experiment[4]. Nevertheless, the library provides a range of useful optimisations for the standard vector space model as described in section 2.4.

In the vector space model, scores of relevance are determined by projecting content and the queries in a multidimensional vector space and ranking the content (in our case, information about entertainment events) based on its distance to the query. Lucene also uses coordination level matching that additionally boosts terms based on their co-occurrence in query and content to provide an extra level of content distinction.

For the user study, a small event content collection was used consisting of a focused set of 187 events extracted from the much larger Reuters Kalends event collection. Each event consisted of both real and simulated pieces of information. The *title*, the *description* and the *interest* category of the event were taken from the Reuters collection and therefore real with respect to the original content. On the other hand, the performance *time* and the venue *location* were simulated pieces of information based on the experiment design. The event performance time was provided in relation to the current time (e.g. "Today 20:00"). The original event locations from the Reuters collection were changed to the six possible event venues as defined by the experiment design and highlighted in figure 6.2 on page 136. This means that for every search task one out of six events from the collection was co-located with the current location of the participant. Appendix E provides a more detailed account of the Reuters Kalends event collection and some examples of typical events with original content and content structure. Figure H.4 on page 277 shows an example of an adapted event with real and simulated elements on screen (5).

[4]At the time when the mobile application was implemented, no other IR library was available that provided similar functionality

H.3 Geographic Information System

Geographic Information Systems (GIS) are systems for the management, analysis and visualisation of geographic information. This includes not only geographic maps but also all information that can be combined (e.g. visualised) with a map; so-called georeferenced information (e.g. points of interest, the user's current position and environment data for particular locations). GIS appeared half a century ago and is the digital counterpart of the long history of paper maps. Section 2.5.3 from the related work chapter provides more details about GIS and map personalisation in general.

The mobile application offers a minimalistic GIS to support geographic navigation for search results of event content. The studies in chapter 3, undertaken in AmbieSense, highlighted the importance of maps and geographic information for mobile users. The application provides, therefore, a user interface that allows the user to spatially navigate the city centre of Aberdeen with the visualised results within; an alternative to traditional ranked result lists. Search results are plotted within the map based on the last search that was performed. This allows participants to explore the city centre spatially and find relevant events more intuitively 'by sight' rather than by browsing a catalogue. As also shown in chapter 4, event content has a very strong connection with location and time - attributes that are particularly well presented by geographic maps. The GIS provides a small number of basic, internal data structures and features common to GIS applications:

- *Layers:* The use of several layers allows Aberdeen's city centre map to be assembled with different types of data - an image layer providing a high quality visualisation of streets and buildings as well as a second layer with event search results (see figure H.2). Events on the search result layer are colour-coded based on their score using the colour schema shown in figure 6.3 on page 136. Scoring information for each event is obtained from the Information Retrieval and Representation Component by accessing the user's last search results. This scoring information is then used to assign one colour for each event result before being added to the layer. Events are ranked in descending order by score for each event location if more than one event is retrieved (e.g. five events for the His Majesty's Theatre and three events for the Art Gallery as shown in figure H.2). In this case, only the most relevant event is displayed in the map but when clicked, a ranked list of all events is presented in another view. More details on the exact process is shown in section H.5 that describes the user interface. The colour schema allowed users to identify relevant

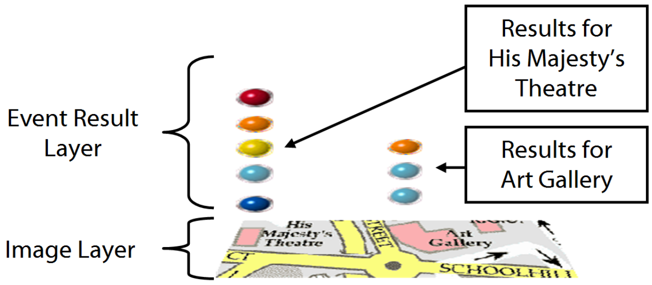


Figure H.2: *Example of a map composition with one image lager and one event result layer. Five results for His Majesty's Theatre and three for the Art Gallery with scores based on the colour schema.*

events easily on the map (see figure H.3) for closer inspection and selection.

