ESPOO 2005

VTT PUBLICATIONS 566

Eija Kaasinen

User acceptance of mobile services - value, ease of use, trust and ease of adoption

VTT

VTT PUBLICATIONS 566

User acceptance of mobile services – value, ease of use, trust and ease of adoption

Eija Kaasinen

VTT Information Technology

Thesis for the degree of Doctor of Technology
to be presented with due permission for the public examination
and criticism in Tietotalo Building, Auditorium TB104
at Tampere University of Technology,
on the 22nd of June 2005 at 12 o´clock noon

ISBN 951–38–6640–8 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–6641–6 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 4374

VTT, Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel 020 722 111, fax 020 722 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 5, P.O.Box 2000, FI–02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 4374

VTT Tietotekniikka, Sinitaival 6, PL 1206, 33101 TAMPERE
puh. vaihde 020 722 111, faksi 020 722 4102

VTT Informationsteknik, Sinitaival 6, PB 1206, 33101 TAMMERFORS
tel. växel 020 722 111, fax 020 722 4102

VTT Information Technology, Sinitaival 6, P.O.Box 1206, FI–33101 TAMPERE, Finland
phone internat. +358 20 722 111, fax +358 20 722 4102

Cover Juha Kolari

Technical editing Leena Ukskoski

Valopaino Oy, Helsinki 2005

Kaasinen, Eija. User acceptance of mobile services – value, ease of use, trust and ease of adoption
Espoo 2005. VTT Publications 566. 151 p. + app. 64 p.

Keywords mobile services, user acceptance, Technology Acceptance Model, human-centred design, mobile Internet, location-based services, value, ease of use, usability, trust, ease of adoption

Abstract

This dissertation studies user acceptance of mobile services based on a series of case studies of mobile Internet and location-aware information services targeted at consumers and accessed mainly by mobile phones. The studies were carried out in connection with technology research projects, where the objectives were to develop technical infrastructures, to pilot the technology in demonstrator services, and to collect user feedback in field studies to assist in focusing future commercial deployment of the technology. A Technology Acceptance Model for Mobile Services is proposed on the basis of the results of the case studies. According to the model, user acceptance is built on three factors: perceived value of the service, perceived ease of use, and trust. A fourth user acceptance factor: perceived ease of adoption is required to get the users from intention-to-use to actual usage. Based on the Technology Acceptance Model for Mobile Services, design implications for each user acceptance factor are proposed.

Instead of implementing collections of useful features, the design of mobile services should be focused on key values provided to the user. The value of mobile services can be built on utility, communication or fun. Successful service content is comprehensive, topical and familiar, and it includes personal and user-generated content. The users appreciate seamless service entities rather than separate services. Ease of use requires a clear overview of the service entity, fluent navigation on a small display, and smooth user interaction with the service. The users should get personally and situationally relevant services and information without needing to expend effort on personalisation. The services should be designed to be adaptive to the wide variety of devices and networks. As the services increasingly support individual users in their daily tasks and the services are increasingly dealing with personal data, user trust in the services is becoming more and more important. The user should be able to assess whether (s)he can rely on the service in the intended contexts of use. The user needs to feel and really be in control, and the privacy of the user must be protected.

Occasional usage and momentary usage sessions on the move are typical of mobile services. In addition, services are increasingly available only locally or in certain contexts of use. This indicates the need for disposable services: services that are easy to find, take into use, use and get rid of when no longer needed. The user needs realistic information about the actual values of the services, so that (s)he can realise how to utilise the service in his/her everyday life and to innovate new usage possibilities.

Designing mobile services that will be accepted by users calls for changes also to the design methods. Value-centred design aims to identify the key targeted value(s) for end users, service providers and other stakeholders, and then focuses on these values in the design. The proposed Technology Acceptance Model for Mobile Services provides a tool to communicate key user acceptance factors and their implications to the design.

Preface

I started to get acquainted with usability research at VTT about 10 years ago. In many research projects usability studies could identify several usability problems, but those problems often could not be solved in the software development. The most common reason for this situation was that major changes would have been needed and there was no more time in the development project. To solve these problems, we started to develop design methods that would strengthen the co-operation between software designers and usability researchers towards a common goal of more usable software.

We were carrying out technical research projects that aimed at developing technical infrastructures and proving their applicability in demonstrator applications. The draft ISO standard for human-centred design, later published as ISO13407:1999, provided us with a good framework to ensure the usability of those demonstrator applications. That was useful but not adequate for our co-operation companies. Their aim was to exploit the technical infrastructures in commercial applications beyond our demonstrators. To focus the forthcoming application development, they needed information about the most important characteristics of the applications from the users' point of view. As we were already accustomed to carrying out field trials with the demonstrators, we could study not only the usability but also the utility of the demonstrator applications. The Technology Acceptance Model that explained how perceived ease of use and perceived utility affected user acceptance of technology gave us a good framework.

All new technology cannot be studied with a single demonstrator. We soon noticed that it was necessary to enhance the circle: it was beneficial to study a certain technical phenomenon and user acceptance of it in several parallel demonstrators. As we obviously could not implement all those demonstrators ourselves, we set up horizontal usability research projects that were in co-operation with several technology research and development projects. This facilitated the collection, analysis and synthesis of user acceptance research results. We could provide the co-operation projects with usability design guidelines so that they did not have to keep reinventing the wheel.

The research on mobile services described in this dissertation was carried out within these continuously enhancing research circles. And still, the circles are expanding: today we are focusing on the co-operation of different research institutes so that we can all learn from each other's experiences.

Acknowledgements

First and foremost I would like to thank my supervisor, Professor Kaisa Väänänen-Vainio-Mattila from Tampere University of Technology. Without her encouragement, patience, support and strictness when needed, I would never have completed this work.

I was honoured to have Professor Kari Kuutti from the University of Oulu and Professor Martti Mäntylä from Helsinki Institute for Information Technology as the pre-examiners of my dissertation. Their insightful and constructive comments helped in clarifying the focus of my work and in recognising deficiencies in the argumentation.

I also want to thank Professor Ilkka Haikala from Tampere University of Technology, whose SOFKO course tempted me back to the university and encouraged me to start my post-graduate studies after so many years in working life. I want to address sincere thanks to Professor Kari-Jouko Räihä from the University of Tampere, who taught me scientific writing in his TAUCHI courses. Those courses greatly supported my writing of the first three papers of this dissertation.

The research work that this dissertation is built on was carried out in four separate research projects at VTT. Many individuals contributed both in setting up those projects and carrying out the research work. I want to address my sincere thanks to Christer Englund, with whom we set up the first mobile Internet research projects at VTT. His insight into the future possibilities of mobile Internet helped us to get started with this research theme well before others. I want to thank Timo Laakko, who was the project manager in the projects studying mobile Internet. His competence and enthusiasm created the good spirit of the projects. I also want to thank the usability research team, who insightfully carried out all the fieldwork needed: Juha Kolari, Eija-Liisa Kasesniemi, Minna Kulju, Suvi Melakoski-Vistbacka and Raisa Suihkonen. I am also grateful to Matti Aaltonen and Tapio Hiltunen, who together with Timo Laakko took care of the technology development and were so responsive to comments and new ideas.

The studies on location-aware services were carried out in the project Key Usability and Ethical Issues in Personal Navigation (KEN). I am grateful to Antti Rainio from Navinova, who saw that horizontal usability research activity was needed in the Personal Navigation (NAVI) research and development programme and helped in setting it up. I also want to thank the members of the NAVI programme, who supported the work of our project and provided us with interesting case studies. I want to thank all the members of the KEN project team for the good co-operation. I was lucky to have that good a project team! Special thanks to Päivi Topo from Stakes and Teija Vainio from the University of Tampere, who made us see how ethical assessment could be connected to human-centred design. I want to address special thanks to the usability team, who professionally carried out all the fieldwork in sun, rain and frost: Ari Ahonen, Veikko Ikonen, Minna Kulju, Sari Lehtola, Tuula Petäkoski-Hult, Pirkko Rämä, Raisa Suihkonen and Tytti Virtanen from VTT, Annu-Maaria Nivala from Finnish Geodetic Institute, Rolf Södergård from Adage and Teija Vainio from the University of Tampere.

I wish to thank the sponsors of the mobile Internet research projects for their fruitful co-operation: the National Technology Agency of Finland (Tekes), Alma Media, Nokia, Radiolinja and Teamware Interactive. The research work on location-aware services was carried out with the financial support of the NAVI programme, hosted by the Ministry of Transport and Communications of Finland. I wish to thank the members of the NAVI management board for their valued co-operation: Benefon, Digia, Fortum, Genimap, Hewlett-Packard, ICL Invia, Nokia, Novo Group, Radiolinja, Satama Interactive, Sonera, Suunto, TietoEnator, VTT, the Ministry of Transport and Communications, the Ministry of Agriculture and Forestry, the Ministry of Education and Culture, the Ministry of Defence, the Ministry of the Interior, the Ministry of Social Affairs and Health, the Finnish National Fund for Research and Development SITRA, Tekes and the Finnish National Road Administration as well as the SME enterprises of the NAVI network.

I want to express warm thanks to my co-members of the Human Technology Interaction (HTI) Key Technology Action team at VTT: Leena Norros, Johan Plomp and Pirkko Rämä, who have greatly supported my work. The co-authoring of VTT's HTI roadmap and our inspiring discussions have enhanced and clarified my thinking.

I want to thank Research Manager Jukka Perälä and Executive Director Pekka Silvennoinen from VTT for their kind support and the possibility to get a leave of absence from VTT to finalise the dissertation.

Juha Kolari has made the cover picture and he has also taken care of other photographic illustrations in this dissertation. I want to thank him very much for his contribution and patience with my changing opinions. I also want to thank Richard Walker who checked the language of this dissertation, Niina Valtaranta and Maria Lahti who checked the language of the final updates and Leena Ukskoski who took care of the technical editing.

I want to thank all my good friends from Tampereen Tennisseura (Tampere Tennis Club). Tennis has helped to refresh my brain and prevented my shoulders from totally stiffening while working with this dissertation. Special thanks to Anita, Habiba and Pirjo for the weekly doubles and Satu for the weekly singles.

I wish to address warm thanks to all my friends, relatives and colleagues, who have supported me in many ways in writing this dissertation. Special thanks to my mother, Elli Höyssä, who taught me to love books and initially set up a thirst for knowledge in me.

Last but not least I want to thank my daughters Susanna and Anne and my son-in-law Richard Wallenius for their loving support. You have always been so understanding for my frequent physical or mental absences. Susanna and Anne have grown from teenagers to smart, self-reliant and warm-hearted young ladies while I have been working with my dissertation. I am proud of you!

Tampere, 29.4.2005

Eija Kaasinen

Contents

Appendix III of this publication is not included in the PDF version. Please order the printed version to get the complete publication (http://www.vtt.fi/inf/pdf/)

List of original publications

This dissertation includes the following original publications:

I. *Kaasinen, E. 1999. Usability Challenges in Agent-Based Systems. In: Zuidweg, H., Campolargo, M., Delgado, J. and Mullery, A. (eds.). Intelligence in Services and Networks. Lecture Notes in Computer Science 1597. Springer-Verlag, Berlin Heidelberg. Pp. 131–142.*

II. *Kaasinen, E., Aaltonen, M. and Laakko, T. 1999. Defining User Requirements for WAP Services. In: Bullinger, H.-J. and Ziegler, J. (eds.). Human-Computer Interaction. Communication, Co-operation and Applications Design. Volume 2 of the Proceedings of the 8th International Conference on Human-Computer Interaction. Lawrence Erlbaum Associates. Pp. 33–37.*

III. *Kaasinen, E., Aaltonen, M., Kolari, J., Melakoski, S. and Laakko, T. 2000. Two Approaches to Bringing Internet Services to WAP Devices. Journal of Computer Networks 33, pp. 231–246. Elsevier Science B. V.*

IV. *Kaasinen, E., Kasesniemi, E.-L, Kolari J., Suihkonen, R. and Laakko, T. 2001. Mobile-Transparent Access to Web Services – Acceptance of Users and Service Providers. Proceedings of International Symposium on Human Factors in Telecommunication. Bergen, Norway. 5.–7.11.2001.*

V. *Kaasinen, E. 2003. User needs for location-aware mobile services. Journal of Personal and Ubiquitous Computing (6), pp. 70–79. Springer-Verlag, London Limited.*

VI. *Kaasinen, E. 2005. User acceptance of location-aware mobile guides based on seven field studies. Journal of Behaviour & Information Technology, Vol. 24, No. 1, January–February 2005, pp. 37–49. Taylor & Francis.*

The author of this dissertation is the sole author of Papers I, V and VI. In Papers II, III and IV she is the main author. All papers describe the results of research projects where the author has been in charge of the human-centred design activities. In Papers II, III and IV Timo Laakko and Matti Aaltonen contributed the technical descriptions of the prototype systems evaluated. Juha Kolari, Eija-

Liisa Kasesniemi and Suvi Melakoski co-authored the descriptions of the evaluation results in Papers III and IV. They also took care of most of the fieldwork reported in the papers. In Paper IV, Raisa Suihkonen co-authored the descriptions of the evaluation results, and carried out most of the interviews with service providers.

Although the author is the sole author of Papers V and VI, the evaluations, the results of which are analysed and synthesised in these papers, were carried out by a large project group, as well as co-operating organisations of the Personal Navigation (NAVI) research and development programme. The author has been in charge of a horizontal research project within the programme. The project carried out several usability studies and also studied ethical issues. The main role of the author of this dissertation has been to design the evaluation activities as a whole as well as the analysis and synthesis of the overall results. In addition, she took part in some of the fieldwork.

List of abbreviations

2G / 2.5G / 3G	2nd generation / 2.5th generation / 3rd generation (mobile system)
CDMA	Code Division Multiple Access (3G mobile network protocol)
CHI	Computer-Human Interaction
cHTML	Compact HTML
e-mail	Electronic mail
FOMA	Freedom of Mobile Multimedia Access (Japanese 3G system)
GPRS	General Packet Radio System (2.5G mobile network)
GPS	Global Positioning System
GSM	Global System for Mobile Communication (2G mobile network)
HCD	Human-Centred Design
HCI	Human-Computer Interaction
HTML	Hyper Text Markup Language
i-mode	Information mode (Japanese mobile Internet system)
IDT	Innovation Diffusion Theory
IP	Internet Protocol
ISO	International Organization for Standardization
ITU	International Telecommunication Union
J-Phone	A Japanese mobile network operator
kbps	Kilobits per second
KDDI	A Japanese mobile network operator
MMS	Multimedia Messaging Service
NTT DoCoMo	A Japanese mobile network operator
PDA	Personal Digital Assistant
POI	Point of Interest
SMS	Short Message Service
TAM	Technology Acceptance Model
UAProf	User Agent Profile
UMTS	Universal Mobile Telecommunications System (3G mobile network)
UTAUT	Unified Theory of Acceptance and Use of Technology
WAP	Wireless Application Protocol
WCDMA	Wideband Code Division Multiple Access (3G mobile network protocol)
WLAN	Wireless Local Area Network
WML	Wireless Markup Language
XHTML	eXtensible Hyper Text Markup Language
XML	eXtensible Markup Language

1. Introduction

User acceptance is crucial to the success of new technologies but it is difficult to predict. User acceptance of new technologies that are not just incremental improvements on existing ones but cause remarkable changes in peoples lives – so-called disruptive technologies – is especially hard to predict because these technologies may take decades or longer to undergo the transition into everyday objects (Norman, 1998). However, some technologies (e.g. short message services on mobile phones) get adopted very quickly even without significant marketing.

This dissertation studies mobile services – in particular mobile Internet and location-aware information services targeted at consumers – based on a series of field studies of different services and complemented with other usability studies. What kind of general attributes can be recognised in the services that affect user acceptance? What kinds of service entities are acceptable both by the end users and the service providers? What kinds of business models are needed to provide the users with these services? How can we balance user needs, the service provider's interests in providing services, and the possibilities offered by the technology?

A lot of engineering work has been wasted over the years in designing and developing products that were neither wanted nor accepted by the users. How could this waste of work be prevented? Could we assess the acceptance of new solutions beforehand, without actually implementing the products? The human-centred design approach (ISO, 1999) is already a commonly adopted solution to integrate user feedback into the design process. However, human-centred design is focused on the design of individual products. User feedback often arrives too late, when it is no longer possible to change key design decisions. The main design decision, the decision to start the human-centred design process to design a certain kind of product, has already been made and it may not be easy to change that decision even if the user feedback is totally negative.

Business and marketing research already have approaches where new technology is studied on a wider scale. The Technology Acceptance Model by Davis (1989) defines a framework to study user acceptance of new technology based on perceived utility and perceived ease of use. Each user perceives the characteristics of the technology in his/her own way, based for instance on his/her personal characteristics, his/her attitudes, his/her previous experiences and his/her social environment. Innovation diffusion theory (Rogers, 1995) explains the adoption of new practices and technologies in the society in different adopter groups. The Technology Acceptance Model and Innovation Diffusion Theory have been evolved and applied widely, but mainly in the context of introducing ready-made products rather than in designing new technologies.

The case studies of this dissertation were made in connection with technology research projects. The objectives of the research projects were to produce both technology and information on user acceptance to guide future development work on commercial services. The projects developed new technical infrastructures for mobile services and piloted those novel technologies in experimental applications and services. The projects gathered feedback on user acceptance based on field trials where different users used the experimental services independently as a part of their everyday life. In addition to end users we also included different service providers in our studies. Service providers are not necessarily in the information technology business. Their business branch may be for instance tourism, roadside assistance, catering or organising events. They could utilise new mobile technology to better serve their customers but, like the end users, they should be informed of the technical possibilities and listened to in the development process. In the studies we combined end user needs with the ideas of the service providers to identify the basis for successful mobile services.

The case studies took place between 1998–2002. This was an interesting period: mobile Internet research started from scratch, industrial and media interest grew and culminated in a peak of inflated hype in 2000, followed by negative hype and a slow new start from 2001 onwards.

The main research method was horizontal usability studies of two types of mobile services: mobile Internet services and location-aware information services. General user acceptance factors were identified by analysing and synthesising the results of several parallel usability studies. The services were studied from three points of view: what are the possibilities and restrictions of the technology, what do the users need, and what kind of services the service providers see profitable.

Mobile Internet has the capability of providing access to information and services anywhere and anytime. The mobile Internet studies in this dissertation work deal with mobile Internet services working on mobile phones and targeted at consumers. The first case studies were carried out in parallel with the specification of the Wireless Application Protocol (WAP); the first approach in Europe to provide mark-up language based services on the mobile phone. By providing mobile-transparent access to web content via a conversion proxy server, our research team could provide the users with a large selection of services, compared to the first WAP portals that were quite modest pilot services with limited content. Unlike many other WAP studies at that time, we were studying whole collections of services provided to the users via a personal portal, rather than just individual services. The evaluation activities described in the papers range from user requirements definition (Paper II), through laboratory evaluations (Paper III) to field trials (Paper IV). In the WAP studies user acceptance is studied through usability and utility. Papers II–IV describe the key findings and conclusions of our research. The detailed results are presented in Kolari et al. (2002).

The more recent case studies deal with location-aware information services. Varying contexts of use are typical of mobile services. Utilising information about the context of use provides a basis for context-aware mobile services that are not just reduced duplicates of fixed network services but services that are specifically intended for mobile use. These services can give the user quick access to the information or services that (s)he needs in his/her current context of use. Mobile services can adapt or even activate themselves based on the usage context. Location and time are the first elements of the context that can be measured, and these two factors alone make it possible to implement various new kinds of context-aware services.

Our research team evaluated with users scenarios of different location-aware service concepts and some of the first commercially available location-aware information services, ranging from simple short message services (SMS) to map-based guidance services on personal digital assistants (PDA) (Paper V). We also had evaluation activities in connection with different research and development projects, ranging from the definition of user requirements to field trials (Paper VI). In the studies on location-aware services trust and ease of adoption were identified as additional factors that affect user acceptance of the services. Papers V and VI describe the key findings and conclusions of our research. The detailed research results are presented in publications of the NAVI programme (Ikonen et al., 2002b; Kaasinen et al., 2002; Kaasinen, 2003).

Based on the above-mentioned empirical studies this dissertation explores factors that affect the user acceptance of mobile Internet and location-aware information services targeted at consumers. In addition, the dissertation analyses the point of view of the service providers and analyses how the different user acceptance factors should be taken into account in the human-centred design process and in related usability design and evaluation methods.

The main research question is:

1. What are the key attributes that affect user acceptance of mobile services, focusing on mobile Internet services and location-aware information services targeted at consumers?

The supplementary research questions are:

2. What factors affect the acceptance of service providers to utilise mobile Internet and positioning in their services, and what kinds of future service concepts does this indicate?
3. What are the implications of research questions 1 and 2 for the human-centred design process of mobile services and related design and evaluation methods?

This dissertation is structured as follows. Chapters 2 to 5 describe the technical and theoretical background of the research as well as related research. Chapters 6 to 9 describe the case studies that the dissertation consists of (Papers I to VI) and further analyse and synthesise the results. Chapter 2 focuses on mobile technologies such as devices, networks and positioning technologies. Chapter 3 focuses on mobile contexts of use, user interaction with mobile devices and services as well as mobile usability. Chapter 4 gives an overview of the evolution of mobile services and, on that basis, identifies challenges for user acceptance studies of mobile services. Chapter 5 describes different approaches to studying user acceptance of technology and analyses the applicability of these approaches to studying mobile services. Chapter 6 describes the individual case studies and the research framework used in them, and also proposes a Technology Acceptance Model for Mobile Services based on the synthesised results of the case studies. In Chapter 7, the Technology Acceptance Model for Mobile Services is then used as a framework for design implications, based on the analysis and synthesis of the results of the individual case studies. The Technology Acceptance Model for Mobile Services classifies the design implications according to four user acceptance factors: perceived value to the user, perceived ease of use, trust and perceived ease of adoption. Chapter 8 complements the analysis from the service providers' point of view and analyses the challenges for the service providers in providing acceptable mobile services to the end users. Chapter 9 describes how focusing on user acceptance affects the human-centred design process and related usability design and evaluation methods. Conclusions and discussion are presented in Chapter 10.

2. Mobile technologies

This chapter gives an overview of the technical infrastructures on which mobile Internet and location-aware services are used: networks, terminal devices and positioning technologies.

2.1 Mobile networks

Major stages in the development of mobile telecommunications are commonly described as generations. First generation networks were analogue cellular systems. Second generation networks are digital cellular systems such as GSM (Global System for Mobile Communication). Second generation networks are able to serve some data services although the network speed limits the amount of data that can be transferred. First and second generation networks were largely developed under a number of proprietary, regional and national standards (ITU, 2002). The development of 3G systems has been characterised by the aim to develop an interoperable, global standard (ITU, 2002).

As a migration path from 2G to 3G, 2.5G systems such as GPRS (General Packet Radio System) have been introduced in many countries. 2.5G systems introduced the "always-on" notion of communications, and the users are charged per volume instead of time. This was a marked improvement on GSM networks, where the connection had to be set up by a call connection to gain access to mobile data services.