- *Navigation and Extended Viewing:* A simplified way of spatial navigation via drag-and-drop for more intuitive map exploration was provided. A viewing aid additionally enabled users to see nearby events that were actually outside the viewable area of the map (see figure H.3). This feature helps overcoming the limited screen size of the PDA. This viewing aid is provided in replacement of a zoom feature that turned out to be too resource intense in terms of performance and memory consumption. The viewing aid is a small dedicated area surrounding the map display (a) that plots all events outside the viewable area of the map to its nearest border (b). These event locations are in reality much further away but can easily be spotted on the border and users can use this information to navigate to these locations.

H.4 Personalisation and Context Management

The personalisation component encapsulates the personalisation model that was described comprehensively in chapter 5. The component also provides means to manage different personalisation models, to configure them and to make personalisation accessible from other system components in an abstract form. For the purpose of this mobile experiment,

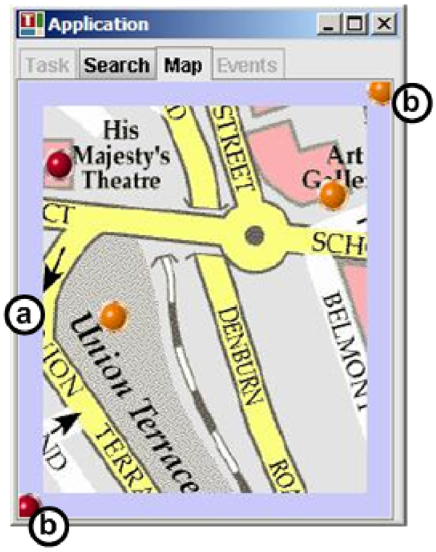


Figure H.3: *Interactive layered map with the extended viewing aid that plots nearby event locations on the closest border.*

the personalisation component was configured with the personalisation model as described in the previous chapter; a model that determines scores of relevance equally based on information retrieval and context. Recalling from the previous chapter, context is represented by a regression model based on the more general theory of causal attribution; a comprehensive framework of how people derive contextual explanations for effects in their environment.

The personalisation model defines and implements the *personalisation input*, the *personalisation method* and the *personalisation output*:

H.4.1 Personalisation Input

The model defines the personalisation input. Firstly, the type and format of the content that the model is able to process. Secondly, the type and format of the information that the model uses in order to adapt the content.

The personalisation model used for the mobile experiment requires event content with an identifier each for the interest category, the time and the location. The interest category has either one of two possible types - either being a musical event or a dance event. The time information is represented as one evening of five possible days. The location is one of six possible event locations as described in section 6.3 on page 133.

The model also requires contextual information to facilitate the personalisation process. Context is represented by one attribute each for the user's current location, the current time and the user's current interest category. This information is administrated by the Context Management Component that provides controlled access to contextual information. For the mobile experiment, current contextual information was automatically set when the experimenter selected the task. Note that each task defined a contextual situation based on the three attributes. In an operational application environment, contextual information would be acquired as described in section 2.3.2 possibly with the help of other application components. These components would gather contextual information with different methods. Context could be sensed from the environment (e.g. detecting location by GPS signal) or the device (e.g. reading time information from the internal clock). Context could also be explicitly obtained from the user (e.g. by application status information). Components could also reason context based on the user's behaviour, such as detecting interest based on the content browsing/viewing history.