Despite the global standardisation efforts, different approaches to 3G technologies have been taken. Most of these approaches (85% according to Funk, 2004) are based on Wideband CDMA (Code Division Multiple Access): e.g. UMTS in Europe and NTT DoCoMo's FOMA (Freedom of Mobile Multimedia Access) in Japan. However, 3G phones using the CDMA 2000 standard selected by the Japanese operator KDDI have diffused in Japan much faster than NTT DoCoMo’s FOMA phones due to lower cost and better coverage (Funk, 2004). Ignoring global standardisation efforts and taking off with an own solution made possible the success of the first Japanese mobile Internet solution: NTT DoCoMo's i-mode. It seems that KDDI has managed to gain at least short-term success by a similar selection.

GSM networks can provide data rates up to 9.6 kbit/s, GPRS networks up to 171 kbit/s and 3G systems up to 2 Mbit/s locally and 144 kbit/s overall (ITU, 2002). In practice, the first 3G networks could provide data rates up to 384 kbit/s. The growing network speed has mainly been marketed as an enabler for transferring multimedia content. GPRS enabled the fluent transfer of images and short recorded video clips. 3G networks enable real-time video connections.

Local area networks have been developed in parallel with wide area radio networks. WLAN has been a rapid global success. WLAN offers locally high-speed network connections without cables. WLAN networks are available in offices and households, shops, hotels and airports as well as in schools, hospitals and factories. Different WLAN solutions are available, offering speeds up to 54 Mbit/s (ITU, 2002). WLAN connections are mainly used in laptops and PDAs, although WLAN connections are also available for some mobile phones.

Bluetooth is a communication protocol for short-range local communication. Bluetooth was originally conceived as a wireless cable replacement solution with which the user could connect peripherals with a mobile device or exchange information between devices close to each other. The Bluetooth specifications currently enable a multitude of applications including hotspots for local connections (Alahuhta et al., 2005).

Different networks are suitable for different usage purposes and different contexts of use. Current mobile phones have typically one network connection (e.g. GPRS) and Bluetooth for local connections. In the future, the devices will increasingly include several network alternatives, giving more flexible connectivity possibilities for the users, but also raising challenges in selecting which network to use.

2.2 Mobile handheld devices

The development of both digital mobile phones and personal digital assistants (PDAs) originated in the early 1990s. Mobile phones were originally designed for communication but as the technology has proceeded, the phones have increasingly been equipped with different applications and access to services. PDAs were originally designed for small personal applications such as notepads and address books. Soon it became possible to synchronise PDA and desktop PC data, and the increasingly available network connections have further enhanced PDA functionality. As this dissertation focuses mainly on phone-type devices, the emphasis in the following is on mobile phones. PDA devices are described mainly in comparison with mobile phones.

Kiljander (2004) categorises mobile terminals according to the primary input mechanism and usage ergonomics as phones, PDAs, communicators and wearables (Figure 1). Phones are operated one-handed, PDAs are used by holding the device in one hand and operating it with the other hand, communicators are held in both hands and operated mainly with the thumbs, and wearables are attached to body or clothing and operated one-handed. Weiss (2002) divides mobile handheld devices into pagers, PDAs and phones, where communicators are a special case of phones. Pagers are widely used in the North American market, but are quite rare in Europe. They are mainly used for two-way e-mail.

Figure 1. Mobile phone, communicator and PDA.

The term smart phone is used to characterise a mobile phone with special computer-enabled features. These features may include email, Internet and Web browsing and personal information management. Typically the functionality of a smart phone can be further enhanced with add-on applications (Alahuhta et al., 2005). The term media phone is often used to describe phones that include cameras and functionality for image messaging (Multimedia Messaging Service, MMS).

Today, the difference between mobile phones and PDAs is getting more blurred as the screen sizes on phones are getting larger and the phones are equipped with different applications, and PDAs increasingly have network connections as a standard function. However, the main difference is still the numeric keypad and one-hand usage of most phones as opposed to the touch screens and two-hand usage of most PDAs.

The first phones with colour displays were introduced in Japan in 1999. High-quality colour displays were utilised in creating the first camera phones in 2001 and enabled a new service branch of providing picture libraries, and photo-editing functions. In Europe mobile phones equipped with colour screens and cameras were introduced to the market in 2002 as media phones. They provided the users with almost PDA-sized colour-screens, a graphical user interface, fast packet-based GPRS connections and an integrated camera. In 2002 the first videophones came on the market in Japan. Now that the novelty of photos and video has worn off, actually quite few people send photos and videos. Nonetheless the market share of the phones including these features has grown (Funk, 2004). The ability to take photos and video has increased the memory needs. Most phones released in 2003 had more than 1 megabyte memory, and several had more than 5 megabytes. Memory sticks that can be used as external memories for phones are an easy and cheap way of multiplying the amount of available memory (Funk, 2004).

The physical lower limit of phones was achieved already in 1998, as the phones at that time weighted less than 60 grams and the physical size could not be smaller for handheld use. With the larger screens of media phones, the device sizes were again getting larger.

Figure 2. The difference between a mobile phone and a PDA is getting more blurred, as phones are equipped with larger screens. The numeric keypad is still typical of mobile phones.

Current mobile devices exhibit a high heterogeneity of hardware and software capabilities, operating systems and supported network technologies. This heterogeneity does not just cause problems for service providers but it also significantly increases configuration hassles for device users (Bellavista et al., 2002). Device manufacturers and other actors have tried to ease the situation by different standardisation activities, and by providing open platforms for easier service creation and application interoperability.

Java-enabled devices have been available since 2001. With Java-enabled devices the user can download small applications – applets – to the device on the move. The applets fall into two categories: stand-alone applets and agent applets. Agent applets are used for timely information alerts and therefore must be connected to a server to provide up-to-date information. (ITU, 2002)

Mobile Internet was introduced to the market in the late 1990s. The Japanese approaches, NTT DoCoMo's i-mode in the forefront, have been remarkable successes, whereas the European WAP approach has not reached wide markets. Mobile Internet access is implemented by having a browser software application on the phone. The browser is able to access different Internet content, depending on the protocol in use. Mobile Internet is described in detail in Chapter 4.

2.3 Positioning techniques

Location information is becoming an integral part of different mobile devices. Location information is necessary for map and way-finding services. In addition location information can be utilised in location-aware services, providing the users with contextually relevant information and services.

From the point of view of the service, the simplest method of locating the user is to let him/her tell the location. From the point of view of the user, this method requires extra effort because the user needs to define his/her location and input it to the system.

If the user’s device includes a GPS (Global Positioning System) module, the user’s location can be defined very accurately (2 to 20 meters). A GPS cannot be used indoors and it may not work outdoors in all places either. For instance in city centres, on streets edged by high buildings (also known as "urban canyons") there is a limited view to the sky and this may hinder the connection to GPS satellites and thus disable positioning. GPS location is calculated in the user device, from where it has to be sent to service providers in order to get location-aware services.

A mobile phone can be located by the telecom operator within the network. The positioning is based on identifying the mobile network cell in which the phone is located, or on measuring distances to overlapping cells. In urban areas, the accuracy can be as low as 50 meters, whereas in rural areas the accuracy may be several kilometres. The telecom operator generates and thus possesses the user location information. Service providers have to agree with the telecom operator on how to get the location information and the telecom operator must have the user's permission to forward the location information.

The user can also be identified at a service point, utilising e.g. WLAN (Wireless Local Area Network), Bluetooth™, infrared or radio frequency tags. The accuracy of these kinds of proximity positioning systems depends on the density of the network of the access points, and can be down to 2 meters. The user needs special equipment, although WLAN and Bluetooth, for instance, are becoming increasingly common in current mobile devices. Because of the required infrastructure, the coverage of such systems is usually limited to a predefined area, such as a shopping centre, an exhibition area or an office building.

2.4 Summary

The variety of mobile networks is growing, providing the users with both remote and local connectivity. The networks are getting fast enough to provide the users with access to even live video. Mobile phones still have many restrictions due to the continuous trade-off between ease of use and portability. The difference between a mobile phone and a PDA is getting more blurred as mobile phones are increasingly equipped with different applications and PDAs are increasingly equipped with network connections.

The most significant technical changes are expected to occur in equipping phones with many kinds of add-on technology to gather information and to interact with the environment. Local connectivity and positioning are the technical forerunners in this area. The growing variety of hardware and software capabilities gives additional possibilities for new kinds of mobile applications and services. However, the variety of devices and infrastructures also introduces new challenges for service providers in providing their services for several different client environments. The growing variety of options in individual mobile devices also increases configuration challenges for end users.

3. Human interaction with mobile devices

Usability cannot be defined in absolute terms but is always dependent on the user and the usage situation: the context of use. This chapter gives an overview of mobile contexts of use and mobile user interfaces. The main factors that affect the usability of mobile devices and applications are also discussed. Two paradigms are dealt with in detail: personalisation and context-awareness. These themes are repeated in research regarding the usability of mobile applications and services as they are expected to improve remarkably the usability and acceptability of mobile services. However, these paradigms also present new usability challenges that current usability design guidelines and practices may not solve.

3.1 Mobile contexts of use

Identifying the targeted contexts of use is the starting point of a human-centred design process as defined by the ISO 13407:1999 standard (ISO, 1999). The standard defines context as the characteristics of the users, tasks and the environment in which the system is used. As user characteristics the standard lists knowledge, skill, experience, education, training, physical attributes, habits, preferences and capabilities. The technical environment includes the hardware, software and materials used. The environment also includes relevant standards, the wider technical environment, the physical environment, the ambient environment (e.g. temperature and humidity), the legislative environment as well as the social and cultural environment. The ISO 13407:1999 standard points out the need to know not just the users but also the contexts of use, because the context affects the usage situation similarly as the user and the technology in use.

Dey et al. (2001) define context as any information that can be used to characterise the situation of entities that are considered relevant to the interaction between a user and an application, including the user and the application themselves. Dey et al. define relevant entities as places, people and things, and introduce four categories or characteristics of context information: identity, location, status and time. Norros et al. (2003) use the term intentional context to describe what the user has been doing and what (s)he is planning to do next.

Mobile devices are not necessarily used while on the move; they can be used at home or at the office as well. However, the possibility to carry the device with the user introduces great variability to the context of use, which may even change in the middle of a usage session.

Väänänen-Vainio-Mattila and Ruuska (2000) list characteristics of mobile contexts of use that should be taken into account in designing mobile devices and services, classified as the technical context, the physical context and the social context:

> The *technical infrastructure* includes mobile networks that have lower bandwidth than fixed networks and may impose restrictions on usage. The user can be out of network coverage or have bad network coverage. The user needs to get feedback of the progress of data transfer and, moreover, (s) he will need to estimate beforehand how long certain operations may take.
>
> Regarding the *physical context*, user mobility may change the context within a usage session. The user may need to use the device in unstable and varying usage positions, and the user may use the device while moving. The usage may take place in a moving environment such as car or train. Other environmental factors such as variable lighting conditions, noise and varying climate may affect the usage situation.
>
> The *social context* may implicate needs to collaborate and share information with others, and a need to keep interaction paths relatively short and/or quiet. The user still needs privacy and discreteness.

The characteristics of mobile contexts of use defined by Väänänen-Vainio-Mattila and Ruuska (2000) emphasise the restrictions on use caused by the mobile contexts of use. The restrictions may be caused by unstable technology, a less-than-ideal physical environment or a social situation in which the user has to use the mobile device.

Tamminen et al. (2004) point out that mobile contexts cannot be taxonomised in the same way as traditional contexts of use. They claim that mobile contexts of use do not just change in an uncontrolled way, but the users may deliberately set up these contexts by subtle actions and available resources. Based on observational studies of 25 urban people in Helsinki, Tamminen et al. recognised characteristics of mobile everyday contexts as well as the actions and resources with which these contexts were set up. Personal and group spaces were set up by utilising available resources, e.g. looking out of a bus window, reading a newspaper or forming a circle of people around a common point of interest. Unplanned context change, e.g. meeting a friend, may lead to unplanned situational acts. These acts were mainly little side steps to the main task on the user. Also temporal tensions such as acceleration, normal proceeding, slowing down and waiting may change the contexts without the user being able to control them.

Tamminen et al. (2004) point out that multitasking is typical of mobile contexts of use. For instance in way finding, the user still has to position him/herself in the social context of people and physical objects. As the main task of the user requires less attention, the extra attention can be devoted to other tasks. Sometimes the main task is so demanding, e.g. hurrying up to catch a bus, that there is no opportunity for other tasks.

Weiss (2002) compares desktop Web and Wireless Web usage situations and stresses that wireless usage situations are focused on instantaneous search and retrieval actions, whereas desktop usage situations may be more surfing-oriented. For instance, a desktop user may compare prices of flights whereas a wireless user may check the status of a particular flight; a desktop user may research a medical condition, whereas a wireless user may monitor a medical condition; a desktop user may read a movie review, whereas a wireless user may purchase tickets to avoid waiting in line.

When designing services for mobile contexts of use, sometimes the context can be dealt with in a similar way to fixed contexts of use, where the contexts of use and the users set the basis for user requirements. This is especially true with services that are targeted at a specific place or at a specific application field such as support systems for mobile workers, museum guides or shopping guides. With services that are targeted at wider everyday use, the context of use has to be dealt with a bit differently. Mobile contexts set generic limitations on usage, as pointed out by Väänänen-Vainio-Mattila and Ruuska (2000). Another aspect is that since the mobile contexts of use may be very different and the context may even change in the middle of a usage session, the services should adapt to the recognised context of use. The study by Tamminen et al. (2004) sets the basis for the kinds of user studies that are needed to identify how the contexts may differ and change and what kind of adaptation needs these changes may bring about.

3.2 User interaction with mobile devices and applications

Kiljander (2004) has defined mobile phone user interface elements for user input and for output as illustrated in Table 1.

Table 1. Mobile phone user interface elements for user input and for output to the user as described by Kiljander (2004).

User input elements	Output elements
• Numeric keypad for entering digits, letters and special characters • Control keys and devices for controlling the device, such as navigation keys, joysticks, rocker keys, rollers, wheels, softkeys, menu keys and other special purpose keys • Call management keys • Volume keys • Power key • Special purpose keys to access dedicated functionality such as camera, Internet access, voice recorder • Microphone for audio input • Digital camera • Sensors e.g. for light or proximity • Touchpad or touch screen for direct manipulation UI control	• Flat-panel display or displays • LEDs to indicate the status of the device: low battery, incoming call, unread messages etc. • Earpiece and possible hands-free loudspeaker • Buzzer for playing ringing tones and other audio • Vibration monitor for tactile output • Laser pointer, or flashlight

Figure 3 illustrates the main user interface elements related to using mobile services: the display, navigation keys for on-screen navigation, selection key, numeric keys, soft keys to select phone or application-specific functions, and control keys such as Application key and Edit key. The exact implementation varies by phone types. The Application key changes the focus between different active applications and the Edit key gives access to detailed edit functions such as activating the dictionary. Since the number of keys on a phone keypad is limited, most keys include several functions, depending on the context. Numeric keys typically include also three letters. The functions connected to the softkeys vary according to the currently active application. There are typically two softkeys on the phone keypad and the softkeys are located next to the display.

The context-dependent meaning of the softkey is shown on the display next to the key. Directional navigation keys act as a pointing device on the display. They scroll the display or move the highlight in two, four or more directions. The directional keypad usually also includes a selection key to select or accept functions.

Figure 3. Main user interface elements on a mobile phone (Nokia 6600).

Phone keypads support triple tap and most phones also T9® text prediction. Triple tap rotates the selection of the letters associated with the key in question. The last selection remains on screen. T9® text prediction systems continuously predict the forthcoming word as the user presses the keys. It takes some time for the user to learn to use T9 but after the learning it is quite effective and popular, e.g. in writing SMS messages. The efficiency is improved as the user teaches the phone his/her favourite words.

Touch screens on PDAs enable point and click on the screen with finger or a stylus. There is a minimum number of physical buttons, mainly for application selection. Most buttons are available on the screen as touchable objects. Some PDAs also have silk-screened buttons that are permanently printed on the touch screen of the device. Alphanumeric keys are accessible as an on-screen keyboard or the device may include handwriting recognition.

For mobile Internet access, mobile devices have specific micro-browsers. The browsers on different phones vary both by manufacturer and by the model. Basic user interaction when accessing mobile services with the browser is to scroll through pages and select links that further activate other pages. Some browsers allow shortcut keys such as numbers to select numbered menu items. Text input is done by selecting the input field for editing, then editing the text and accepting the result. Forms can be implemented as pages with several input fields. Alternatively, forms can be implemented using a wizard approach – each input field is on a separate page and the user is guided through them one by one.

Mobile Internet services include more simple elements than their desktop Internet counterparts. Typical user interface elements include menus, forms, dialog boxes and articles (readable pages) (Weiss, 2002). As mobile Internet pages are relatively short, these elements are rarely mixed within a single page. Menus include lists of choices and they are often successive to let the user fine-tune their selection, e.g. restaurants-ethnic-Indian. Forms consist of labels and text entry fields. Dialog boxes appear on top of an existing page, requiring immediate attention. Dialog boxes are used for instance to define preferences or choices. Articles are accessible materials on the site such as text, pictures, audio and video. Mobile Internet services are discussed in more detail in Chapter 4.

3.3 Mobile usability

Nielsen (1993) defines usability and utility as subcomponents of usefulness, which itself is a subcomponent of acceptability (Figure 4).

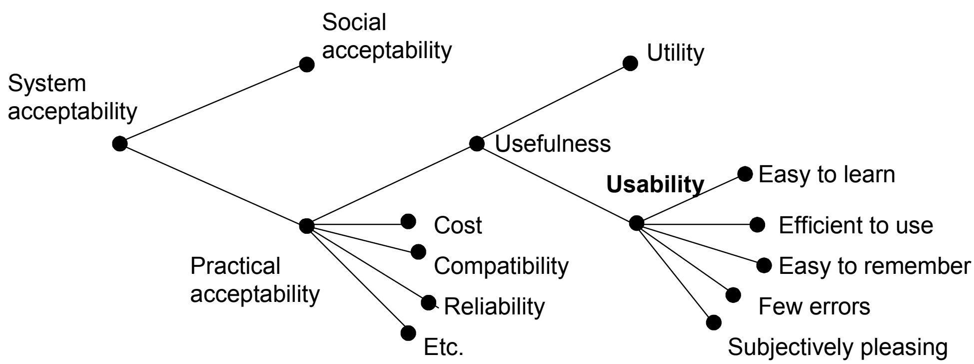


Figure 4. Nielsen's definition of usability as part of acceptability (Nielsen, 1993).

The ISO 9241-11:1998 standard (ISO, 1998) emphasises the context in defining usability: "*Usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*" Usability cannot be measured in absolute terms since it always depends on the context. The ISO definition enhances Nielsen's definition that basically deals with user interfaces rather than application or service entities.

The main usability challenges for mobile devices and services targeted at consumers are related to the wide variation of users, the restrictions of the devices and technical infrastructures, and the demanding contexts of use. Dunlop and Brewster (2002) list five main challenges for human-computer interaction design for mobile devices:

1. **Designing for mobility**: the environment may be a far-from-ideal working environment and the environment may change drastically as the user moves.
2. **Designing for widespread population**: users consider mobile technologies as devices rather than computers.
3. **Designing for limited input/output facilities**: small screens, poor sound output quality, restricted voice recognition, limited keyboards both in size and number of keys, and pointing devices are often hard to use while on the move.
4. **Designing for (incomplete and varying) context information**: mobile devices can be made aware of their context, giving new information to the systems but also bringing problems associated with implying task and user level activities from sensor information and unreliable or patchy sensor coverage.
5. **Designing for user multitasking** at levels unfamiliar to most desktop users: the opportunities for and frequency of interruptions is much higher than in desktop environments.

The guidelines by Weiss (2002) for user interface design for mobile devices emphasise remembering the difficulties that users may have in text input, using consistent terminology and interaction schema throughout the application and also between applications, and feedback about what the application is and how to navigate from each page. Weiss points out that while consistency can increase ease of use, it may also hinder good design when the user interface is migrated from platform to platform.

Kiljander (2004) uses the concept of "*usability knee*" to describe how Nokia has dealt with the problem of functionality versus complexity in their mobile phones (Figure 5). Each device can be described by a curve that describes the relationship between ease-of-use and functionality. As functionality is increased, at some point the ease of use drops dramatically (the knee shape of the curve). The knee describes the limits of each user interface category. To include more features (typically additional applications), a new user interface category needs to be developed. Kiljander (2004) defines Internet browsing as one of the breakpoints where the most simple mobile phone user interface categories cannot provide enough ease of use. According to Kiljander Nokia phones from Series 40 onward provide enough ease of use for Internet browsing. Series 40 phones have a relatively small display (5 to 8 rows for Internet content), a 4-way navigation key and two soft keys. Mobile Internet does not therefore necessarily require very advanced devices, but of course this requires well-designed content. The most demanding functionalities require a lot of user input, such as typing text or drawing.

Ziefle and Bay (2004) have studied mental models of cellular phone menus with users of different ages. Their studies demonstrated that users' mental model of how a mobile phone menu is structured significantly influenced their navigation performance. Their results indicate that older users may need more transparent menu structures with hints that give feedback of the menu structure and where in the structure the user currently is. In their earlier studies, Ziefle and Bay found that children and teenagers might also need this kind of additional navigation support.

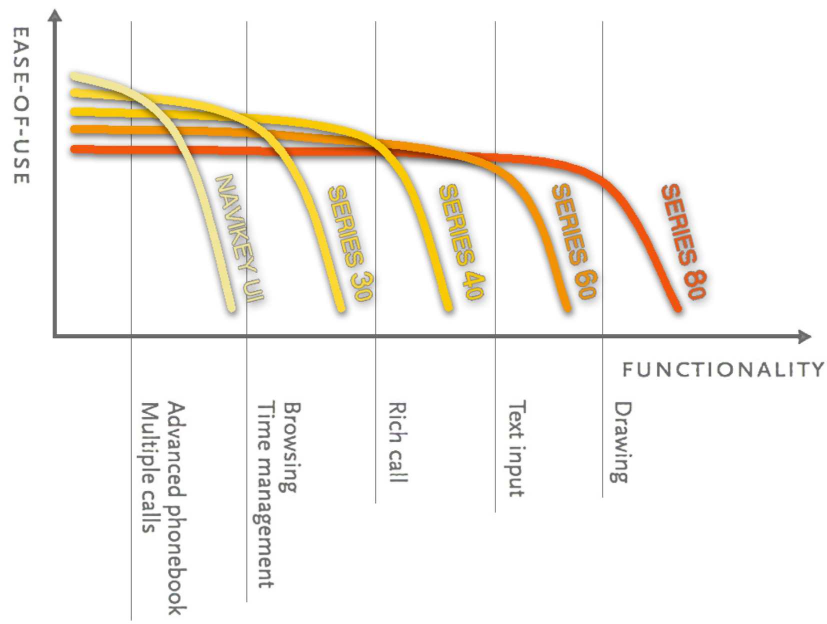


Figure 5. Mobile phone usability knee (Kiljander, 2004).

Buranatrived and Vickers (2004) have studied application and device effects on usability when implementing applications with Java on WAP phone and Palm PDA platforms. Their studies indicate that the complexity and nature of the application had more influence on usability than the platform. By good task and interaction design, applications can be written once and rolled out across multiple platforms just by utilising device profiles for each target platform. However, with some tasks Buranatrived and Vickers identified higher error rates with mobile phones, for instance when scrolling through long lists. They propose that device adaptation should be brought to a higher level to improve usability and to be able to utilise particular characteristics of individual devices. This would require redefining the interface, not just as user interface widget/data pairs but as a description for each task about the input and output requirements. These requirements could then further be rendered to structures of user interface widgets.

3.4 Personalisation

Personalisation is often proposed as a solution to cope with information overflow in mobile services by making personally relevant information or services more easily available to the user. The main challenge is making the personalisation effortless to the user, because even though (s)he would appreciate the outcome there is little or no desire to expend much effort in setting up personalisation. In mobile environments this challenge is especially demanding, as user interaction with mobile devices is tedious.

Personalisation of software means making it more responsive to the unique and individual needs of each user. Personalisation may affect the means of interaction, the presentation or the content. Personalisation can be system or user-initiated, the former often being described as customisation. Both kinds of personalisation can be found in Web services (Perkowitz and Etzioni, 2000; Kohrs and Merialdo, 2001). Personalisation can be done manually, by utilising user or group profiles or the system can learn from the previous behaviour of the user (Norros et al., 2003).

Yahoo! was one of the first sites on the Web to use personalisation on a large scale. Manber, Patel and Robinson (2000) have described their experiences with designing My Yahoo. The majority of active My Yahoo users did not personalise their pages. Manber et al. suggest that a major challenge for large-scale personalisation is to lower the entry bar, making it easier for less-experienced users to customise their pages, and making it clear to novice users that customisation is possible. However, the power users should also be kept in mind, and their needs should not be underestimated. In other Web studies it has also been found that users often express the need for personalisation and enjoy the results, but are not prepared to do much about it (Nielsen, 1998).

Personalisation is an important feature in mobile phones today. Personalisation mainly takes the form of personally selected covers, ring tones and logos. Blom (2000) has studied the motivations and effects of mobile phone personalisation. The main motivations turned out to be ease of use, reflecting one's own or group identity, feeling in control and pure fun.

The early WAP studies presented quite contradictory views on personalisation. Hollensberg and Vind Nielsen (2000) suggested that in WAP, personalisation is a highly advanced feature, which WAP users will explore and use later, if ever. The first-time experience is always the non-personalised version. However, the authors admitted that personalisation offers great advantages and is a strong feature for the expert user. On the other hand, Ramsay and Nielsen (2000) pointed out that the lack of personal and localised services were very significant failings in WAP. They claimed that apart from entertaining someone with time to kill, personal and localised information was where WAP should have been able to make an impact.

3.5 Context-awareness

An efficient way of improving the usability of mobile services and applications is to adapt the content and presentation of the service to each individual user and his/her current context of use. In this way, the amount of user interaction will be minimised: the user has quick access to the information or services that (s)he needs in his/her current context of use. The services can even be invoked and the information provided to the user automatically.

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task (Dey, 2001). The main problem with context adaptation is that the context cannot be easily identified or measured. The location of the user is an element of the context that currently can be measured more or less accurately depending on the positioning system in use. Location-aware services are a concrete step towards context-awareness. Other aspects of context-awareness will follow as soon as the corresponding elements of the context, such as weather or the social situation, can be measured and the adaptivity needs can be identified.

A context-aware user interface can select the appropriate modes for service interaction. A context-aware user interface can be implemented, e.g. as context-aware text prediction (Ancona et al., 2001) or a location-aware remote control for the environment (Fano, 1998). A major challenge for the context-aware user interface is that the context may be continuously changing. This raises the problem of integrating changes into the user interface in such a way that the user remains in control (Cheverst et al., 2000). Moving can also be seen as one mode of interaction with the system. This interaction mode is quite challenging because it is difficult to know the user's intention: is the movement really taking place in order to interact with the system (Svanæs, 2001).

Context-aware content has been studied in different application areas, e.g. tourist guidance (Cheverst et al., 2000), exhibition guidance (Bieber and Giersich, 2001), museum guide (Ciavarella and Paternò, 2004), e-mail (Ueda et al., 2000), shopping (Fano, 1998), mobile network administration, medical care and office visitor information (Chávez et al., 1999). In these studies, the location of the user is the main attribute used in the context-adaptation and the device platform is in most cases a PDA. In well-defined application areas, it is possible to predict the other elements of the context according to the location of the user. Designing for more general user groups and wider contexts of use will be much more challenging.

Sorvari et al. (2004) point out that proximity of people and objects can be measured by identifying nearby Bluetooth devices. This enables the identification of people who are present and in some cases also the environment, e.g. an identified Bluetooth car kit may indicate that the user is in a car. Sorvari et al. (2004) suggest that automatically generated context metadata can be used to manage personal information content such as photos. However, they admit that converting location and proximity data into meaningful context metadata is a major design challenge. Iacucci et al. (2004) have done a related study where they have developed a system that provides mobile support to record and organise multimedia in order to preserve more of the experience of a tourist visit. In their study the context included time and location. Funk (2004) points out that the problem with location-based marketing is that it is often impossible to know whether the user is on his/her way to or already at his/her destination. If the users get all kinds of offerings already before getting to where (s)he is heading, (s)he gets an over-dosage of messages. Commercial messages need to be based on a more exact knowledge of the context of the user, and should preferably be based on agreements related to regular customership. In many countries this is obligatory because pushing messages to the user without his/her permission is illegal.

3.6 Summary

The main usability challenges with mobile devices, applications and services are related to the widespread user population, the restrictions of the devices and technical infrastructures as well as the varying and demanding contexts of use. Short instantaneous usage sessions are typical of mobile usage and the usage often takes place in parallel with some other activity. Especially actions that include excess user input are prone to usability problems due to the limited user interface facilities on mobile devices. Transparent menus and clear feedback are especially important for younger and older users, but they will benefit all users. With mobile consumer services, the widespread population indicates a large variety of client devices, bringing in the need to adapt the services for different devices. However, limitations on device functionality may make some tasks unsuitable for some device types.

Mobile usage situations vary a lot in terms of the technical, physical and social context of use. Thus the services cannot be designed only for specific contexts of use but need to be adapted to different users and usage situations. Together, personalisation and context-awareness can greatly improve the usability of mobile services by providing the users with contextually and personally relevant information and functions. A major design challenge is to predict the personal needs in varying contexts of use without requiring constant attention from the user in the form of personalisation dialogues and refining questions. Personalisation in itself may introduce additional usability problems.

Usability issues related to mobile services are discussed further in the next chapter.

4. Mobile service concepts – history and future visions

Mobile Internet was launched on the market in the late 1990s. This technology provided the infrastructure to enhance the mobile phone from traditional communication tasks to a more versatile tool for accessing information and services. The expectations were high but, apart from Japanese mobile Internet solutions, i-mode by NTT DoCoMo in the forefront, the initial introduction of the services on the market failed, ending in lost investments and even bankruptcies. This chapter describes the history of mobile Internet – focusing on the two mobile Internet solutions, WAP and i-mode, as well as the current situation and future visions of mobile services.

4.1 Mobile Internet

Mobile Internet refers to the convergence of mobile communication technologies with information and data communication services. The WAP (Wireless Application Protocol) approach was to use an XML-based language, WML (Wireless Markup Language), specially designed for small phone-type devices. The first Wireless Application Protocol (WAP) services in 1999 were separate applications, and little attention if any was paid to their co-existence with a corresponding Web service (Hjelm, 2000). In Japan, telecom operators were each implementing their own mobile Internet approach. Most operators were using limited sets of the HTML language, thus providing the users with access to some Web pages as well. This far, WAP has not been a remarkable success, whereas in Japan the various mobile Internet approaches soon became very popular, i-mode by NTT DoCoMo being the best known. In the following I use the term WAP to refer to the European deployment of WAP. WAP was also used by some Japanese operators but their technical infrastructures and business models resembled more the i-mode approach.

Mobile Internet usage is typically characterised by short bursts of activity. The browsing behaviour is object-driven rather than exploratory, and on average 80% of usage sessions are shorter than 10 minutes (ITU, 2002). The basic usability problems with phone-based Internet access are well-known: problems in getting an overview of the content on a small screen, navigation problems due to the small screen and limited input methods, especially slow and cumbersome text input. WAP usability problems also included slow connection set-up and response times that were not such critical problems in the Japanese approaches due to their packet-based networks.

ITU (2002) crystallises the lessons learned during the 6-year period of practising with mobile Internet services by stating that the development always takes longer than expected and that the pioneers tend to get burnt fingers. The media and the market follow different cycles – the hype often falls flat before market development begins to take off.

Mobile phones launched to the market in 2003 began to include XHTML (eXtensible Hyper Text Markup Language) browsers (e.g. Nokia 6600). This made it possible to provide the same content both on the Web and on mobile devices. However, due to the limitations of mobile devices it has turned out to be almost inevitable that alternative implementations of the services for mobile devices are to be provided. In addition, the growing variety of mobile devices requires several alternative implementations.

In the following sections, WAP and i-mode are described in more detail.

4.1.1 Wireless Application Protocol (WAP)

WAP was planned to be an open standard, available to all vendors and targeted at devices ranging from small mobile phones to PDAs. The WAP specification defined the User Agent Profile (UAProf), which included device hardware and software characteristics as well as application and user preferences for the client device. Based on the UAProf attributes, the services were supposed to be able to adapt the content and the presentation for each client. The definition of the UAProf included over 50 parameters and thus constituted a major adaptation challenge for service designers.

Although WAP itself was a standard protocol, the implementation of first WAP browsers on different devices varied as device and browser manufacturers compromised between device look & feel and browser functionality in different ways. A number of browsers supported phone-type use. WAP-specific keys were added to the keypad (e.g. Back, Clear and Options) and short cuts were available for many functions. Other browsers supported Web-type use. No WAP-specific keys were included in the keypad. Most actions were done by selecting a link on the screen. Since no dedicated Back key was available, the

Back function had to be available as a link on each screen page. Figure 6 illustrates some of the first WAP-enabled mobile phones.

Figure 6. Some of the first WAP phones by Ericsson, Nokia and Siemens. The same service looked and felt quite different on different WAP phones.

The first WAP usability design guidelines were device specific, and did not include advice on how to ensure that the service would work on other platforms as well. Later on, when more devices were launched on the market, a concept of designing generic WAP applications was published. The idea was to give instructions on designing minimal services that would work on any device, even on the most modest one. However, it was soon realised that the variety of available devices was so big that these kinds of general guidelines did not work and the number of separate guidelines for different devices and browsers continued growing. All these guidelines emphasised simplicity, navigation support and consistency but the recommendations regarding the way in which these could be achieved in practice varied according to the device in question.

In many cases the browser and the mobile Internet functions were mishmashed with the phone's functions, causing usability problems (Funk, 2004). For instance when selecting "Options" while browsing the mobile Internet the user could get a selection of optional functions, some of which were phone functions, some browser functions and some related to the mobile Internet page that was currently selected. XHTML guidelines propose keeping phone, browser and mobile Internet functionalities separate, thus solving these usability problems.

The first WAP services were provided as operator portals, so-called walled gardens that did not allow access to the services of other portals or third party services. Unlike i-mode (see 4.1.2), the selection of services was modest – typically the first services were simple demonstrator services produced by the operators themselves. The WAP services in Europe were provided on the GSM network, thus requiring the users to first set up a call connection before they could access the WAP content. The price of WAP services was quite high and the price consisted of both telecommunications costs and service costs, making it difficult for the users to estimate real costs. Quite soon, though, WAP portals were opened so that the users could access any WAP services. Installation has been made much easier since the early days, but still the services are not readily installed at the time of purchasing the device. The user has to know how to order the settings with a text message and needs to set the necessary bookmarks.

In the first WAP services, personalisation was usually implemented as a Web service, where the user could select personal links from a selection of alternatives. While being in a WAP service, the users could also make bookmarks in a similar way as when browsing on the Web. As the list of alternatives was not very extensive, the outcome cannot be described as providing truly personalised content for the user.

Version 2 of the Wireless Application Protocol was released in 2001. WAP2 devices could access content written in XHTML Basic in addition to WML, which was the only content format in the original WAP. Services written in XHTML Basic are accessible both on the Web and on WAP.

Figure 7. On media phones colours and images can be utilised in services (Sonera Plaza, wap.sonera.fi).

In 2002 many mobile phone manufacturers introduced their media phones (Figure 7). These phones provided the users with almost PDA-sized colour screens, graphical user interfaces, fast packet-based GPRS connections and integrated cameras. With the bigger screen size and always-on and faster GPRS connections many of the usability problems with mobile Internet encountered with earlier phone models could be overcome. Better devices have encouraged service providers to develop new WAP services and the selection of available WAP services has grown little by little. Still, mobile Internet services on the phone have not yet become very popular in Europe. There are certainly several reasons for that but at least the bad reputation due to the unsuccessful initial introduction of WAP, high pricing and user efforts needed in installations have affected the situation.

4.1.2 i-mode

Launched in 1999 in Japan, i-mode by NTT DoCoMo has been one of the few success stories of mobile Internet. In December 2000 over 50% of NTT DoCoMo's cellular subscribers had subscribed to i-mode. By June 2003, the percentage had arisen to 87%, and i-mode comprised 60% of the Japanese mobile Internet market in terms of the number of users. The most notable competitors of NTT DoCoMo's i-mode are the mobile Internet solutions provided by network operators J-Phone and KDDI. As these and other competitors have released new services that in many cases have been more innovative than i-mode services, i-mode has lost its appeal among new subscribers. In the second quarter of 2003, only 32 % of new cellular subscribers in Japan selected NTT DoCoMo. (Mitsuyama, 2003)

i-mode was initially available on 9.6 kbps packet-based networks. In 2002 the network was upgraded up to 28.8 kbps. Already earlier, in October 2001 NTT DoCoMo launched their 3G service FOMA that offered speeds up to 384 kbps and the subscribers could access enhanced i-mode services known as i-motion, referring to the availability of multimedia. At first, FOMA was available only in very restricted areas. (Mitsuyama, 2003)

The data exchange protocol in i-mode is Compact HTML (cHTML), a subset of HTML, enabling access to existing simple Web pages. Because of the packet-switched network protocol, the Internet access mode has been "always on" from the very beginning and users do not have to set up the connection to access i-mode services. I-mode has been marketed as a consumer service, avoiding technical details in marketing and offering a compact packet of terminal device, network and services on a turnkey basis to the consumers.

Figure 8. Two early i-mode phones.

In i-mode, the personalisation can be done within the mobile service itself. The user can register sites to a personal page, "My Menu", which is accessible from the i-mode front page. The user can also select "Add to My Menu" to add a link while accessing the service. The latter approach does not require the user to actively start personalising but (s)he can perform the personalisation as a sub-task while accessing the services. There are both official i-mode sites with the certification by NTT DoCoMo (3,462 official sites in March 2003) and unofficial sites (64,207 in March 2003) (Mitsuyama, 2003). Since July 2000, sites in the English language have also been available. Charging is based on monthly basic subscription fee (300 yens – about 2 euros – in 2003) and monthly subscription fees to information providers (100–300 yens – 70 cents to 2 euros – per service in 2003) (Mitsuyama, 2003). E-mail has been the most important i-mode service, and probably the key success factor behind i-mode. Before i-mode, mobile messaging in Japan was based on pagers because SMS was not available. As the number of i-mode users has grown, also the amount of usage per user was saturated by early 2003 (Funk, 2004).

NTT DoCoMo has defined their targeted value to the customers as providing at affordable price:

- immediacy – information available as the end user needs it with low latency
- ubiquity – information available anywhere
- mobility – information available on the move
- utility – usefulness and ease of obtaining information.

NTT DoCoMo has paid a lot of attention to ensure the quality of i-mode services. The guidelines for content providers (Mitsuyama, 2003) include e.g. that all content must be updated several times a day so that it is as current as possible and that the content must be as comprehensive as possible, providing access to other levels of related information. In addition, content design must be attractive and both the information provided and the services must be highly readable and understandable. The i-mode service guideline (NTT DoCoMo, 2002) identifies the minimum requirements for i-mode services. The service providers need to follow the guidelines by NTT DoCoMo in order to provide services for the i-mode portal. However, NTT DoCoMo has the right to finally accept or reject candidates for i-mode services. NTT DoCoMo has also created guidelines for i-mode handset manufactures to ensure that devices include the necessary features required to access i-mode services. For service providers, today it is no longer as important as before to be one of the official i-mode sites, because they have adopted other ways to distribute information about the availability of their service. Also consumers have become more aware of alternative sites.

The new FOMA services for 3G networks include distribution of videos, music and electronic books, videoconferences as well as one-to-many video streaming. The introduction of the 3G FOMA services was not as big a success for NTT DoCoMo as the original i-mode. The main reasons for the slow start include:

- The service area for WCDMA (wideband CDMA) that FOMA was using was very limited in the beginning, and there were dead zones even within the reported service areas.
- The first FOMA phones were bigger and thus clumsier than the subscribers were used to with i-mode.
- Video calls and other 3G features consumed a lot of energy, and the users had problems with charging their devices often enough.
- New services and features were introduced gradually, so the introduction of FOMA did not bring in a revolutionary change for the users.
- FOMA handsets were considerably more expensive than the earlier i-mode devices.

Mainly because of the better coverage, NTT DoCoMo's competitors had a better start for their 3G services. For instance, KDDI was using a different network technology that could utilise existing network spectra (CDMA2000) and could thus provide better coverage from the beginning.

4.1.3 Mobile Internet usability

The basic usability problems with mobile Internet are related to the general mobile usability issues described in Chapter 3.3. In a mobile phone, mobile Internet is an add-on to a ready-made communications device. This causes usability challenges in integrating the Internet functionalities into the device menu structures so that they do not disturb the device's main functions of calling and messaging. As described in Chapter 4.1.1, some early WAP phones left the phone keypad intact, and others added WAP-specific keys such as Start WAP, Back and Browser Options. The former approach kept the phone’s basic functions easy to use but made using the new WAP functions more complex. The latter approach made the phone more complex in all tasks, but made WAP usage easier.

One of the early WAP usability papers by Buchanan et al. (2001) sums up the basic design guidelines:

- Provide direct, simple access to focused, valuable content.
- Trim the page for minimum navigation needs; use simple hierarchies familiar from phone menus.
- Reduce the amount of vertical scrolling by simplifying the text to be displayed.
- Reduce the number of keystrokes the user needs to make, e.g. replace text input with selection lists.

Weiss (2002) gives more detailed guidelines to ease navigation by duplicating the most popular links, putting links consistently on the bottom of the pages, supporting number short cuts, keeping menus flat and providing a way out on each menu. He also advocates avoiding text entry when possible, providing content on each page to avoid endless navigation, and making the case in the first 12 letters.

These guidelines have a lot in common with the many early WAP usability guidelines provided by device manufacturers and other organisations in the field. There have been both device-specific guidelines and efforts to provide more general guidelines. The problem with general guidelines is that their observance tends to generate services that work in practice on any device but do not enable utilisation of the advanced features of more versatile devices.

The guidelines by Sony Ericsson (2002) extend the view from managing the technical limitations to making use of mobility. Sony Ericsson emphasises presenting content relevant in mobile contexts – "*here, now, fun and cool*" – and targeting the content at a wide audience. The content needs to be fresh, the minimum content update rate according to Sony Ericsson (2002) is weekly, and an ideal update rate every 5 minutes.

In addition to separately designed mobile Internet services, many Web pages provide small screen / slow connection versions of their pages. On these pages, the layout, display and sometimes also the content has been changed or reduced to fit the restrictions of small devices and slow networks. Web pages can also be transformed on line at the site or at the browser by linear transformation or overview transformation (MacKay et al., 2004). Linear transformation converts the site structure to a long linear list that fits with the width constraints of the client device. User interaction on the transformed page basically consists of scrolling the list and selecting links. This transformation is suitable for devices with very small screens such as mobile phones. Overview transformation is based on shrinking the original page. The user interaction consists of magnifying parts of the screen and then interacting on the sub-screen. Overview transformation is more suitable for devices with larger screens and more processing power.

One of the most recent mobile Internet usability guidelines by Nokia lists five key considerations when creating usable XHTML content for Series 60 platforms (Nokia, 2004):

> **Display size.** Fit all components, such as images and tables onto the screen. Avoid the need for horizontal scrolling. Use a brief and descriptive page title that fits on one line. If a page is long, use anchors to enable jumping inside the page.
>
> **Images.** Minimise the number and size of images. Make sure your service is understandable also without the images.
>
> **Page layout.** Do not use frames or pop-up menus, as they are difficult to handle on a mobile browser display. Insert a Back link to each page.
>
> **Forms.** Avoid requiring letters and numbers in the same input field. Prefer selection lists over text fields. Avoid long text entries and provide good default values when possible. Avoid errors by accepting several value formats and make error correction simple.

Style. Check that the text has good contrast with the background. Limit the use of different colours and use the same colour for the same XHTML elements. Keep the content of the site consistent and apply the same style on all pages.

Mobile Internet usability research is mainly focused on coping with user interface restrictions such as on-screen navigation (Ramsay and Nielsen, 2000; Kaikkonen and Roto, 2003; Chitarro and Dal Chin, 2002; Chae and Kim, 2004). The guidelines by Nokia (2004) well represent the current comprehension of key issues related to mobile Internet usability, which is mainly related to the restrictions of the devices and networks. That is why most guidelines focus on avoiding or at least minimising certain types of elements and content. As user input is usually quite tedious, the guidelines tend to recommend minimising input efforts.

4.2 Location-based services

Positioning is becoming an integral part of new mobile devices and networks. Location-awareness in services can be used to provide the user with topical, location-based information, targeted at mobile users, not just a fallback to Internet access. Different technology components for location-based computing are currently available but the lack of integration architecture and the slow development of business models has hindered the spread of location-based services (Rainio, 2003). The large number and variety of location technologies is likely to inhibit the growth of location-based services, as it will be difficult to ensure a truly global service when technologies and standards differ from region to region (ITU, 2002).

Market research in 1999 forecasted fast development and big business potential for location-based services. In connection with the collapse of mobile Internet hype, the expectations for markets of location-based services have also settled down. The business is not expected to be significant until after 2005 when location-based services are estimated to represent around 7-10% of the mobile service market (Rainio, 2003). Currently, the consumer service market in Europe is quite modest although services such as friend finders, maps and location-aware directory services are available both as mobile Internet services and as downloadable applications. The market introduction of location-based services has certainly suffered from the bad reputation of WAP both in the media and in the minds of potential customers. Also the marketing of these services has not been very visible. For professional use there are already quite a lot of successful services, e.g. for fleet management (Rainio, 2003).