H.4.2 Personalisation Method

The personalisation model implements the personalisation method that produces the output of the model. In this thesis, the personalisation method assigned a personalised score to each event based on an information retrieval score and a context score. The information retrieval score is based on a variation of the vector space model. The context score is determined with the following regression function as described previously in chapter 5:

$$Score_{ContextModel} = \begin{cases} e^{1.564} e^{0.217T - 0.106L - 0.885I - 0.147TI} & -2 <= T <= 0 \\ e^{1.460} e^{-0.628T - 0.114L - 0.807I + 0.362TI + 0.088TLI} & 0 < T <= 2 \end{cases} \quad \text{(H.1)}$$

The regression function represents a causal schema of how people explain usefulness (i.e. the amount of situational relevance) of event content in relation to time, location and interest. To enable personalised information retrieval, information retrieval score and context score are combined into a single score. In other words, content-based relevance and contextual (situational) relevance are merged using a single personalisation model:

$$Score_{PersonalisationModel} = 1.0Score_{ContextModel} + 1.0Score_{IRModel} \qquad \text{(H.2)}$$

The two scores are combined by addition, a common way of combining elements that are regarded as independent. The two types of relevance are combined with equal weights. As shown in section 5.4, a change in balance between information retrieval and context considerably affects ranking behaviour. This demonstrates the importance of careful adjustment. The equal weights used in this experiment are a first step of investigating the personalisation model.

H.4.3 Personalisation Output

The personalisation output is the type of personalised result the model produces during the personalisation process. The model that was used for the mobile experiment generated a selection of event content (representing search results from the Information Retrieval and Representation Component) with adjusted scores based on the personalisation method. The modified scores of the personalised results re-ranked the event search results and also modified the visualisation of these events in the map. The next section describes the user interface that connects the application components for the experiment procedure.

H.5 User Interface

The user interface largely reflects the experiment procedure from the application side. Figure H.4 shows an example of one search task in the personalised system mode[5] reproduced from one of the automatic user logs that were recorded during the mobile experiment. The user interface consisted of four different views - "Task"(1), "Search"(2), "Map"(3) and "Events"(4)(5)[6]. These views provided a wide range of personal freedom by offering (relatively) unrestricted navigation between the functionalities of searching, browsing (both geographic and by using ranked lists) and viewing of events. The process

[5]The two system modes of the mobile application were opaqued by naming. The "Blue" system presented the non-personalised system mode and the "Green" system the personalised mode.

[6]For the chosen example, screen (5) expands over two display lengths. This is highlighted with the two border lines in figure H.4.