As in the mobile Internet generally, Japan has also been the forerunner in location-based services. All Japanese operators have some location-based services, and the services were accessed half a million times a day in 2003. Location-based services include searches for local services such as shops, restaurants and hotels, coupons with special offers, local news, maps, traffic information and weather forecasts (ITU, 2002; Rainio, 2003). DoCoMo i-area service as well as J-Phone's J-Navi provide these local services simply based on the handset's area code (ITU, 2002). One cultural specificity is that in Japan employees get transferred frequently, and they use mobile location-based services to familiarise themselves with their new surroundings (Rao and Minakakis, 2003).

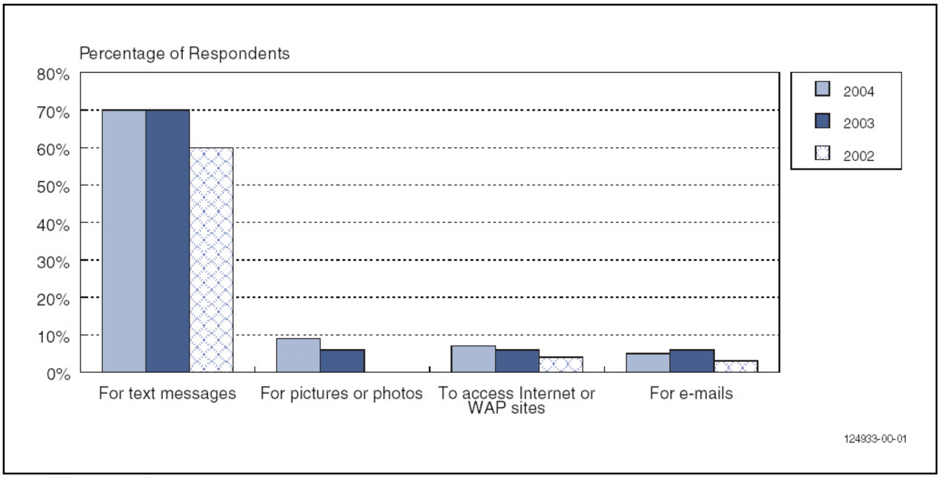
Rao and Minakakis (2003) point out that the main obstacles to the customer value proposition for location-based services are the hype surrounding them, the current limitations of technical solutions, potential intrusions into user privacy, and the lack of sustainable business models that sufficiently address and enhance the customer experience. Rao and Minakakis (2003) expect that the major drivers for the demand of location-based services will be, in addition to the consumer and business users, the large variety of objects and inventories that can be identified and tracked in physical space. Business opportunities for consumer applications include route guidance, personal information delivery at points where needed, and niche applications with rich but narrow content for groups such as golfers, outdoor enthusiasts and families.

4.3 Mobile service business

4.3.1 Mobile service markets

Figure 9 illustrates the results of a user survey by Gartner (Pittet, 2004) of mobile service subscribers in six European countries (France, Germany, Italy, Spain, Sweden and the United Kingdom). The respondents were asked for which purposes they were using their mobile phones. Compared with text messages that were used by 70% of the respondents, access to Internet and WAP sites was quite rare, less than 10% of the respondents had been using their phone for that purpose. The usage of mobile Internet also appears to have been growing very slowly since 2002.

A survey by Gartner in Japan (Liew et al., 2004) shows remarkably higher usage percentages (Figure 10). Information services are used at least once a month by 50% of the respondents and location-based services by 30% of the respondents. Email is still the most popular application, as almost 70% of the respondents use it at least once a week. It is worth noting that this is the same percentage as text message users in the European survey.



Base: 749 respondents
WAP = Wireless Application Protocol
Source: Gartner Dataquest (May 2004)

Figure 9. Mobile phone usage for different purposes in six European countries (Pittet, 2004).

The biggest challenge for mobile Internet is to provide consumers with services and applications that they really want and that they are willing to pay for. ITU summarises the main reasons behind the relative failure of WAP as a mobile Internet platform on 2G networks as slow downloading, ineffective billing models, lack of content and inappropriate interfaces in mobile phones (ITU, 2002). One of the main obstacles with WAP was the ambitious goal of making WAP a universal standard and making a big compromise between all the players, e.g. device manufacturers and operators.

Funk (2004) lists the lessons learned from mobile Internet. He disproves the popular belief that Internet penetration, popularity of public transportation, the type of mark-up language, the packet-based network or the service provider's control over phone specifications are factors behind the high volume of mobile Internet usage in Japan. Rather, he stresses that targeting the right kind of customers, providing the right kinds of content (especially email), making the services easily available on a working service network, and easy micro-payment-based charging have been the real factors behind the success of i-mode in Japan. Other success factors have been the ready-made installations and reliability, because the service entity has been built and tested with predefined service content and for predefined devices (Funk, 2004).

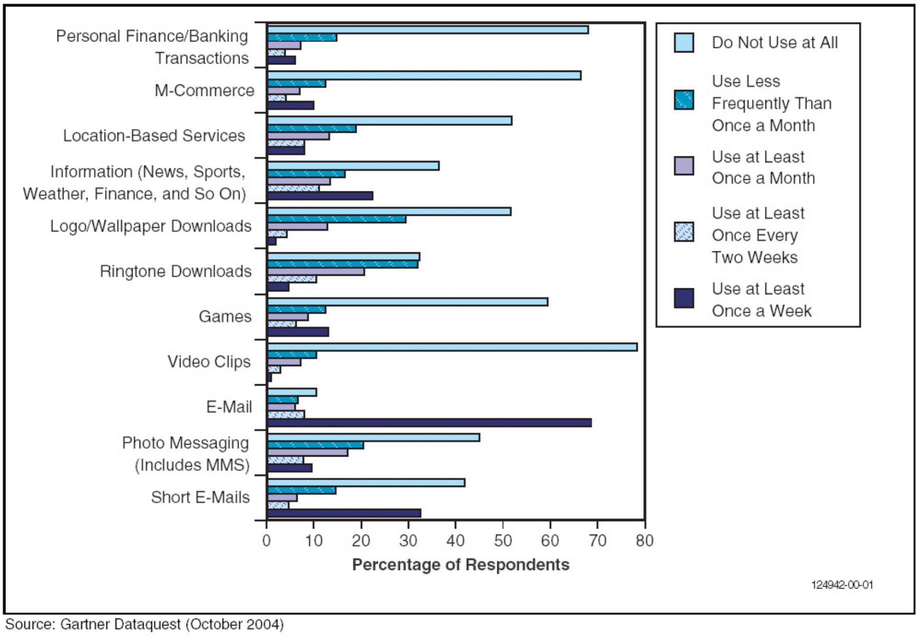


Figure 10. Usage of mobile data in Japan (Liew et al., 2004).

When the actual quality of WAP services failed to live up to the marketing hype, people soon labelled WAP as a "big failure", and attitudes towards WAP services turned negative. Also the service providers were disappointed with the actual possibilities and the teething problems of the technical infrastructures of WAP. As user attitudes towards WAP services were so negative, many service providers soon turned back to SMS-based services because SMS was reliable technology and they could reach large potential user groups with these traditional services. SMS services are still popular ways of providing for instance news and event information services. These services are relatively easy to use because they just need to be ordered once, after which the subscriber receives regular information updates automatically.

It is interesting, though, to note the similarities in the reasons for the poor performance of WAP in Europe and the slow start of NTT DoCoMo's 3G network (FOMA) in Japan: high cost of services and handsets, limited coverage area, lack of suitable content and technical problems. As FOMA is based on an almost universal standard, WCDMA, it could be argued that some of the above-mentioned problems may be related to difficulties in adopting this universal solution, especially as KDDI's competitive and less standard 3G solution, CDMA 2000, has been more successful on the market.

4.3.2 Business models for mobile services

Kallio (2004) has in her dissertation studied wireless business models. She has identified ten different business roles as described in Table 2. Wireless business networks can get quite complex, as there may be several actors in each category involved. Some actors focus on content and some on the infrastructures and technologies. Content providers need to have an insight into the possibilities of the technology and the technology providers need to have an insight into the needs of the content providers. In the Japanese i-mode model, the network provider is the node that connects all other actors, is in charge of central business roles (content integrator, infrastructure provider and network provider) and is also in charge of the whole service entity. In open business networks, which are typical of the Western business culture, the situation is much more complex.

Table 2. Wireless business roles (Kallio, 2004).

Business role	*Function*
application provider	develops horizontal or vertical applications
application service provider	provides outsourced computing power on a rental basis
content aggregator	catalyses content and markets it to other actors
content integrator	integrates content and provides it to other actors
content provider	provides content to service providers, content aggregators and integrators
infrastructure provider	provides application, network or system infrastructure
network provider	sells wireless network capacity to consumers via service providers
service provider	provides services to end users
terminal manufacturer	manufactures terminal devices
end user	individual or company that uses the service

Perhaps the most meaningful difference between WAP and i-mode has been in the business models. In Japan, the mobile Internet solutions have been strongly controlled by the operators. The operators market service entities that include the mobile phone, selected services and connections. The client gets the service entity on a turnkey basis. The selection of the services has been wide from the very beginning and the pricing, which is also controlled by the operator, is moderate.

The i-mode business model can be seen as a networked organisation, where the nucleus of the network, NTT DoCoMo, has taken a strong role in managing and controlling the network. NTT has strong alliances and partnerships that ensure that the quantity and quality of i-mode services remains good. Key partners include game companies, banks and credit card companies, database companies and entertainment businesses. The operator defines the specifications for the terminal devices and selects the terminal providers as well as the services to be provided. The operator provides consumers with ready-made packets of device and services. The service providers do not need to worry about device adaptation. However, the service providers have to think about which operator's portal they want to be in and whether they can be in several operator portals at the same time. The service has to be tailored separately for each operator portal.

In Japan, NTT DoCoMo's decision to allow users to access freely any sites enabled the creation of an entire industry of independent sites. Most of these sites are free of charge. Personal mobile Internet home pages are very popular in Japan, and the inclusion of a link to the home page in your mail signature is especially popular (Funk, 2004). As the number of official i-mode sites has grown, it has become less important to be one of the official sites, and content providers have found other ways to attract users to their sites. Portals and search engines are examples of these new service types. Funk (2004) points out that new companies rather than companies familiar from traditional Internet have succeeded in this market because of the disruptive nature of the mobile Internet market. One example is Girls Walker, which is a mail magazine based on input from 1,700 different writers, most of them writing as a hobby. Another example is Magic Island, which is a portal offering content for young people as well as the possibility to create own home pages. As people include their home page address in their mail messages and as each home page includes a link to Magic Island, the portal regularly gets new visitors.

The growing market and increasing technical sophistication of mobile entertainment content has caused a complex value chain of activities to begin to emerge in entertainment content and will likely spread to other content in Japan

(Funk, 2004). Funk presents a model of this kind of value chain (Figure 11). Service providers collect content fees from users and pass on a percentage of these fees to the content providers. At the other end of the value chain, other firms offer content providers the raw content or services such as content production or site management. In the Japanese mobile Internet market, examples of companies on the other end of the chain include companies focused on mark-up language transformation, mail system transformation, ring tone producers, and image processing services. Content production with Java is the fastest growing market as most new entertainment content is written in Java.

Funk (2004) points out that mobile Internet offers an easy and cheap way to organise for instance membership systems (such as regular customers), to offer customers discount coupons or free samples and to make surveys among customers. These are promising ways to enhance existing business with mobile services.

Rao and Minakakis (2003) point out that the key driver of location-based services will be a degree of fit between the system's technical feasibility and the overall marketing strategy guiding its usage. Several technologies and platforms need to be connected and integrated with the wireless network infrastructures, ranging from different servers to back-end databases. Solving the complex puzzle of disparate software, hardware and connectivity components poses a major challenge for the design of low-cost, reliable and high-quality systems. However, Rao and Minakakis point out that success in this area will accelerate networking effects, leading to widespread adoption, expansion of the customer base, and lower operating costs.

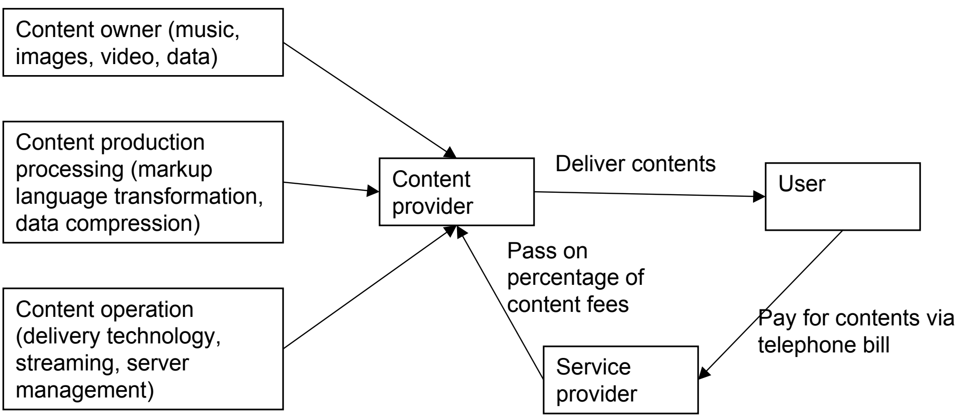


Figure 11. An example of the mobile value chain for entertainment services (Funk, 2000).

In their interviews with technical enablers of mobile services and the potential users of mobile business services in Finland, Alahuhta et al. (2005) found that a clear and shared understanding of the mobile value networks was still missing. The market is emerging and the interviewees expressed the wish to join forces in research and development to extend the value chains beyond niche technical application areas. Also the service provider interviews carried out by Kallio (2004) point out that the service providers considered the applications and services as easy to implement but found it very complicated to manage the whole delivery of the service from design to marketing. On the technical side, the interviews by Alahuhta et al. (2005) pointed out that browser-based applications seem to be the best solution for most services. Mobile services were thought to be best suited to real-time information without high requirements for large displays or input capabilities. The life span of a mobile application is around 18 months, including development. This indicates a need for faster, iterative design processes, as pointed out also by Ketola (2002) and Kallio and Kekäläinen (2004).

Mobile services have got off to quite different starts in Japan and in Europe. The Japanese operator-centric business model has been easy to handle, but more flexible business models will be needed to keep the market open for all kinds of actors. The European market has suffered from the failed introduction of mobile Internet and that is reflected in the low-profile marketing of mobile Internet services still today.

4.4 Future visions of mobile service concepts

Marcus and Chen (2002) point out two alternative trends in the development of mobile devices: the device can either be an information appliance especially designed for a certain user task or a "Swiss-army-knife", useful in many different situations. Both solutions have their benefits and drawbacks. Integrating many different functions into a single device may make the user interface complicated and the device non-intuitive (Norman, 1998). On the other hand, it might be impractical to have several different information appliances for different purposes, especially if the usage need is occasional.

According to Marcus and Chen, context-awareness, especially location-awareness, can solve problems related to baby-face displays and tedious text input by giving the user easy access to the services relevant in his/her current context. Time-shifting: "start now, finish later" will be important in mobile services, which are often used during casual empty moments. Also context-based reminders like "reply to this at home" will be useful.

Norman (Bergman and Norman, 2000) points out that future IT products will be designed for everyday persons, not for early adopters or technologists. The products are meant to fit one's lifestyle and give value and convenience. This will require a different approach to design: you have to observe the way that people live so that you can design the product to fit naturally and seamlessly into their lives.

Although the above-mentioned visionary publications are getting a bit old, the question of information appliance or Swiss-army knife still remains unanswered. Mobile phones are being equipped with more and more different functions and applications, but at the same time other everyday devices are getting equipped with communication capabilities. For instance, some new digital camera models have embedded network connections to transfer the photos more easily.

Plomp et al. (2002) predict that in the future mobile devices will increasingly interact with users and will be part of smart environments. These smart devices will consist of either a unique handset gathering various functions (computing, communication, sensing, multimedia processing, etc.) or various context-specific add-on modules (GPS, video streaming, camera, actuation, health monitoring etc.). These systems will be context-aware – they will identify the current context of use and adapt their behaviour accordingly.

Norros et al. (2003) predict that in the future, mobile devices will not just be entities in their own right but they will become tools to get information about the environment and to interact with the environment. These characteristics will become possible as more and more of the context can be measured, identified or predicted. Norros et al. suggest that introducing ambient intelligence as enhancements to current mobile services will ease the adoption of ambient intelligence and facilitate solutions where the user feels and really is in control.

Islam and Fayad (2003) point out four main challenges that need to be overcome before ubiquitous access to information anytime and anywhere can gain worldwide acceptance:

1. Designing applications for multiple devices is challenging. Successful applications will exploit the uniqueness of each device. Different devices are appropriate for different tasks and this should also guide the application design.

2. Increasing memory capacity, e.g. in the form of memory sticks, gives new potential to making mobile devices truly personal as they are able to store large amounts of personal data. Unfortunately, battery technology lags far behind the technical enhancements to mobile devices. Each hardware add-on requires additional power. This raises the need for power conscious hardware and software solutions.
3. Network heterogeneity poses a challenge for roaming. The applications have to adapt to changing network environments in real time, and connectivity management becomes a major task in the mobile devices.
4. In spite of increasing alternatives for mobile networks, the users still need the possibility to use their applications offline. This raises the major challenge of resynchronisation.

Islam and Fayad (2003) point out the necessity to develop devices in parallel with the mobile infrastructures: new types of content require new types of devices, new network options require devices that can choose the most suitable network in each context of use, increasing hardware add-ons require additional battery power, more applications and personal content require more memory capacity and so on.

An important issue for mobile services is service discovery. Today information about mobile services (e.g., ring tones, logos, and Java applications) can be found from traditional media such as newspapers. In order to have successful mobile business in the future, users should be able to find services easily from networks and from the surrounding environment in real time. The environment has to inform the user about the services available around him/her, give easy ways to take the services into use, and to adapt the services and user interface according to the usage context (Alahuhta et al., 2004).

Alahuhta et al. (2004) see a transition from network centricity towards terminal centricity taking place. Services will increasingly exist on the user's terminal instead of on network servers, and the users will be able to download applications onto their devices. Open software platforms such as Java, Symbian and Linux will enable this development. Terminals will be used for many new purposes, such as MP3 player, TV, radio, camera, remote controller, data storage and GPS.

Agent technology has long been under discussion, and in mobile services agent technology is expected to be a necessary tool to help the users cope with the vast amount of information and services. Norman (1998) points out that the value of agents is high when they work inconspicuously in the background and then, when they surface, they do it gently, offering suggestions that the human user can decide to explore or ignore. According to him, the danger of agents arises when they start wresting away control, doing things behind the user's back, making decisions on behalf the user, taking actions and in general, taking over. Norman (1998) finds the goals of agent technology attractive and points out that the agents hold great promise at some, yet unspecified, future time. Agent technology will be acceptable only when it is sufficiently intelligent and reliable. This development will take time and current technology roadmaps estimate that large-scale deployments of agent technologies will not take place until around 2012 (Alahuhta et al., 2004).

The future of mobile services seems to have several alternatives that may all come true. Mobile services will increasingly be available locally or contextually, providing users with services that optimally suit their current context of use. As service adaptation techniques develop, fixed network services will increasingly be available and automatically adapted also for mobile users. Not all services require a network connection. Mobile users will be provided with disposable applications that they can easily download and take into use when needed. When these are no longer needed, they will simply be thrown away. In addition to all this there will probably be small dedicated information appliances that are specially designed for a certain purpose, and that include all the necessary software and hardware ready to use.

4.5 Challenges to user acceptance of mobile services

With the success story of i-mode in Japan and the failure of WAP in Europe, it is tempting to conclude that user acceptance of mobile services could be guaranteed by repeating what NTT DoCoMo did and by avoiding the mistakes made with WAP. A lot can certainly be learned from the past but repeating the i-mode success is not an easy task, as NTT DoCoMo's slow start with FOMA shows. Actually, the failure with the introduction of FOMA shares many similarities with the introduction of WAP in Europe: modest selection of services, technical problems and high prices.

E-mail boosted all other i-mode services. As i-mode e-mails could include links to i-mode services, there was a ready-made channel to inform users about interesting sites and thus to encourage the use of the services. In Europe the users already have the mobile messaging service, SMS, in good use and it remains unclear whether some other service could be the one to boost mobile Internet in Europe. Messaging provides good value to the users as it enhances the role of the mobile phone as a messaging device. It is not easy to imagine another service that would provide as much value to the users.

WAP was introduced to the market as an unfinished solution using unfinished infrastructures. The marketing message did not give a realistic description of the possibilities available, installation of the browsers turned out to be complicated and the available content was very modest. Slow connections, high costs and repeated technical problems completed the failure. The users would have needed more realistic information about the services available and how to take the services into use. Compared with the i-mode approach, where the users got a ready-made package of a phone and services ready to use, the difference is obvious. The failed initial introduction of the services still affects the attitudes of the users and can be seen in the usage figures of mobile services in Europe.

The usability of mobile services has been studied a lot and several guidelines are available. Although mobile devices have developed, user interface restrictions still remain and cause usability problems that can easiest be overcome by avoiding certain kinds of elements and content in the services. The goal of universal services that would work optimally on any device and network has turned out to be extremely ambitious, as the variety of devices and technical infrastructures is constantly growing. Furthermore, the services need to adapt to changes in technical infrastructures, such as the available network on the move. Personalisation and context-awareness have been proposed as solutions to usability problems but they may bring in new ones. Perhaps even bigger problems than ease of use have been the technical reliability of the services as well as difficulties in installing and configuring the services.

Future visions of mobile services present solutions where information about the user is increasingly collected and transferred to the services, for instance to provide the users with contextually and personally relevant services. This trend highlights the need for privacy protection. As mobile devices are increasingly used to interact with the environment, service discovery and adoption become more important.

The life spans of individual mobile services are quite short and they do not allow extensive human-centred design activities. Lack of sustainable business models makes managing the delivery of the service from design to marketing complicated and thus limits the number and quality of services available.

The challenges for user acceptance of mobile services arise from the increasingly complex technical and business environments, where the users access the services with mobile devices that have several user interface restrictions. At the same time the pressure for faster development cycles does not allow extensive usability studies, and the services may end up on the market with severe usability and technical problems as well as inadequate content. In this challenging design environment there is a need to extend the focus of current usability-oriented design guidelines to better cover the design decisions that affect user acceptance.

5. Related research

5.1 Approaches to study user acceptance of technology

This chapter gives an overview of different approaches to study user acceptance of technology: the Human-Centred Design process, which aims at producing usable and thus acceptable products; the Technology Acceptance Model, which predicts usage behaviour based on user and environment characteristics as well as perceived product attributes; and the Innovation Diffusion Theory, which studies the likelihood and the rate of an innovation being adopted by different user categories. This chapter also introduces the Technology Hype Cycle, which describes the almost inevitable progress of new technologies from over-inflated expectations created by hype and over-enthusiasm to disappointments and negative hype and then again to a new start from more realistic ground.

5.1.1 Human-centred design process

Nielsen defined usability engineering practices in his book Usability engineering (Nielsen, 1993), which has become one of basic textbooks on usability engineering. Nielsen's usability concept (Chapter 3.3) was mainly related to user interfaces and thus the usability engineering practices that he proposed were also focused on designing user interfaces. Later on, the focus of human-centred design has been shifting more towards the contexts of use. The ISO (1998) definition of usability (Chapter 3.3) emphasises context of use, and also the definition of human-centred design process in ISO 13407:1999 (ISO, 1999) has a strong emphasis on the intended contexts of use. ISO13407:1999 defines the design process to make systems usable. The standard is targeted at project managers and it defines how to organise and manage a human-centred design process.

According to ISO 13407:1999 (ISO, 1999), the incorporation of a human-centred approach is characterised by:

a) active involvement of users and clear understanding of user and task requirements,
b) an appropriate allocation of functions between users and technology
c) the iteration of design solutions
d) multidisciplinary design.

According to the standard (ISO, 1999), the goals of individual evaluation activities can be to assess how well the system meets the goals of the user or the organisation, to diagnose potential problems and identify needs for improvements, to select the proposed design option that best fits the user and organisational goals and/or to elicit feedback and further requirements from the users.

The human-centred design process consists of successive cycles of specifying the intended context of use, specifying user and organisational requirements, producing design solutions, evaluating the solutions against the requirements, and then again refining the context of use as illustrated in Figure 12.

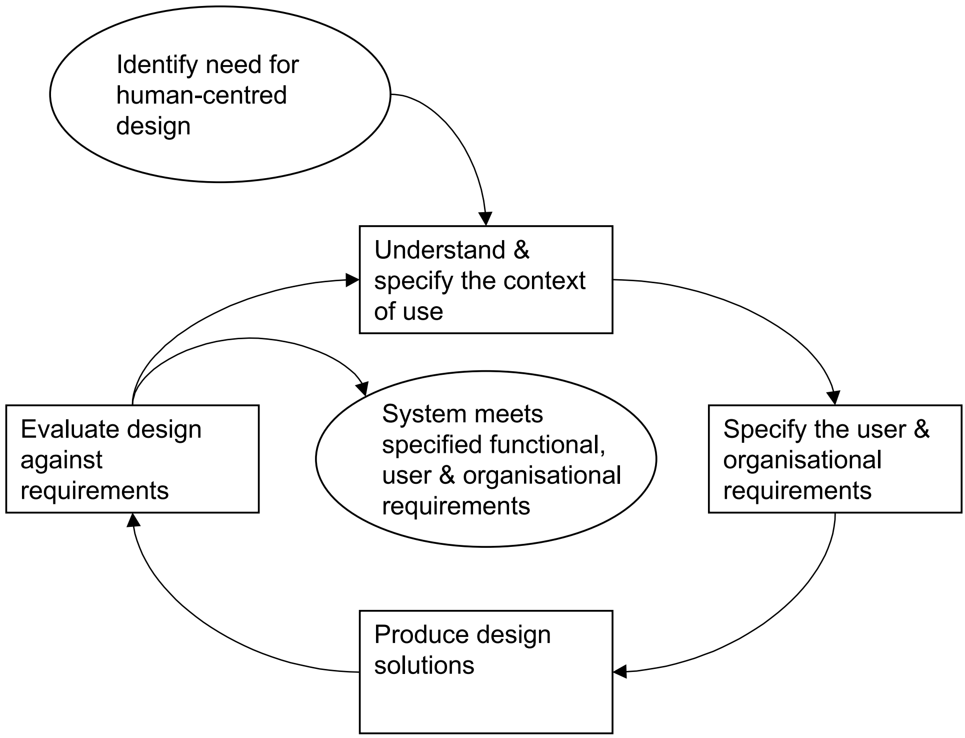


Figure 12. Human-centred design process (ISO, 1999).

ISO 13407:1999 has been criticised because it does not take into account current software engineering practices. Software engineering is today based on system modelling, and prototyping suggested by the human-centred design approach is technically very difficult, especially as the systems are getting more complex, services are distributed and design work is concurrent. Gulliksen et al. (2003) claim that this basic difference in the approaches is an obstacle to the adoption of human-centred design in software engineering. Ketola raised the same theme is his dissertation (Ketola, 2002).

In some approaches the main focus of the design is on the usage situations. Contextual design (Beyer and Holzblatt, 1998) emphasises the significance of clarifying the task context, including the goals and conditions of action. Information technology products need to be designed in parallel with the redesign of working practices. User experience research focuses on studying how different users experience the usage of a product in different situations (Battarbee, 2004).

Norman (1998) points out that designers need answers in hours, not months. According to him applied science does not need the precision of traditional scientific methods used in social and behavioural sciences. In product design, the focus is on seeking big phenomena and they can be found with simpler methods. Norman (1998) proposes rapid ethnography as a research method to invent new product classes. Rapid ethnography includes not just observing potential users but making these people participants in the discovery process of learning what their real needs are.

Kallio and Kekäläinen (2004) also point out that product lifecycles of mobile services are short, varying from one month to one year. In this kind of development environment there are needs for less time-consuming and less expensive methods. Kallio and Kekäläinen propose an approach where the collaboration between the Human-Computer Interaction (HCI) department and the marketing department within a company is improved. As marketing departments are carrying out focus group interviews, the participants of the focus groups can be involved in usability tests right after the group sessions. To facilitate having several test users in parallel, the tests are conducted by having user-pairs in the tests and non-expert moderators. In the study by Kallio and Kekäläinen (2004), the user-pair tests revealed nearly half of the problems that were identified in individual usability tests.

Regarding technology research, the shortcoming of both human-centred design approach and contextual design is the product-centredness of these approaches. Both approaches are designed to guide a design team that already has the mission of designing a certain kind of product. User feedback is collected to guide the design of this individual product. In technology research the aim is rather to identify how a certain technical solution could be utilised in different applications, demonstrate these applications, and study user acceptance of these applications to get information to guide future commercial development. This indicates that the focus needs to be enhanced to business and marketing research methods such as the Technology Acceptance Model and the Innovation Diffusion Theory, which are described in the next chapters.

5.1.2 Technology Acceptance Model

Davis presented in 1989 the Technology Acceptance Model (TAM) to explain the determinants of user acceptance of a wide range of end-user computing technologies (Davis 1989). The model is based on the Theory of Reasoned Action by Ajzen and Fishbein (1980). TAM points out that perceived ease of use and perceived usefulness affect the intention to use. Davis (1989) defines perceived ease of use as "*the degree to which a person believes that using a particular system would be free from effort*" and perceived usefulness as "*the degree to which a person believes that using a particular system would enhance his or her job performance*". Perceived ease of use also affects the perceived usefulness (Figure 13). The intention to use affects the real usage behaviour. TAM has been tested and extended by many researchers, including Davis himself. TAM has been designed to study information systems at work to predict if the users will actually take a certain system into use in their jobs. The model provides a tool to study the impact of external variables on internal beliefs, attitudes and intentions. TAM has been applied mostly in studying office software usage and in that application area the model can explain about 40% of system use (Legris et al., 2003).

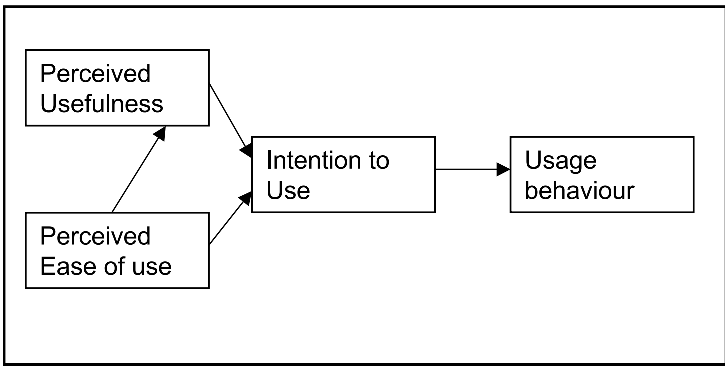


Figure 13. Technology Acceptance Model (Davis, 1989).

TAM deals with perceptions; it is not based on observing real usage but users reporting their conceptions. The instruments used in connection with TAM are surveys, where the questions are constructed in such a way that they reflect the different aspects of TAM. The survey questions related to usefulness can be for instance "Using this system improves the quality of the work I do" or "Using this system saves my time". The survey questions related to ease of use can be for instance "The system often behaves in unexpected ways" or "It is easy for me to remember how to perform tasks using this system".

Venkatesh and Davis (2000) have enhanced the model to TAM2 (Figure 14), which provides a detailed account of the key forces underlying judgements of perceived usefulness, explaining up to 60% of the variance in this driver of usage intentions. TAM2 reflects the impacts of three interrelated social forces: subjective norm, voluntariness and image. Venkatesh and Davis (2000) define subjective norm according to Fishbein and Ajzen (1975) as "*person's perception that most people who are important to him think that he should or should not perform the behaviour in question*". Venkatesh and Davis (2000) define voluntariness as "*the extent to which potential adopters perceive the adoption decision to be non-mandatory*" and image as "*the degree to which use of an innovation is perceived to enhance one's status in the social system.* TAM2 showed that both social influence processes (subjective norm, voluntariness and image) and cognitive instrumental processes (job relevance, output quality, result demonstrability and perceived ease of use) significantly influenced user acceptance.

Mathieson et al. (2001) have extended the Technology Acceptance Model by analysing the influence of perceived user resources. They claim that there may be many situations in which an individual wants to use an information system, but is prevented by lack of time, money, expertise and so on. Mathieson et al. classify resource-related attributes into four categories: user attributes, support from others, system attributes and general control-related attributes that concern an individual's overall beliefs about his/her control over system use. In their extended model, external variables affect perceived resources that further affect perceived ease of use and behavioural intention to use.

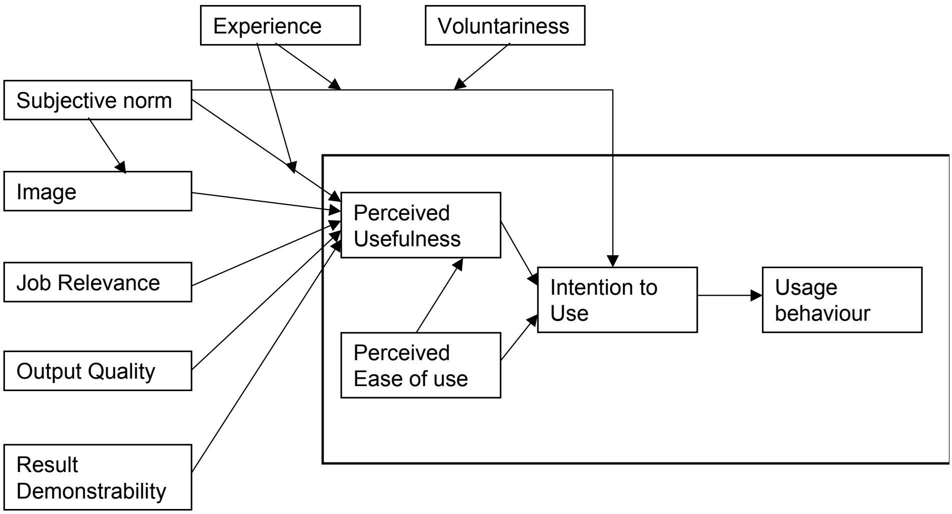


Figure 14. Enhanced Technology Acceptance Model (TAM2) by Venkatesh and Davis (2000).

TAM was originally developed for studying technology at work. Later it has been often used to study user acceptance of Internet services (Gefen, 2000; Gefen and Devine, 2001; Gefen et al., 2003; Barnes and Huff, 2003; Chen et al., 2004;). Gefen et al. (2003) have studied TAM in connection with e-commerce. They have extended TAM for this application area and propose that trust should be included in the research model to predict the purchase intentions of on-line customers. Gefen et al. define trust as the expectation that the trusted party will behave in an ethical, dependable and socially appropriate manner and will fulfil their expected commitments in conditions of interdependence and potential vulnerability. Chen et al. (2004) also include trust together with perceived service quality and compatibility in their TAM-based model for user acceptance of virtual stores.

Venkatesh et al. (2003) have proposed a unified view for the user acceptance model. They have combined the original TAM with seven other user acceptance research approaches, including the Innovation Diffusion Theory (described in the next sub chapter 5.1.3). The Unified Theory of Acceptance and Use of Technology (UTAUT) includes four key determinants of intention and usage as illustrated in Figure 15: Performance expectancy, Effort expectancy, Social Influence and Facilitating conditions. Performance expectancy replaces the Perceived utility in the original TAM and Effort expectancy replaces Perceived ease of use. Social influence and Facilitating conditions have been adopted from the other research approaches.

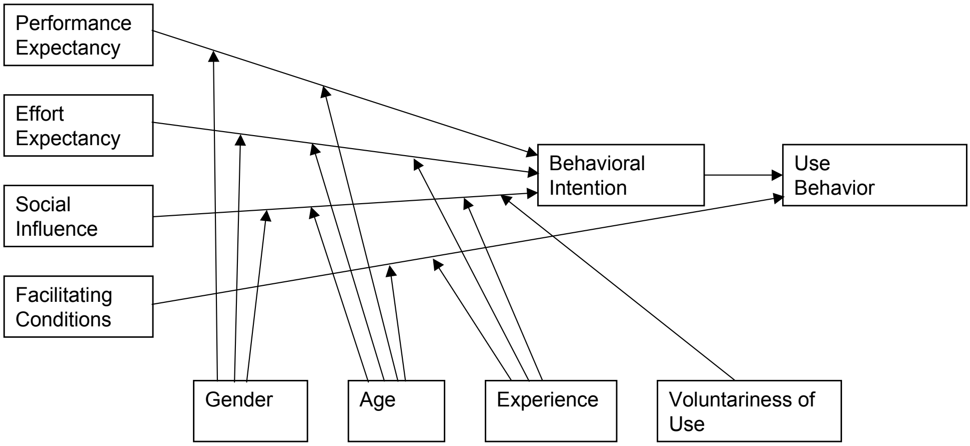


Figure 15. Unified Theory of Acceptance and Use (Venkatesh et al., 2003).

TAM has been widely used as a framework for surveys where existing products are assessed. Davis and Venkatesh (2004) proved out that stable and behaviourally predictive measures of perceived usefulness of information systems can be made by using mock-ups. Traditionally, mock-ups are used in human-centred design to evaluate proposed designs with users for ease of use. Davis and Venkatesh (2004) suggest that mock-ups should increasingly be used to assess the usefulness of the proposed system at a pre-prototype phase of the project.

The Technology Acceptance Model constitutes a solid framework to identify issues that may affect user acceptance of technical solutions. As Davis and Venkatesh (2004) have proved, the model can be enhanced from the original purpose of studying user acceptance of existing products to study planned product concepts e.g. in the form of mock-ups. This indicates that TAM could also be used in connection with technology development projects and processes to assess the usefulness of proposed solutions. Applied in this way, the model also supports the human-centred design approach.

5.1.3 Innovation Diffusion Theory

The Innovation Diffusion Theory (IDT) introduced in 1962 and later refined by Rogers (1995) is similar to TAM based on the Theory of Reasoned Action (Ajzen and Fishbein, 1980). IDT is a well-established theory to study user adoption of different innovations in target populations, not only technical but all kinds of new ways to act. The theory explains the process of the innovation decision process, the determinants of the rate of adoption, and various categories of adopters. The theory aims at predicting the likelihood and the rate of an innovation being adopted by different adopter categories. Rogers (1995) defines five factors that explain 49–87 per cent of the variance in the rate of the adoption of an innovation:

- Relative advantage is the degree to which the innovation is perceived as being better than the practice it supersedes.
- Compatibility is the extent to which adopting the innovation is compatible with what people do.
- Complexity is the degree to which an innovation is perceived as relatively difficult to understand and use.
- Trialability is the degree to which an innovation may be experimented with before making the adoption or rejection decision.
- Observability is the degree to which the results of an innovation are visible to others.

Other researchers have extended Roger's work (Karahanna et al., 1999; Barnes and Huff, 2003), suggesting additional factors for the model:

- Image is the degree to which adoption and use of the innovation is perceived to enhance one's image or status.
- Trust is the extent to which the innovation adopter perceives the innovation provider to be trustworthy.

Rogers (1995) defines five adopter categories: innovators, early adopters, early majority, late majority and laggards (Figure 16). These categories illustrate variability around the mean, when half of the target population has adopted an innovation. Rogers describes innovators as venturesome risk-takers who serve as gatekeepers for those who follow. Early adopters are opinion leaders who are the first within their group to adopt, and willing to maintain their position by evaluating innovations for the others. Early majority includes users who are deliberate in their adoption decision but want to wait until others have assessed the innovation. However, they do not want to be the last to change. The late majority includes sceptical users who prefer to wait until most others have adopted the innovation. The last to adopt are the laggards, who base their decisions on the past rather than the future. Rogers regrets the selection of the term "*laggard*" and emphasises that it would be a mistake to imply that laggards would be somehow at fault for being late to adopt. Resistance to the innovation may be entirely rational from the laggard's point of view. For instance, his/her economic position may force him/her to be extremely cautious in adopting innovations.

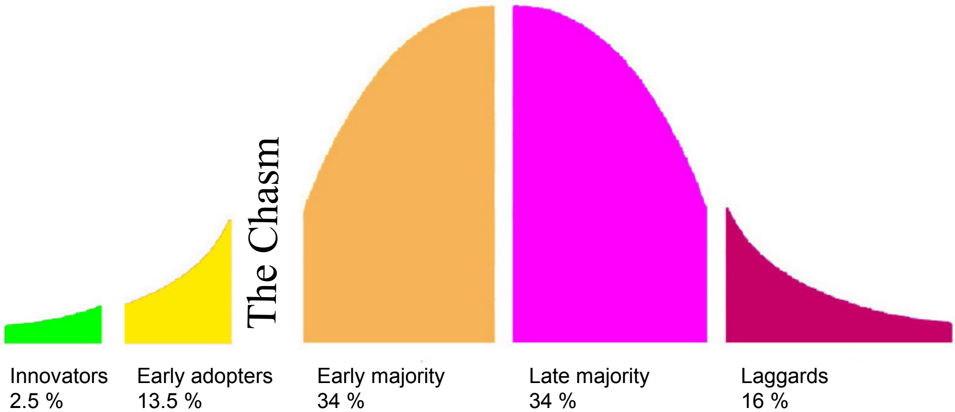


Figure 16. Innovation adopter categories (Rogers, 1995) with the chasm as defined by Moore (1999).

Rogers (1995) points out that often innovations are not viewed singularly by individuals but instead they may be perceived as an interrelated bundle of new ideas. The adoption of one idea may trigger the adoption of others. Rogers states that unfortunately the effects of using a package approach have seldom been used in diffusion research, even though it would make sense to study whole technology clusters or the user's perceptions of interrelated innovations.

According to Rogers (1995) the critical mass occurs at the point at which enough individuals in a system have adopted an innovation so that the innovation's further rate of adoption becomes self-sustaining. The critical mass is particularly important in the diffusion of interactive innovations such as e-mail, where each additional adopter increases the utility of adopting the innovation for all adopters.

Moore (1999) extends Rogers' work by referring to differences across categories as "cracks in the bell curve". His research suggests that innovations that succeed among innovators or early adopters may fail among the early majority or late majority, if the innovation lacks characteristics that appeal to these groups. Moore claims that the chasm – the different needs of early majority compared to early adopters – needs to be bridged if an innovation is going to be successful in the mass market (Figure 16).

Norman (1998) points out that all new technologies take a long time to affect the lives of ordinary people. Especially with disruptive technologies the adoption takes a long time and is a complicated process. By disruptive technologies Norman means technologies that are not just incremental changes to current technologies but technologies that cause revolutionary changes in people's lives. Disruptive technologies may change the entire course of an industry, spawning new companies and killing off existing ones. Norman claims that new approaches always start small, simple and weak. At this stage the innovators and early adopters as defined by Rogers (1995) adopt because they value technology and performance. After these adopters the new innovation has to bridge the chasm as defined by Moore (1999). The early and late majorities value new solutions and convenience. Norman (1998) claims that at this stage it is not enough to fine-tune the product and change the marketing; rather, the whole product has to be different. So, he suggests that a whole new product development process needs to be carried out to cross the chasm.

Even though the Innovation Diffusion Theory targets the innovation adoption phase, which is usually beyond the technical research work, it can be utilised also in the early research phases. Test users can be categorised into different adopter groups and these groups can be weighted in the evaluations according to the objectives of the evaluation. The results of early evaluations of prototypes can give clues to the characteristics of the product that cause the chasm defined by Moore (1999). The suggestion of Norman (1998) that totally new products are needed for the early majority is also worthy of consideration. However, it has to be kept in mind that the innovators may not always be technology freaks as the Innovation Diffusion Theory is often interpreted. Sometimes for these early adopters the value of the product is so high that they are ready to adopt in spite of foreseen usability and technical problems.

5.1.4 Hype cycle of technologies

New emerging technologies tend to have a high profile in the media and public debate compared with more mature technologies. This often leads to over-inflated expectations, so-called hype, around a particular technology. Gartner Research illustrates this phenomenon by hype cycles that aim at giving an overview of the relative maturity of technologies in a certain domain compared with their visibility (Linden and Fenn, 2003). The hype cycle characterises the typical progression of an emerging technology from business and media over-enthusiasm through a period of disillusionment to an eventual understanding of the technology's relevance and its role in a market or a domain. Figure 17 illustrates the general structure of a hype cycle.

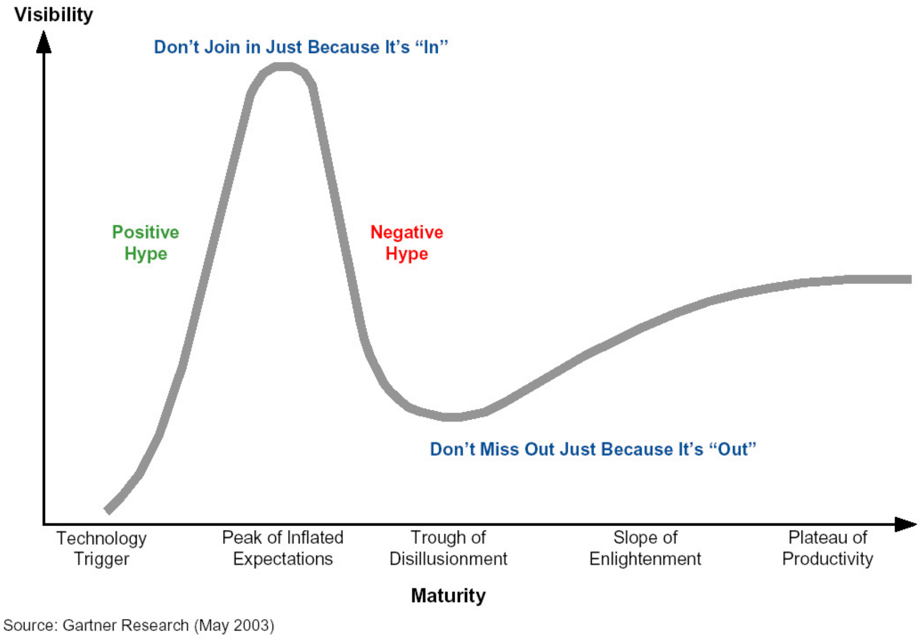


Figure 17. The hype cycle (Linden and Fenn, 2003).