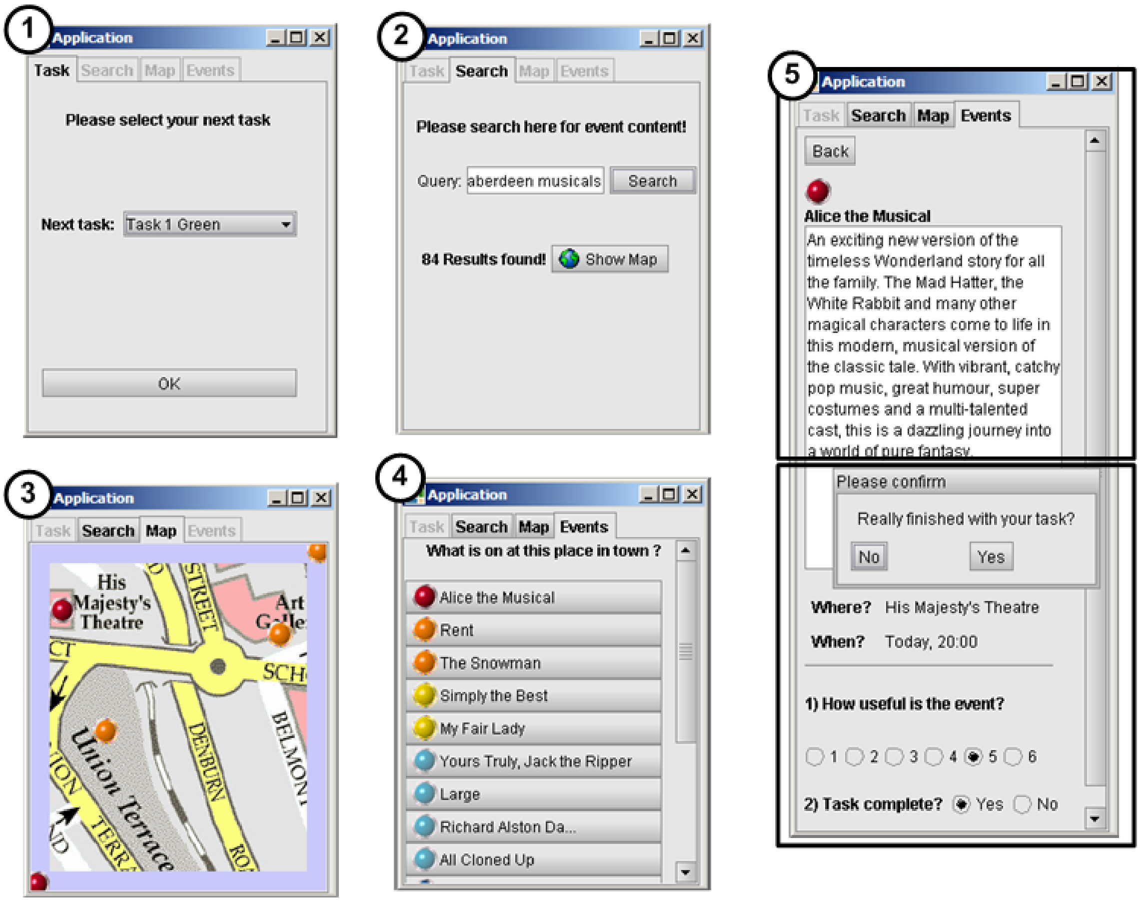


Figure H.4: *User interface views of the mobile application based on an example of search task 1 (musical events); represented with different views for task selection (1), search (2), map browsing (3) and event viewing (4)(5).*

started when the experimenter selected a new task together with either of the system modes (1) (see figure H.4). After this selection, the application switched into the search view (2). At this point, the device was handed over to the participant. Here, the participant submitted one or more search queries (e.g. "aberdeen musicals") to retrieve events based on the provided stimulus (background scenario, situation description and task statement). After every query submission, the number of results was displayed at the bottom of the search view (e.g. 84 in the example) together with a button to switch to the map view (3), a small geographic browser. The map displayed all retrieved events as geographic points based on relevance as determined by the current system mode. In this view, the participant freely navigated the map via drag-and-drop. Event locations with results were represented by the top ranked event at this location (i.e. the event with the highest score) based on the colour schema as shown in figure 6.3 on page 136. Event locations without results did not appear in the map. All those events that were located outside the viewable area of the map were visualised using an extended viewing feature; a small border surrounding the viewable area of the map on which nearby events were displayed. This viewing aid allowed people to identify events from the entire map even if the display was limited to a small part. Section H.3 describes this feature in more detail. The participant eventually selected an event location on the map and the application switched to the 'Events' view (4). This view presented a ranked list of all events available at the selected event location. Events were shown with titles and scoring information based on the colour schema as described in section 6.3. Upon selection of one entry, more detailed information about the event was displayed (5); the event description, the performance time and the venue. At this point, the participant was required to provide a rating of usefulness (understood as situational relevance as described in [Borlund, 2003a]) on a 6-point scale. After that, users were free to continue browsing and searching until task completion. Navigation was supported with tabs that allowed users to freely switch between search (2), map (3) and event views (4)(5) - the only condition being that every newly viewed event had to be rated[7]. The completion of the task was declared by the participant and could be set every time when viewing an event. An additional dialog box ensured that task completion was not selected by mistake. After the task was completed, the view changed back to the task view (1). At that point, the PDA was returned to the experimenter who selected the next task to continue with the experiment procedure. The participant completed one training task in one of the two system modes as well as two experiment tasks for both system modes, meaning that the procedure was repeated five times.

[7]This small restriction was mainly to ensure that users provided enough data.

...what's next?

Fear and Loathing in Las Vegas

Hunter S. Thompson