Linden and Fenn (2003) introduce the phases of the hype cycle in the following way:

> *Technology trigger:* a breakthrough, public demonstration, product launch or other events generate significant press and industry interest.
>
> *Peak of inflated expectations:* over-enthusiasm and unrealistic projections result in some successes but more failures, as the technology is pushed to its limits.
>
> *Trough of disillusionment:* because the technology does not live up to its over-inflated expectations it rapidly becomes unfashionable. Media interest wanes, except for a few cautionary tales.
>
> *Slope of enlightenment:* focused experimentation and solid hard work by an increasingly diverse range of organisations lead to a true understanding of the technology's applicability, risks and benefits.

Plateau of productivity: the real-world benefits of the technology are demonstrated and accepted. The final height of the plateau varies according to whether the technology is broadly applicable or benefits only a niche market.

The hype cycle typically occurs early in the technology's life cycle. The peak of hype occurs when there is almost no adoption in the marketplace and the performance of products (if there are any) is poor. Some technologies experience multiple rounds of vacuous hype before beginning a serious growth path. Another "mini-peak" of hype may occur in connection with climbing the slope of enlightenment. Each technology's hype cycle is different depending on the technology's perceived importance in the business and the society. (Linden and Fenn, 2003)

The typical timescale for a particular technology to be really mature and make a profit is 5–10 years after the first launch (Alahuhta et al., 2004). New technologies have often already been analysed, worn out and condemned in public debate before the technology is actually mature for large-scale deployment. Hype is generated by several actors: researchers, industry and media, pressing industrial actors for early launches of technologies. Launching immature technologies with incomplete business models then turns the hype curve into steep descent. The new start with a more mature version of the technology is then characterised by a gentle upslope.

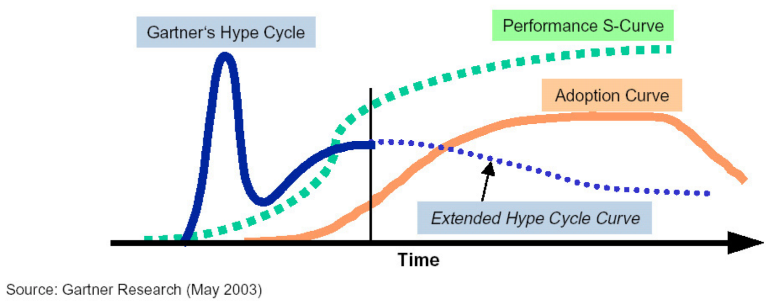


Figure 18. The relative timing of the hype cycle and the Innovation Diffusion (adoption) curve together with the rising curve of technology performance (Linden and Fenn, 2003).

Figure 18 (Linden and Fenn, 2003) presents the Extended Hype Cycle Curve as well as the Adoption Curve of the Innovation Diffusion Theory on the same time axis. Figure 18 also includes the Performance S-Curve, which indicates the performance and thus maturity of the technology. Although the curves are presented on the same time axis, it must be pointed out that the y-axis of Figure 18 indicates different things, depending on the curve in question.

The trough of disillusionment of the hype cycle coincides with the start of the innovation diffusion curve by Rogers (1995), as illustrated in Figure 18. Innovators may start adopting at the trough of disillusionment, but early adopters start adopting only at the slope of enlightenment. The visibility of the technology starts to decrease as the early majority start adopting – the technology starts to get embedded in everyday life so that it does not raise much discussion any more. This phenomenon is seen as the Extended Hype Cycle is in retreat. At that time the performance of the technology has improved remarkably compared with the level at which it was when the innovators were adopting, described by the Technology Performance S-Curve.

Linden and Fenn (2003) identify three adoption speeds for technologies:

1. Fast-track technologies go through the hype cycle within two or three years. These technologies are typically adopted without many fanfares, bypassing the peak of inflated expectations and the trough of disillusionment. High value to the users, simplicity of use, several strong vendors and use of current infrastructures are typical of fast-track technologies. SMS (Short Message Services) is perhaps the most famous example of a fast-track technology.
2. Long-fuse technologies may take one or two decades to traverse the hype cycle. Science-fiction style fascination of the technology in the media, inherent complexity, reliance on a new infrastructure and required changes in business processes are typical of long-fuse technologies. For instance e-mail and Internet have gained their current success via a long-fuse hype cycle.
3. "Normal" technologies usually traverse the hype cycle in five to eight years. These technologies usually introduce only small changes to existing user practices and technical infrastructures.

Not all technologies ever get to the plateau of productivity. The technology may get embedded in another one, the technology may get split into several sub-concepts, or the technology may become a "ghost" that remains on hold in the trough of disillusionment. "Phoenix" technologies continually cycle through enthusiasm and disillusionment. These technologies are usually extremely slow-moving technologies with principal methodological or scientific challenges (Linden and Fenn, 2003). Intelligent agents, studied in Paper I, are one example of phoenix technologies.

The hype cycle is a good tool to understand the inevitable pattern of excitement and disillusionment about technologies. It helps to understand why the recommendations from technology planning groups may be different than what is heard or read in the media. The lesson to learn is that enterprises should not invest in technologies just because they are being hyped. On the other hand, technologies should not be ignored just because they are currently not living up to early over-inflated expectations. (Linden and Fenn, 2003)

5.2 Studies on user acceptance of mobile devices and services

Kiljander (2004) points out that mobile phones need to be designed to be both intuitive for first-time use and efficient in long-term use. The devices need to be designed for gradual learning so that the user can explore the functionalities and gradually become an expert in using the device. Kiljander (2004) points out that purchasing decisions are strongly affected by the features of the device even though the consumer may never actually use all that functionality. In real use, the usability of the most frequently used features becomes more important.

As mobile phones are increasingly becoming everyday tools rather than status equipment, Kiljander's suggestion of feature-phased purchasing decisions can be questioned. According to a user survey by Gartner in France, Germany, Italy, Spain, Sweden and United Kingdom, battery life, size, weight and price were the most important considerations for people buying mobile phones (Pittet, 2004).

Funk (2004) claims that mobile Internet is one of the disruptive technologies that improve some aspects of performance compared with earlier technologies, but at the same time reduce others. This causes a change in the customer population: the first adopters may be totally different to the lead customers of the old technology. Mobile Internet is most appropriate for customers who value portability over limitations in display, keyboard and memory. Funk points out that one of the i-mode success factors was the creation of a totally new development and marketing organisation, with new leaders and a new focus on totally new customers – in the case of i-mode, young people. Focusing on new customers made it possible to focus on new kinds of services, mainly entertainment, shopping and games, which in addition to email constituted the main part of the Japanese mobile Internet market in 2002.

Sarker and Wells (2003) have studied key factors affecting the use and adoption of handheld hybrid mobile devices that offer both voice and data features. Rather than instantiating existing theories such as the Technology Acceptance Model and Diffusion of Innovation framework, they have taken an approach grounded in practice. Sarker and Wells claim that existing theories might ignore unique issues associated with mobile devices. The shortcoming of their study is the modest research material: the study is based on a three-week field trial with 21 participants in a rural university setting in the USA. Sarker and Wells (2003) propose an Input-Process-Output model-based framework to study the adoption of mobile handheld devices. Input includes factors related to the intention of use, such as user characteristics, task characteristics, technology characteristics, modality of mobility and the surrounding context. Process includes factors related to actual use divided into two interacting use subprocesses *Exploration and experimentation* and *Assessment of the experiences*. A positive experience with the use process, reflected in favourable assessment in terms of functional, psychosocial and relational outcomes, influenced the adoption decision and behaviour, i.e. the Output of the process. Sarker and Wells claim that this pattern is likely to hold true especially in a voluntary technology adoption scenario.

Järvenpää et al. (2003) have carried out a cross-cultural study of mobile handheld devices and services involving 32 focus groups with nearly 200 active urban mobile device users in Finland, Japan, Hong Kong and the USA in 2001. According to their study, Western users placed more significance on freedoms related to effectiveness at work and those in the individual sphere. Asian participants valued especially new freedoms of interpersonal relationships and emotional expression. In Finland and the USA individual needs, values and goals were emphasised in the results, whereas in Japan and Hong Kong the needs, values and goals of a group took precedence over those of the individual. Järvenpää et al. (2003) cite as an example that in Japan the mobile phone has not yet become part of many serious professional communications. In a collectivist culture, people are unlikely to use a mobile device to interact with others who do not belong to the same group.

In all four countries, the study by Järvenpää et al. (2003) revealed a predominant perception that the quality of mobile services is low and mobile services are difficult to use. However, in Japan 45% of the focus group members were using mobile Internet sites regularly, whereas in Finland the percentage was 3 and in Hong Kong 4 (unfortunately in this study most figures from the American focus groups include wired access and are thus not comparable with the figures from other countries). In Finland banking services were used regularly by 15% of the focus group members, whereas the percentage was only 8 both in Hong Kong and in Japan. Messaging was very commonly used in all the countries, only the format varied from paging to text messages and email. The members of the focus groups were also asked about the perceived importance and value of different mobile services using a scale from 1 (useless) to 10 (very valuable). In all the countries different messaging services were assessed moderately valuable and scaled around 7. Banking was valued around 5 and games around 3–4 in all the countries. The value of mobile Internet sites was assessed very differently: In Finland the value was around 3 whereas in Hong Kong it was as high as 8 and in Japan 6. These figures are in line with the survey results by Gartner presented in Chapter 4.3.1.

Based on their study, Järvenpää et al. (2003) claim that consumers may lack a compelling motivation to adopt new pay-per-use service offerings unless those services create new choices where mobility really matters and manage to affect positively people's lives. Järvenpää et al. point out the need for flexibility and malleability in technology, so that users can innovatively shape the technology to their individual and group needs in various social and business contexts.

Barnes and Huff (2003) have used the enhanced innovation diffusion theory to explain the success of Japanese i-mode:

- Relative advantage is achieved by the unmet Internet demand and low price. Other researchers (e.g., Funk, 2004) do not agree with this point. Actually, the relative advantage has probably been the email because it made possible text-based mobile messaging that was previously possible only with pagers.
- Compatibility is achieved by using the familiar mobile phone as the platform, and using a micro-browser that is a scaled-back version of traditional browsers.
- Complexity is avoided by using a simple menu-based interface and intuitive command navigation button.
- Trialability is encouraged by pay-per-use pricing and low costs that encourage users to share their devices with others for trial.
- Observability is not a key success factor since most i-mode usage is quite invisible. This can also be disagreed with because, as Funk (2004) describes, email signatures turned out to be an efficient way to make mobile Internet sites observable, thus diffusing the innovation.
- Image and trust are provided by using the established and respected service provider, NTT DoCoMo.

Barnes and Huff (2003) claim that the technology acceptance theory operates principally at the level of an individual. There are also wider economic and technical factors that play important roles. The market situation and the central role of the operator (NTT DoCoMo) were already discussed in Chapters 4.1.2 and 4.3. In addition Barnes and Huff point out the effect of self-enforcing – the services boost each other, e.g. information access generates more phone calls through getting contact data to different points of interest.

Kindberg et al. (2004) have studied users' perception and reasoning of security and trust in mobile interactions with a case study of an electronic wallet used for payment in a restaurant. In web-based e-commerce services user trust is built on user's previous experience or familiarity with a particular site or vendor. Kindberg et al. point out that in mobile settings the users have to make dynamic decisions about the trustworthiness of the services with little if anything known a priori about the other parties in the interaction. The study by Kindberg et al. revealed that there was only a loose mapping between the actual technical risks in the system and the subjects' perception of them. The users were more willing to trust the technology than the human beings that might misuse the systems. In the test set-up most users selected their preferred payment method based on convenience of use, with trust and social acceptability being less important selection criteria.

Billsus et al. (2002) emphasise that a literal transition from the Web to mobile Internet is inadequate because the presentation of information needs to be tailored for mobile devices. Agents that select information for the user are convenience in the desktop environment, but in mobile use they are essential. Billsus et al. have studied adaptation and personalisation in connection with mobile news and advertisements. According to them, the need for adaptive personalisation arises from three key features of mobile use. Firstly, manual personalisation usually cannot be done with the mobile device but it has to be done on the Web, heightening the already high threshold of actually bothering to personalise. Secondly, the personalisation options tend to be too coarse-grained, e.g. "International news" instead of "Telecom industry in Korea". Thirdly, user interests change over time. For instance, many people are not interested in sports as such, but they may gain interest in some sports news during the Olympics.

Billsus et al. (2002) propose the following guidelines for user acceptance of adaptive personalisation:

- The first use should provide an acceptable, non-personalised experience. The transition from a non-personalised version to a personalised experience should be quick and smooth.
- Adapt quickly to changing interests but avoid brittleness.

- Avoid tunnel vision, personalisation should not get in the way of getting novel information.
- Approaches where content providers have to hand-tag their content do not work in practice. Look for automated solutions.
- Support multiple modes of information access, e.g. links to related stories turned out to be a popular way to access news items.
- Respect privacy; leave the users the freedom of choice to remain anonymous.

Pousman et al. (2004) have studied a location-aware event planner on PDA devices. Based on a small-scale field study, they propose some considerations that may increase the acceptability of location-aware services. In their study, interaction elements derived from the desktop (e.g. buttons and lists) were one of the primary sources of mistakes for users, and that is why they propose using device-specific interaction elements that may be more usable. Resuming tasks easily after an interruption is a fundamental need in mobile environments. This can be supported e.g. by atomic interaction sessions, by appropriate timeouts on unfinished operations, and by a stateless interaction model. Integration of different features is needed because of the lack of temporary storage. Integration should provide hooks for moving from one interaction to another with intermediate data. Also, data should be integrated to reduce the need to access multiple sources.

As context is predicted based on measurements and the user's previous behaviour, context information is seldom accurate. That is why context-aware systems have several error possibilities: the system may offer the user wrong things either because it predicted the context wrongly or because it predicted the context correctly but predicted the user’s needs in that context wrongly. Antifakos et al. (2004) have studied the effect of displaying uncertainty in context-aware systems. Their test set-up was quite a simple task of organising numbers into a given order. The users got guidance from the system with information on the uncertainty of the advice. Giving the uncertainty information improved hit rates. Antifakos et al. point out that further work will be needed to study the trade-off between the cognitive load and the added value of displaying uncertainty information. However, a remarkable finding of their study was that several participants reported that as uncertainty information was displayed, it was easier to understand what the system was doing and how well it was doing it. Thus displaying uncertainty may improve the acceptability of context-aware systems by making them more intelligible (Antifakos et al., 2004).

5.3 Applicability of current approaches to studying user acceptance of mobile services

Human-centred design is a well-established practice to design individual products for usability. The design process includes successive evaluations of the proposed solutions with intended users. In the early phases of the design, the evaluations in laboratory conditions give feedback mainly on ease of use but as the users are provided with realistic prototypes that they can use in long-term field trials, feedback can also be collected about utility. When field trials are included in the design process, the focus of usability studies is enhanced from the traditional usability conception of specified users performing specified tasks in specified contexts of use. In field trials the users can use the prototype services as part of their everyday life. The research framework can then be enhanced to identify the actual tasks that users want to perform and the actual contexts of use. Technology acceptance models provide a framework for these kinds of studies.

Technology acceptance models aim at studying how individual perceptions affect the intentions to use information technology and further the actual usage (Figure 19). The original Technology Acceptance Model, TAM (Davis, 1989) was developed for studying user acceptance of information systems in the office environment, and most of the later enhancements and modifications to the model are also focused on business software applications. Legris et al. (2003) have made a review of published studies on the applications of TAM, and made a wider analysis of 22 studies. They identified three significant limits in TAM research: the narrow focus of the applications with the main emphasis on office automation, narrow user groups with many studies made with students as users, and measurements based on self-reported use rather than observing actual usage. By applying TAM in connection with field trials, user acceptance can be studied with real users and actual usage situations.

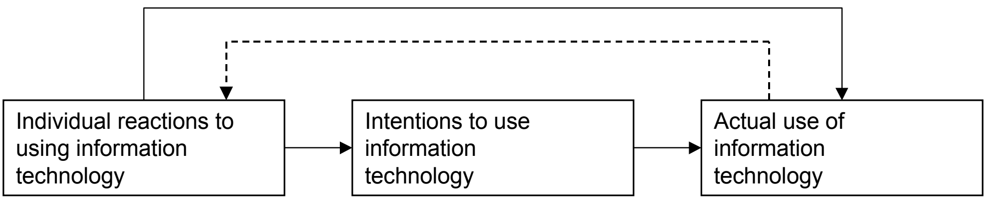


Figure 19. The basic concept underlying user acceptance models (Venkatesh et al., 2003).

To study user acceptance of mobile services, a framework is needed to analyse and synthesise the individual findings on factors that affect user acceptance. This framework can then be used to communicate key user acceptance factors and their implications for the design of future services. The focus thus differs from the focus of TAM, which takes the information system to be studied as a given fact and focuses on explaining the underlying forces affecting user acceptance. These underlying forces are mainly related to user characteristics and the social and organisational environment. The Innovation Diffusion Theory (IDT) that has been used in studying user acceptance of consumer products is also focused on the dissemination of ready-made products, and thus it also concentrates in user characteristics and the social environment rather than product characteristics. TAM and IDT have been widely used and adapted to predict potential user acceptance, but as such they do not provide assistance in the design and development of systems that would be accepted by the users.

As described in chapters 3 and 4, mobile services targeted at consumers have several specific characteristics that may mean that their user acceptance cannot be studied using the same models as with information systems in the workplace. When dealing with consumer services, individuals make voluntary adoption decisions and thus the acceptance includes assessing the benefits provided compared with either competing solutions or the non-acquisition of the service in question. In a marketing situation mobile services have to compete not only with other mobile services but also with alternative solutions in other media. As pointed out by Funk (2004), mobile services are disruptive technology that may find their innovator adopters elsewhere than expected, as highlighted by experiences with the Japanese i-mode. Focusing too early on only limited user groups may miss possible early adopters. As proposed by Norman (1998), it is important to cover different user groups in user acceptance studies to identify the factors that cause the chasm between early adopters and early majority. With the Japanese i-mode, other services were boosted through e-mail and personal home pages. This suggests that the focus of user acceptance studies of mobile services should be extended to interrelated innovations as proposed by Rogers (1995).

Consumer applications are acquired by the user him/herself rather than being provided ready to use in environments such as the workplace. An individual consumer needs to get information about the product, be able to assess it, acquire it, install and learn to use it. In workplaces the users usually get support for installation, but with consumer services they have to manage on their own. Mobile devices include several user interface limitations that pose usability challenges for mobile services. Complex technical infrastructures and service networks cause problems for the users in understanding where to find the services, who is responsible for providing the services and what kind of technology the services will require. The demanding technical environment with an increasing variety of mobile devices and technical infrastructures together with the demand for faster development cycles makes mobile services prone to technical problems.

From the numerous enhancements and modifications already proposed to TAM, some issues can be identified that could be adapted to study the above-mentioned characteristics that affect user acceptance of mobile services. These issues include value, trust and adoption of the service.

Perceived usefulness included in TAM may not indicate an adequate purchase intention in a market situation. Product value has been proposed as a wider design target both in software engineering and HCI approaches. Value-centred software engineering approach was proposed by Boehm (2003) to define more clearly what the design process is targeting at, and identifying the values that different stakeholders – including end users – expect of the product. Rapid ethnography proposed by Norman (1998) emphasises the importance of identifying big phenomena related to user needs and communicating them early on to the design. Although Norman does not use the term value, his suggestion of concentrating on big phenomena is close to the idea of value-centred design. Cockton (2004b) points out that in value-centred HCI existing HCI research components, design guidance, quality in use and fit to context, need to be reshaped to subordinate them to the delivery of product value to end users and other stakeholders. According to Cockton, value is not just profits and sales but can be political, personal, organisational, cultural, experimental or spiritual.

Mobile services are increasingly handling personal information of the user, for instance due to the personalisation and context-awareness of the services. The functionalities of the increasingly complex systems are not easy for the users to comprehend. Context-aware services may include uncertainty factors that the users should be able to assess. Mobile service networks are getting quite complex and the user may not know with whom (s)he is transacting. Technical infrastructures as well as the rapidly developed services are prone to errors. All these issues raise trust as a user acceptance factor similar to TAM applied in e-commerce (Gefen et al., 2003; Chen et al., 2004). Trust has been proposed as an additional acceptance criterion for mobile services by Kindberg et al. (2004) and Barnes and Huff (2003). Trust has also been included in studies of personalisation in mobile services (Billsus et al., 2002) and studies of context-aware services (Antifakos et al., 2004).

Ease of adoption is included in the studies by Sarker and Wells (2003) and Barnes and Huff (2003). Sarker and Wells (2003) propose a totally new acceptance model that is based on user adoption. Barnes and Huff (2003) cover adoption in their model within the wider themes of compatibility and trialability. *Perceived user resources* in the extension of TAM by Mathieson et al. (2001) and *Facilitating conditions*, in the Unified Theory of Acceptance and Use (Venkatesh et al., 2003) also include elements related to ease of adoption.

The purchasing decision in the consumer market is strongly affected by marketing messages and the media. Gartner's hype cycle is a good reminder of the almost inevitable curve of over-enthusiasm and disappointments with new technologies. It is useful to analyse the position of the technology on the cycle to be able to assess its possibilities and to be able to interpret the messages from media and public debate. Paper I examines agent technology that seems to be in a continuous hype cycle from enthusiasm to disillusionment without significant commercial deployment. The mobile Internet studies described in Papers II–IV took place on both sides of the peak of inflated expectations of the hype cycle. The positive and negative hype certainly affect the end users' perceptions, and may change them rapidly.

In Chapter 6, Section 6.2, I will propose a Technology Acceptance Model for Mobile Services based on the results of the case studies described in Papers I–VI. The model aims at taking into account the aforementioned special characteristics of mobile consumer services, and previous studies on user acceptance described in this chapter. The purpose of the model is to provide a framework to analyse and organise user acceptance factors identified when studying different mobile services. The framework can then be utilised when designing new services and assessing them to ensure that key user acceptance factors are considered in the design.

6. The research framework

This chapter gives an overview of the research framework used in the research projects, the results of which are described in Papers I–VI of this dissertation. This chapter also introduces the Technology Acceptance Model for Mobile Services, which was constructed on the basis of the synthesised results of the case studies. The model is a modification and extension of the original Technology Acceptance Model by Davis (1989) and it identifies four factors that affect user acceptance of mobile services: perceived value to the user, perceived ease of use, trust and perceived ease of adoption. The aim of the model is to provide a framework to design and evaluate mobile services for user acceptance. Later, in Chapter 7, I will present design implications for each of the four user acceptance factors based on the synthesised results of the case studies. This chapter ends with an overview of the case studies described in Papers I–VI.

6.1 The research approach

Human-centred design as defined by ISO 13407:1999 (Chapter 5.1.1) has been the basis of the research projects at VTT, the results of which are described in Papers I–VI of this dissertation. We have integrated the ISO 13407:1999 approach with software engineering practices, as illustrated in Figure 20. When carrying out different usability design and evaluation activities, we are always well aware of which stage of the software development process we are in: context definition, user requirements, design requirements, design or implementation. User evaluation results may suggest returning to earlier phases, e.g. to refine user requirements or intended contexts of use. In practice, returning to earlier phases is not always possible because of the schedule, budget, technical problems or other issues. The project team decides on necessary changes based on user feedback and maintains a design rationale (Moran and Carrol, 1996) to record design decisions and the justifications for them.

We start our human-centred design process by defining preliminary target users, contexts of use and user requirements. The methods used in this phase include literature reviews, user observations, interviews and focus groups as well as studies and user evaluations of earlier or corresponding products. A lot can be learned just by getting acquainted with what has already been studied and what has been learned about user acceptance. Paper I is an example of this kind of literature study. Paper II also focuses on the definition of user requirements. The product evaluations described in Paper V were evaluations of commercially available location-aware products. By evaluating these products with users, we learned a lot for the design of more advanced products in this application field.

In the requirements definition phase we typically evaluate usage scenarios in the form of short stories of possible usage situations (Paper V) and user interface mock-ups. In the design phase we evaluate paper prototypes, simple mock-ups and later prototypes in laboratory conditions (Paper III). The final prototype is typically evaluated in a medium- or long-term field trial to get feedback on the usability and utility of the prototype in everyday use (Papers IV and VI). With location-aware services we also had short user evaluations in the field, where we were using principally the same methods as in the laboratory evaluations (Papers V and VI). The evaluations were carried out in the field because using the system indoors would have been very restricted (e.g. a map application) and because in the field we could use more realistic test tasks.

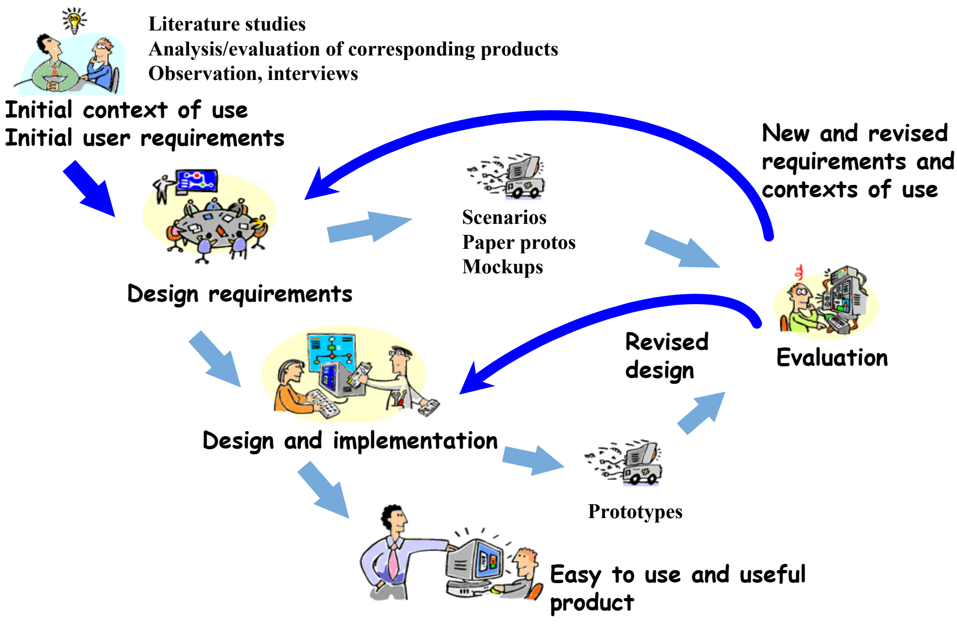


Figure 20. VTT's human-centred design approach.

The principal evaluation methods in the case studies were focus groups (Concejero et al., 2000), use-oriented heuristic evaluation (Po et al., 2004), design walkthroughs (Kaasinen et al., 2000) as well as user testing (Nielsen, 1993), short-term field evaluations (Dix et al., 1997) and long-term field trials

(Kaasinen et al., 2000). Design for all approach (EDeAN, 2005) appears in our work in such a way that we involve different user groups (with regard to age, gender, abilities, skills, background etc.) in the design and target at technical solutions that do not exclude any user group by default.

The case studies described in Papers I–VI were carried out in connection with research projects targeted at prototypes leading to commercial product development. In this kind of environment it is essential to deliver as the results of the project not just the prototype but also feedback on user acceptance of different solutions and the design rationale. User acceptance studies identify the kinds of solutions that were widely accepted by the users and thus worth implementing in the actual products. The studies also identify solutions that were less successful and thus the ways to implement these features or their existence in general should be considered in the forthcoming products.

Especially in the field trials the focus in the research projects was on predicting user acceptance of the proposed technical solutions. In the mobile Internet studies (Papers II–IV) the focus was on usability and utility according to the Technology Acceptance Model by Davis (1989). In the studies of location-aware systems (Papers V and VI) we had an additional focus on ethical issues. Based on the research results, we focused the theme of ethics to user trust, including issues such as privacy protection as well as reliability and controllability of the services. In our field evaluations, we installed our prototype services ready for the users. In addition to our services, the users could freely use any other services available. In user interviews we found that taking commercial services into use was a major problem for the users. Even if they felt that they might have needed some service, the foreseen problems in getting the service installed and set up constituted a major obstacle to starting to use the service. In addition, the users seemed to have quite unclear and often erroneous conceptions about the technologies and services available. In the studies of location-aware systems we included a special research theme of studying how the users could be assisted in taking the services into use.

The technology hype cycles (Chapter 5.1.4) that illustrate the maturity of technologies in relation to their visibility in the media partially defined the social environment in which our case studies took place. During the studies described in Papers II to IV, WAP went through the peak of inflated expectations to the through of disillusionment. Location-aware services have been in the trigger state for quite a long time already, but the time for the plateau is expected to be only 2 to 5 years

(Gartner, 2004). Positive and negative hype certainly affects the attitudes of test users and service providers and this has shown up in the evaluation results.

In this dissertation, user acceptance of mobile services has been studied in connection with the human-centred design approach. The case studies took place within technology research projects well before actual product development, on both sides of the hype cycle defined by Linden and Fenn (2003). Still the focus was on the forthcoming diffusion of the technology to all kinds of users, not just innovators and early adopters. We included in our evaluations users from diverse user groups and eased the adoption of the test prototypes as much as possible to let the users experience how the technology would look and feel at more mature stages.

6.2 Technology Acceptance Model for Mobile Services

Based on the results of the case studies, we have identified key factors that affect user acceptance of mobile services, in particular mobile Internet services and location-aware services targeted for consumers. The first mobile Internet Papers II and III focus on ease of use but also issues related to utility are identified. The later mobile Internet Paper IV gets more focused on utility as the mobile Internet services have been evaluated in long-term field trials. The two last Papers V and VI focus on location-aware mobile services, adding trust and perceived ease of adoption to the factors of user acceptance. Paper I contributes to the model by its trust-related design implications.

The starting point has been the original Technology Acceptance Model by Davis (1989) presented in chapter 5.1.2. The model by Davis introduces two perceived product characteristics that affect the intention to use a product and also the usage behaviour: perceived ease of use and perceived usefulness. Later enhancements to the Davis model, such as TAM2 proposed by Venkatesh and Davis (2000), the model suggested by Mathieson et al. (2001) and the unified UTAUT model by Venkatesh et al. (2003), have identified key forces underlying the judgements of perceived usefulness, e.g. subjective norm, voluntariness, image, user gender and age. The original TAM as well as most of the enhancements and modifications to it are focused on predicting user acceptance, and do not provide much assistance to the design and development of systems that would be accepted by users. The original Technology Acceptance Model was chosen as the starting point for our work because it provided a framework for connecting our field study findings of ease of use and usefulness. Usefulness was not included in our usability evaluation framework (ISO, 1998) where usability was defined as effectiveness, efficiency and satisfaction of specified users in specified contexts of use and carrying out specified tasks. In the ISO definition effectiveness is quite close to usefulness but effectiveness is focused on whether the necessary functions and characteristics for specified user tasks are included in the design. Our research questions for the field trials were wider: we wanted to study how different users were using the mobile services in their everyday lives and what features made the services acceptable in actual usage. Based on the results of both the case studies presented in this dissertation and of previous studies analysed in chapter 5, I propose the Technology Acceptance Model for Mobile Services. The new model extends the original core model by Davis (1989) by identifying two new perceived product characteristics that affect the intention to use: trust and ease of adoption and by redefining the theme of usefulness as value to the user.

The framework suggests that perceived ease of use, perceived value and trust affect the intention to use a mobile service. Figure 21 illustrates the framework and shows modifications and extensions to the original TAM in bold text and as boxes with stronger lines. To get from an intention to use to real usage, the user has to take the service into use. This transition is affected by the perceived ease of adoption. Perceived value, perceived ease of use, trust and perceived ease of adoption need to be studied in order to assess user acceptance of mobile services.

The forces underlying the judgements of these factors, such as the characteristics of individual users or the social environment, are not the focus of this dissertation. Identifying these forces will require further studies and user surveys. The theoretical aim of this dissertation is to define a framework to synthesise the key findings of our field studies and to provide a framework for future user acceptance studies. The framework helps designers of mobile services to identify key issues that should be focused on in the design to ensure user acceptance. Thus the motivation of the model is different than the motivation of the original TAM that was built to explain user acceptance and underlying forces for existing technical solutions. The case studies (Papers I–VI) that the model is based on deal with mobile Internet services and location-aware information services. Entertainment and communication services were in minor roles. For these kinds of services the model can be applied as a starting point, but it may need modifications.

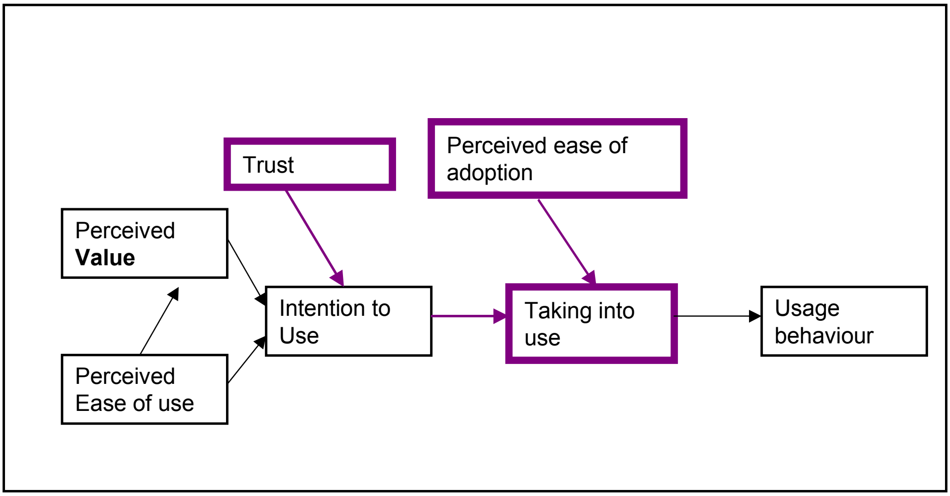


Figure 21. Technology Acceptance Model for Mobile Services as an extension and modification of TAM by Davis (1989).

Perceived ease of use was included in the original TAM and it is also included in the proposed model. Davis (1989) defined perceived ease of use as "*the degree to which a person believes that using a particular system would be free from effort*". At first perceived ease of use is based on external factors such as the user's attitude towards technology in general, experiences of using similar services and information from other people. In actual use and sustained use, perceived ease of use is increasingly affected by the user's own experiences of using the system in different contexts of use.

In the case of mobile services that are used on small devices such as mobile phones or PDAs, the limitations of the device have a major influence on perceived ease of use. The limitations include the small screen, small and limited keyboard, the absence or limited functionality of pointing devices, limited amount of memory, limited battery power and slow connections. As new devices and mobile networks are being introduced to the market, these limitations have somewhat diminished but still mobile networks are slower than fixed ones and the requirements for ease of carrying and holding the device do not allow very large screens or large keyboards. Designing mobile services for ease of use is to a large extent about coping with the limitations of the device. In addition, the design should adapt to the variety of client devices and available networks and other infrastructures.

The ease of use of mobile services has been studied quite a lot and different usability guidelines are available. It is a pleasure to note that many of the usability problems identified in our mobile Internet studies have already been corrected in current mobile devices, browsers and services. Also many issues that we pointed out in the studies are already included in mobile Internet usability guidelines. However, location-aware services pose even more challenges for ease of use. Location-aware services are not just mobile in the sense that they can be easily carried around but, typically, they are used while the user is moving. These kinds of usage situations require extreme ease of use. Personalisation and context-awareness are expected to improve ease of use, but they may also introduce new usability problems, e.g. in the form of personalisation dialogues.

Perceived value replaces perceived usefulness in the proposed model. In the consumer market, perceived usefulness may not indicate adequate motivation to acquire the mobile service. As the focus group studies by Järvenpää et al. (2003) point out, consumers may lack a compelling motivation to adopt new mobile services unless those services create new choices where mobility really matters and manage to affect people's lives positively. In a value-neutral setting each requirement is treated as equally important in the design (Boehm, 2003). This easily leads to featurism – the product becomes a collection of useful features but as a whole it may not provide enough value to the user. Value not only includes rational utility but also defines the key features of the product that are appreciated by the users and other stakeholders, the main reasons why the users are interested in the new product. Values are made explicit by the identification of objectives, which are statements about what the user wants to achieve. Fundamental objectives are directly related to the user's current problem or situation at hand, whereas means objectives help to achieve the fundamental objectives (Nah et al., 2005).

Defining the targeted values and concentrating on them in design and evaluation helps to focus the design on the most essential issues. This is in line with the concept of value-centred software engineering proposed by Boehm (2003) and value-centred HCI proposed by Cockton (2004a,b). Focusing on perceived value in user acceptance studies supports the wider scope of value-centred design, where user value can be studied in parallel with business value and strategic value as proposed by Henderson (2005).

Trust is added as a new element of user acceptance in the proposed model. The original TAM (Davis, 1989) was defined for information systems at work, and in those usage environments the end users could rely on the information and services provided and the ways their personal data was used. When assessing user acceptance of e-commerce applications, Gefen et al. (2003) proposed to enhance TAM with trust in the service provider, as in their studies trust-related issues turned out to have a considerable effect on user acceptance. When consumers are using mobile services that are provided to them via complex mobile service networks, trust in the service providers also becomes an issue. As mobile services collect and use more and more information about the usage environment and the user, ethical issues need more attention, especially ensuring the privacy of the user. Also, as the users get more and more dependent on mobile services, reliability of the technology and conveying information about reliability to the user becomes more important.

In our studies of location-aware services, ethical issues were set as one of the research viewpoints. When analysing the findings of the studies, we identified trust as a wider theme that affected user acceptance of the services. Fogg and Tseng (1999) define trust as an indicator of a positive belief about the perceived reliability of, dependability of, and confidence in a person, object or process. Shneiderman (2000) refers to the definition of trust by Fogg and Tseng and further emphasises that trust is about expectations of the future. Shneiderman defines trust as the positive expectation a person has for another person or an organisation based on past performance and truthful guarantees. In the Technology Acceptance Model for Mobile Services, trust is defined according to Fogg and Tseng (1999). User trust in mobile services includes perceived reliability of the technology and the service provider, reliance on the service in planned usage situations, and the user's confidence that (s)he can keep the service under control and that the service will not misuse his/her personal data.

Perceived ease of adoption is related to taking the services into use. In the original TAM settings with information systems at work, this certainly was not an issue as users typically got their applications ready installed. In our case studies, the interviewees often said that one of the main obstacles in adopting commercial mobile Internet services was that they were not aware of available services, nor did they know how to take the services or even the mobile Internet into use. Furthermore, since the usage need was typically quite occasional, people did not have enough motivation to find out about these issues. And finally, configuration and personalisation seemed to require almost overwhelming efforts. Similar problems with taking services into use have been reported by Kasesniemi et al. (2003). Our studies in the NAVI programme also pointed out that introducing the services to users would require more attention in service design (Kaasinen et al., 2002).

As mobile services are typically used occasionally and some services may be available only locally in certain usage environments, ease of taking the services into use gets even more important. The user should easily get information about available services and should be able to install and start to use the services easily. Finally, (s)he should be able to get rid of unnecessary services. When evaluating prototype services, the users typically get a ready-installed package with personal guidance. This is certainly necessary to get the user trial going, but in practice the users may not get commercial services into use that easily. Firstly, they have to get and understand the marketing messages. Secondly, they have to acquire the service, install and configure it. Studying this activity is gaining more interest in usability research and it is commonly called "out-of-the-box experience" (IBM, 2005).

Compared with the original TAM (Davis, 1989), the Technology Acceptance Model for Mobile Services includes an additional phase between the intention to use and the actual usage behaviour. Our studies pointed out that taking a service into use may constitute a major gap that may hinder the transfer from usage intention to actual usage. Perceived ease of adoption is added to the model at the stage when the user's attention shifts from intention to use to actually taking the service into use.

The characteristics of the user and his/her social environment affect how the user perceives the service. These issues are not included in this study since the focus of this dissertation work was to build the core model: to identify key characteristics of mobile services that generally affect user acceptance of mobile services. Further research is needed to fit previous TAM enhancements such as TAM2 (Venkatesh and Davis, 2000) and UTAUT (Venkatesh et al., 2003) to the model to identify external factors such as characteristics of the users and their social environment that affect the user acceptance factors in the model. In our studies, at least the positive and negative hype turned out to be important external factors affecting user attitudes.

The Technology Acceptance Model for Mobile Services proposed in this dissertation can be used for designing mobile services to make sure that each user acceptance factor gets the necessary attention in the design. In chapter 7 I will further analyse each acceptance factor of the Technology Acceptance Model for Mobile Services and suggest design implications for each factor based on the synthesised results of the case studies. The Technology Acceptance Model for Mobile Services together with the design implications communicate previous user acceptance findings to the design of future mobile services.

6.3 The case studies

Paper I describes the results of a literature analysis typical of the user requirements definition phases of research projects. The study was carried out in a horizontal usability support project (Usability in ACTS, USINACTS) of the European Advanced Communications Technologies and Services research programme (ACTS) in 1999. At that time there was a lot of hype surrounding agent technologies: overoptimistic ideas about the technical possibilities and fears based partly on facts and partly on misunderstandings as well as negative hype. The objective of the study presented in Paper I was to present an impartial overview of what was already known about usability issues in agent-based applications. Moreover, the study identified key challenges for usability issues in agent-based systems so that other projects that were starting to design and implement agent-based systems could utilise these as usability guidelines in their design. Soon after the study described in Paper I – conforming to the hype cycle model by Gartner (Linden and Fenn, 2003) – the hype surrounding agent technology started to dive into the trough of disillusionment. Linden and Fenn (2003) name agent technology as one example of "Phoenix" technologies, which continually cycle through enthusiasm and disillusionment. These technologies are usually extremely slow-moving technologies with principal methodological or scientific challenges, but interesting enough for the media to build up another peak of hype every now and then. In mobile services, agent technology could provide benefits especially for personalisation but there are still considerable technical challenges ahead. Large-scale deployments of mobile agents are not currently expected before around 2012 (Alahuhta et al., 2004).

The case studies presented in Papers II, III and IV deal with technical research projects where human-centred design approach has been adopted in the design of mobile Internet services. The objectives of these projects were to provide as the results both working prototypes and concrete information on user acceptance for future product development. In the early days of mobile Internet we saw that it would be difficult to recognise a single "killer application" for mobile Internet, but the wide access to information and services as such might be the main success factor of mobile Internet. The user interfaces in mobile devices include many restrictions and the mobile networks are slower than fixed ones. That is why Internet services need to be adapted for mobile access. The case studies dealt with usability and utility of Web services that were automatically converted and adapted with a conversion proxy for mobile use and for different devices (Figure 22). In parallel with that, the case studies dealt with mobile-aware services that were specially designed for mobile use. The third category of services that we studied were service entities where a mobile service complements a fixed network service and vice versa. The research focus was on service entities rather than on individual services. From the user's point of view, all three kinds of services were mobile services as they were available on his/her personal mobile device. The automatic adaptation of Web services for mobile use was an efficient technical way to provide the users with versatile content. We identified the characteristics of successful services and predicted usage cultures that could be built around these services.

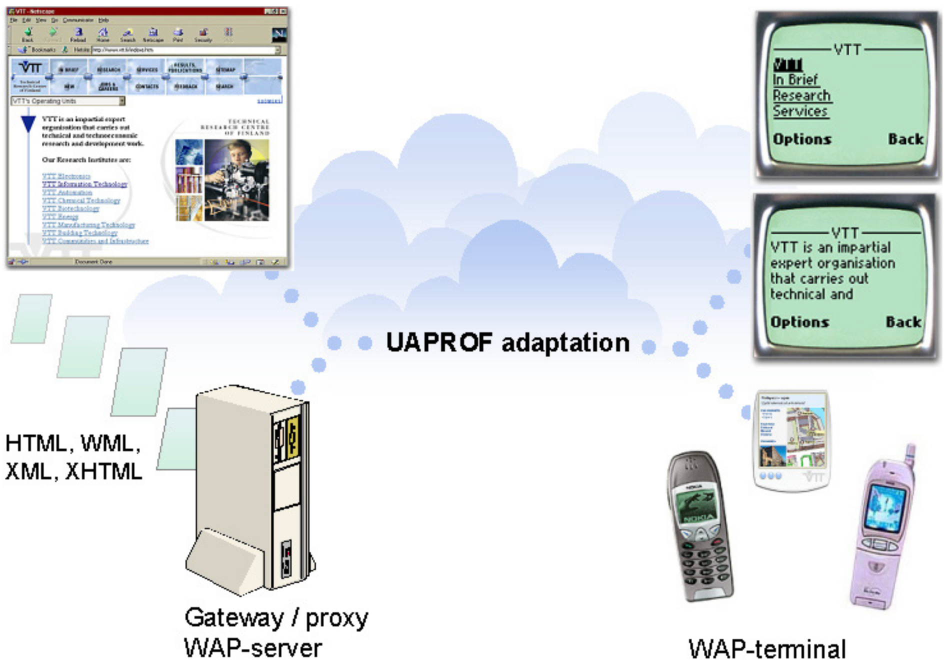


Figure 22. Our mobile Internet solution converted Web content to WML and adapted the presentation according to each individual client device using the User Agent Profile (UAProf) that defines device characteristics and user preferences.

Feedback from the users and service providers guided the design of our mobile Internet services from the very beginning and throughout the design processes. Final user evaluations were carried out with working prototypes and real service content that the users were using in their everyday lives in long-term field trials. We used the Technology Acceptance Model (TAM) as a basis for assessing user acceptance of the mobile Internet services. The evaluations studied both perceived ease of use and perceived utility, thus predicting user acceptance. Especially the qualitative evaluation methods aimed to be open enough to be able to identify unforeseen issues as well. When analysing the evaluation results we identified central user acceptance themes, under which we then grouped the findings.

Paper II describes early design phases: definition of initial user requirements, scenarios and use cases, evaluating them in design walkthrough meetings as well as maintaining the design rationale. Paper III describes the technical framework of the conversion proxy server and the results of first user evaluations with simulated mobile devices in laboratory conditions. We carried out two long-term field trials with different generations of our mobile Internet services. Paper IV describes the results of these field trials and related interviews with service providers.

In the first field trial the test users each had their own tailor-made WAP front page with links to personally interesting mobile aware (WAP) services and converted Web services. The second field trial focused on group communication, and the selection of services was enhanced by a community-specific message board that the members of each group could access both via Web and WAP. In addition the users could personalise their WAP front page themselves, and they had more content available as the conversion proxy had been developed further. The long-term field studies in real contexts of use made it possible to follow the usage patterns beyond the first weeks of "honeymoon" to the real established use, thus getting reliable feedback on the usage routines and user acceptance of the services. The studies were mostly qualitative but also quantitative elements were included – mostly as log files that recorded the actual usage of the services during the test periods. Both field trials lasted for about two months and they both included 40 test users. In parallel with the field trials, service providers were also interviewed in two phases.

Context-awareness gives new possibilities for mobile services. The services can be adapted or even activated according to the context, providing the user with topical and personally meaningful content or services. Location is one of first elements of the context that can be measured and thus utilised in adapting the services. Papers V and VI deal with location-aware mobile services and they take a wider view of analysing user acceptance of mobile services. The research was carried out as a part of the Personal Navigation (NAVI) research and development programme in Finland (Rainio, 2003). The aim of the programme was to facilitate co-operation between different actors who were developing personal navigation products and services. Within a horizontal usability support project, our research group supported individual projects in usability and ethical issues and, beyond this, identified general guidelines for acceptable personal navigation services. Paper V describes empirical studies that we carried out during the first phase of our project. We studied user attitudes and preliminary acceptance by evaluating different service scenarios in focus groups. In addition we evaluated in user and expert evaluations some of the first commercial location-aware services. Paper VI describes field studies of mobile location-aware guidance services. The field studies were carried out in co-operation with the projects of the NAVI programme that were developing those guidance services. Using the Technology Acceptance Model as the general framework, Papers V and VI aim to identify factors that affect the user acceptance of location-aware mobile services by analysing and combining the results of the field studies and other evaluation activities.

In our WAP studies, the point of view of the service providers was studied in interviews, where the possibilities of the services were illustrated as service prototypes. In the studies of location-aware mobile services, service providers participated in the design and evaluation of the services. Looking back, it now seems that the service providers should have had an even stronger role in the projects.

Table 3 gives an overview of all the services studied, the users involved and the evaluation methods employed in each paper.

Table 3. Overview of the services studied, research methods used and users involved in each paper.

Paper	Service, application or device	Research methods	Users
I	Agent-based systems	Literature research	-
II	WAP services WAP-converted Web services	Literature studies Studies of existing mobile applications Design walkthrough	-
III	WAP services	Laboratory evaluation with phone simulator	6
	WAP-converted Web services	Laboratory evaluation with phone simulator	4
IV	WAP services WAP-converted Web services	Field trial 2 months	40
		Interviews with service providers	25
	WAP services WAP-converted Web services Web/WAP Message board for group communication	Field trial 2 months	40
		Interviews with service providers	11
V	Scenarios of personal navigation services	Group interviews	55
	Benefon GPS phone	Field evaluation	6
	Sonera Pointer location-aware WAP services	Laboratory evaluation	5
	Garmin GPS device	Field evaluation	5
	Magellan GPS device	Field evaluation	5
	CeBIT Fair Guide on PDA	Expert evaluation	-
	Pocket Streetmap on PDA	Expert evaluation	-
	Vindigo location-aware service guide on PDA	Expert evaluation	-
VI	Location-aware SMS services	Field evaluation	6
	Weather and road conditions by SMS	Field trial 1 month	10
	Location-aware integrated service directory	Field trial 3 weeks	7
	Mobile topographic maps	Field evaluation	6
	Mobile 3D maps	Laboratory evaluation	6
		Field evaluation	4
	Location-aware tourist information	Web survey	300
		Survey at the ski resort	70
	Scenarios of context-aware consumer services	Interviews in anticipated contexts of use	28

7. Design implications for user acceptance

In the following, I will further analyse the four factors of user acceptance included in the Technology Acceptance Model for Mobile Services presented in chapter 6.2: perceived value, perceived ease of use, trust and perceived ease of adoption. The construction of the model was based on the results of a series of user interviews, evaluations and field studies of mobile Internet and location-aware information services targeted for consumer use. In the following sections 7.1–7.4 I will suggest design implications for each of the four user acceptance factors, based on the synthesised results of the case studies.

Although the case study material is quite extensive, it does not cover all kinds of mobile services. The case studies are mainly focused on information services and entertainment and communication services are touched upon only slightly. Because of the quality of the case study material, the design principles cover best mobile information services targeted at consumers. For other kinds of services, the Technology Acceptance Model for Mobile Services as well as the design implications may need to be revised.

7.1 Perceived value

Values define the key features of the services that are appreciated by the users and other stakeholders, the main reasons why the users are interested in the new services. Defining the targeted values helps in focusing the design on the most essential issues. Although the service may be equipped with several useful features, it is important to identify the key user values of the service and keep focused on them in the design and evaluation. In the following, I will describe the characteristics of mobile services that in our case studies were generally valued by the users. When designing mobile services in the future, these characteristics are worth considering as candidates for service concepts that may provide value to the users.

7.1.1 Successful service content is comprehensive, topical and familiar

In the early days of WAP, service providers often thought that a small device would require a small amount of content. As an extreme, there was even a news service that provided only the news headlines to the mobile user. Our results point out that restricting the information content of individual services is not the right way to assist the users in coping with information overload. The users need access to all relevant information, as deep as they are ready to go, but the information has to be structured in such a way that the user can choose to get the information in small portions. In our evaluations, users appreciated mobile-transparent access to the Web despite its technical problems because it enhanced both the selection of available services and the amount of information within each service, compared with the very limited portals provided at that time by WAP operators.

Especially community-type services will require a critical mass of users to get going. In the mobile Internet trials, communities within which only part of the community had access to the community-based service did not succeed well. It seems that if each member of the group cannot access a mobile community service, a better approach would be to provide mobile access to the existing Web community service for mobile users.

Our studies of user needs for location-aware services highlight the need for comprehensive services in terms of geographic coverage, breadth (number of services included) and depth (enough information in each individual service). A complementary need regarding comprehensive content is easy access to the particular piece of information that the user needs in his/her current context of use. Detailed search options turned out to be important in location-aware service directories. For instance, the user might not want to get just a list of nearby restaurants but the nearest restaurant serving Indian food at a moderate price. Utilising location-awareness can provide the user with topical, location-based information, targeted at users at a certain time at a certain location, not just a fallback to Internet access.

In the evaluations of location-aware directory services (Paper VI) we found that the content of service directories is typically designed for home and office users and they lack some services or whole service categories that would be important for mobile users. Services that the users were missing included gas stations, kiosks, ATMs and bus stops. Commercial directory services did not give access to useful non-profit services and important landmarks. According to our test users, this kind of content would have enriched the services, and would probably have attracted more users.

As alternatives to other services, mobile services have the benefit that they are always available. This can be utilised in providing the users with topical information. This kind of information is likely such that the mobile service is the best way to keep up to date with what is going on. During our first WAP trial, a bus strike happened to be in progress. Our observations showed that quite a few users began to access news services several times a day to check how the negotiations were progressing. Other examples of successful topical content include weather forecasts, traffic information, news topics and event information. Topical travel information, for instance, does not just give timetables but informs about delays and traffic jams as well as recommends alternative routes. A topical roadside help system would not just provide the user with a service catalogue, but would inform the user about the nearest available roadside assistant. A topical event information service would include last-minute changes to the programme and information on free seats.

The user may perceive the service as being familiar because it resembles other mobile services that (s)he has been using or because it resembles the same service or brand in a different environment such as Web, TV or newspaper. In our study with Web conversion for WAP, the users adopted teletext services quite easily. The reason for this was their familiarity with the service in the TV environment. Also the converted Web pages of large television companies were popular because many users were already using these pages regularly on the Web. Familiarity is also an attribute for ease of use as described in chapter 7.2.2, but in multiple media services the value of familiarity is that the user can have his/her favourite service available also on his/her personal mobile phone. Examples of these kinds of services include news, TV schedules, traffic schedules and banking services. Familiarity is also related to the provider of the service. In the mobile Internet trials, the users appreciated more news provided by a familiar newspaper than news provided by unknown service providers. Familiarity is also a trust-related issue as will be described in chapter 7.3.

7.1.2 The service should provide personal and user-generated content

In our WAP trials, personalisation was not just about selecting services but also making the user's own personal items available. These personal items could be the home pages of friends and relatives, the menu of a favourite restaurant, the timetable of the user's bus to work, the user's own notes and similar very personal items. Some users actively used the possibilities to store and access all kinds of personal data, such as memos, to-do-lists, shopping lists and even poems. Access to the stored data from both the Web and the mobile service was especially welcome. In the trials with maps (Paper VI), the users appreciated the possibility to complement the map data with their own information such as important places, favourite routes and self-written notes.

Personal links (Figure 23) were interesting and fun, and often the users could reach information that was not available by other means. The tailor-made personal WAP front pages that we provided the users with in the WAP trial (Paper IV) helped the users in familiarising themselves with the services available and gave them an example of how the personalisation could be utilised. After getting used to the personal front page, they expressed wishes to change the front page and would have been willing to do these changes by themselves. Location-awareness as such also seemed to create an impression of a personal entry to the service (Paper V).

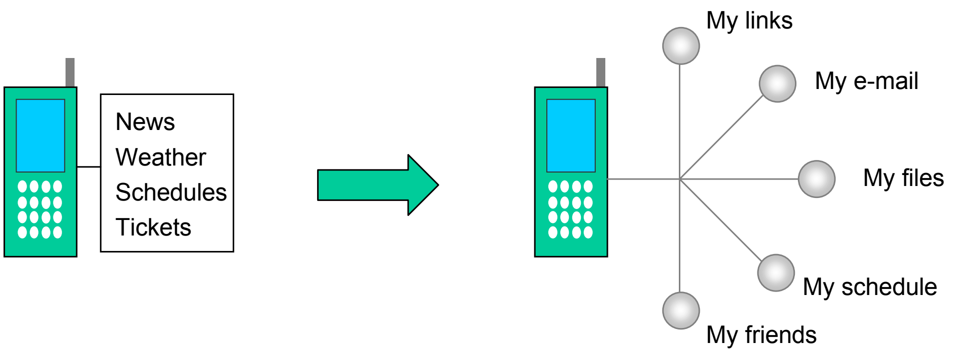


Figure 23. The shift from common to personal increased the appeal of the services.

In our early WAP trials the users faced technical problems with almost all the devices when using long input fields, e.g. when participating in a Web discussion group. Active user input had not been considered much in the design of the first WAP devices and browsers, where users were rather seen as passive information consumers. In fact, at that time usability design guidelines by device and browser manufacturers recommended avoiding or minimising text input by the user.

Despite the technical problems, the test users were quite eager to use the message board service for group communication (Paper IV). According to user interviews of location-aware services (Paper V), letting the mobile users contribute to content creation could enhance many services. Such content may enrich the service, bring in additional users and encourage a sense of community among users. Location-based communities introduce interesting service possibilities. The dynamically changing information generated by other users at a particular location may be more up-to-date and better suited to the needs of the next visitors than a more general type of information that is provided externally. The interactiveness of the users should, however, not be overestimated. The users may not always be willing to respond to contacts from other users or the service (Kolari et al., 2002).

7.1.3 The users appreciate seamless service entities rather than separate services

In our evaluations of location-aware services, some of the evaluated products were designed for route guidance and some for location-aware services. On the one hand, when using route guidance services, the users often missed information about places and nearby services. On the other hand, when using location-aware services, the users expressed a need to get route guidance to the place of interest. They also wanted to have easy access to phone numbers, e.g. the possibility to call the given restaurant to make a reservation. In the first WAP evaluations (Paper III) the users were using a contact database. After finding a contact, they spontaneously suggested that they should be able to call the person or send him/her an email. They also wanted to store the contact in their personal address book. These examples illustrate the needs to continue the usage situation seamlessly from one service or application to another.

The information and the options should be accessible directly from the point in the service where the need for that piece of information or option arises. The design should aim for a seamless solution whereby the user is supported throughout the whole usage situation, e.g. looking for nearby services, getting information on the services, contacting the services, and getting route guidance to find those services.

Another example of the need for seamless entities is the first mobile Internet search engines, which the users were eager to use in the field trials (Paper IV). Unfortunately, many mobile Internet service providers had not defined the title for their pages. This caused search results that mainly included links called

"Untitled". This kind of a service entity was not very appealing to the users, although the search engines as such were initially valued by the users.

In the early days of WAP, service providers believed that they could find a gold mine – a killer application for the mobile Internet. The only really popular application found this far has been e-mail. Our studies suggest that perhaps the main success factor for the mobile Internet will be a wide variety of services including access to already existing Internet services and complementary mobile features as well as mobile-aware and context-aware services that provide the users with contextually relevant information and services.

The selection of mobile services that the users will need regularly seems to be a very personal matter. In our WAP field trial (Paper IV) each of the 10 most active users had 2 to 5 services that they were using regularly. All other available services were used only occasionally. Already the users in our laboratory trials of simulated WAP services (Paper III) anticipated that they would need mobile Internet services only occasionally. Also the case studies with location-aware services (Papers V and VI) indicate that the usage needs for individual services may be quite occasional, even if the users assess the services as being very useful in those occasional situations.

In our field evaluations of mobile Internet services (Paper IV), the users utilised the Web services of popular television shows on their phones while watching television. These kinds of distributed services have further possibilities in the future. The users can participate using their personal mobile devices or they may pick parts of the service to carry with them, for example to ask for continuous updates of a particular news item that they want to follow while on the move. The services should support seamless usage that extends even from one terminal device to another.

7.1.4 The services need to provide utility, communication or fun

Killing time is often mentioned as one of the main contexts of use for the mobile Internet. This motivation for use often came up also in our studies. For example, the users accessed WAP services when taking a break at work, waiting for the bus, waiting for their friends, or travelling by car or on the bus. However, with early mobile Internet services the empty moment to fill had to be long enough because connecting to the service and navigating through the content took time.

Mobile networks are currently much faster but navigating through the service to reach the necessary functionality may still take a considerable amount of time. Another problem with services that are basically filling in empty moments is that the willingness to pay for that kind of a service may not be very high.

In addition to filling empty moments, usage situations in which the mobile Internet was a handier alternative than the office or home PC, newspaper, radio or television were common. These contexts depended strongly on each user and his/her lifestyle and were not necessarily mobile. Indeed, in our WAP trials (Paper IV), the test users were mainly using the mobile services at home, e.g. while watching television or late in the evening before going to bed. The WAP device was readily at hand, and the motivation to use a mobile service was to avoid having to move unnecessarily. The content that the users were looking for included both utility and entertainment.

In our field trials of mobile Internet services we had as test users also novice users who were neither familiar with the Internet nor with mobile phone functions, except for making and receiving calls. For the novice users, in many situations the phone was their only available way to access the Internet. The users who had access to the Internet at work and often at home also could not envisage many situations where the mobile Internet would be useful for them – in most cases they could wait to get home or get to the office to access the Internet. These groups included quite a few business users who in our field trials, turned out to be among the most passive WAP user groups. This indicates that for mobile Internet, the innovators and early adopters may not be technology freaks but people for whom mobile Internet is a good alternative for their occasional need to access Internet services. This is in line with the studies of the adoption of the Japanese i-mode service, where "ordinary consumers" were selected as the target user group.

In addition to personally selected content, interactive services also took the mobile Web to a more personal level, providing the users with new ways of communicating and participating. Typical interactive services that the users were interested in were e-mail, message boards, and storing and retrieving their own notes. A mobile phone is basically a communication device and thus services that enhance or enrich communication have good potential. This proved to be true in our WAP field trials with discussion groups (Paper IV), for whom the mobile alternative to participate turned out to be popular regardless of the technical problems in using the services.

In our studies location-awareness could provide the users with services that were really intended for mobile use, not just secondary access points to Web services. Examples of such services include traffic information, weather forecasts, route guidance, travel information, event information and help services in emergency situations. Location-awareness made these services easier to use and also more personal.

7.1.5 Summary

To conclude, attention should be paid to the following issues when considering the values that mobile services could provide to the users:

- Mobile devices are above all personal communication devices. That is why it is no wonder that key values include personally relevant and interesting content and communication. The communication value can be related to communication-based services such as discussion groups but it may also be related to the possibility to participate and spice up the service with user-generated content.
- Mobile services need to provide the users with topical information. If the information is not topical, the user can have it elsewhere and at other times.
- Mobile services need to provide the user with enough information. If the information is not comprehensive, the user has to get the rest of the information elsewhere, and soon (s)he learns to go elsewhere in the first place.
- Seamless service entities support the user throughout an activity, even from one service to another and from one device to another.
- The usage needs are often occasional, even if the service is very useful in those occasional usage situations.

7.2 Perceived ease of use

Perceived ease of use was defined by Davis (1989) as "*the degree to which a person believes that using a particular system would be free from effort*". Ease of use of mobile services has been studied a lot and the limitations of mobile devices and usage situations are well known, as described in chapter 3. Several usability guidelines are already available. However, as mobile services are getting increasingly complex and enhanced with new characteristics such as personalisation and context-awareness, new usability challenges are raised. Based on the results of the case studies, in the following I will describe the identified design principles that affect the perceived ease of use of mobile services.

7.2.1 Clear overview of the service entity

When starting to use a service or when accessing a service entity, the user will need a clear and intelligible overview of the whole range of available information, services and functions. The first impression may encourage and motivate the user or frighten him/her away. With the first WAP portals that we evaluated, the services were often structured according to artificial classifications. The user easily got lost from the very beginning, e.g. when trying to establish whether the television schedule was utility or entertainment. It is a major challenge to design the front page in such a way that it is simple enough but still gives a clear overview of the whole content of the service.

With the converted Web content as well, a major problem was the complex front pages that tended to be overloaded with images, links, advertisements, animations and so forth. After getting through the front page, navigation was usually much easier and simpler. Some services provided a separate, lighter version of the front page for mobile users. This turned out to be a good solution if the users knew where the lighter version of the front page could be found and if the page did not exclude important content.

The information and the services need to be structured in such a way that the information and functions that the user will most probably need are easily available. By proceeding further, the user should be able to access any information available within the service. If the whole content of the Web service cannot be offered to the mobile user, the user should easily understand what content (s)he can and cannot expect to find in the mobile version. Web usage statistics can give preliminary guidance on which services should be made available. However, user needs in mobile contexts may be different from those in other contexts.

In our mobile Internet studies, user needs for the future included new kinds of visualisations of the service entity. The content could be illustrated e.g. as a figure of a zoomed out page, as suggested by Holmqvist (1999) and demonstrated by MacKay et al. (2004). With today's fast networks, overview transformations could already be implemented in real services.

Occasional usage, which according to our studies seems typical of mobile services, emphasises the need for a clear overview of available services, especially as the available services and their supply may vary according to the location. Occasional usage may indicate that the user has to learn the structure of the service again in each usage session.

When integrating several location-aware services together, we were able to provide the users with comprehensive content (Paper VI). However, the users had difficulties in understanding the coverage and extensiveness of this diverse content. Different services and databases had different types of content and different search methods, and they showed the search results in different ways. When accessing the service, the users could see that they were accessing a directory service but it remained unclear to them what kinds of content were included and from where they originated. Many usability problems faced in using the service could be traced back to the missing overview of the service entity.

To conclude, the following things help the user to get an overview of the service:

- The service should support occasional users – it should be clearly indicated what the service provides and how it should be used.
- An overview of the service content should be provided with information where the content comes from, how often it is updated and how comprehensive it is.
- Enough design efforts and user evaluations should be invested in designing the main structure and the front page of the service.
- Consistency with existing services in other media can help to provide the overview, but then small differences may turn into big usability problems.

7.2.2 Fluent navigation on a small screen

In mobile Internet the basic interaction model is selecting links and navigating from one part of the service to another. Basically, there are two design alternatives: long pages with less navigation between pages but more scrolling, and short pages with more navigation between the pages but less scrolling. In our studies we found that both long and short pages had benefits and drawbacks. Users were generally in favour of longer pages because scrolling was felt to be easier than changing the page. This was especially true in the case of devices with large screens and slow browsers. On the other hand, on devices with small screens and fast browsers, shorter pages eased navigation.

In our user trials, the users were especially reluctant to scroll when looking for something within the site structure. When they had already found what they were looking for, for instance a news article, they could scroll through several screenfuls of information while reading what they found to be of interest.

Ramsay and Nielsen (2000) identified a similar phenomenon in their study. The smaller the screen, the more important it is to arrange information in such a way that the user gets the most important piece of information first. The usability of the sites could be further improved by making the structure adaptive according to each user. A novice user may want to get instructions first, whereas more experienced users may want to go straight into the service. For frequent users, the structure could be adaptive so that the most recently or most often used items are easily available. However, user needs may vary according to the usage situation and then there is a trade-off to be solved between consistency and easy access.

In our WAP trials services familiar from other media such as teletext, newspapers and television schedules were less prone to usability problems and the users were bolder in accessing these services that they had prior knowledge of – even in another presentation format in another media. The appearance of the service was not the most relevant thing as long as the navigation path remained intact. In some instances, being familiar with the service made it possible for the user to use short cuts. For example, when using teletext services on their phone, the users were able to utilise the page numbers that they remembered, e.g. 303 for the television schedule of the MTV3 channel. Also in the early WAP studies (Paper III) the test users expressed needs to utilise both predefined short cuts and user-defined personal short cuts.

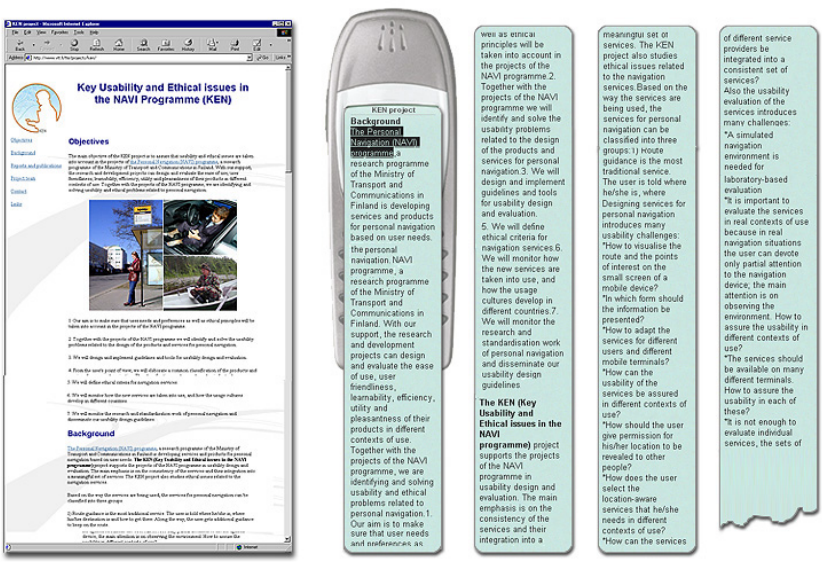


Figure 24. On a small screen, there is a lot to scroll, even when accessing a simple Web page.

Without a consistent site structure, the users were often deterred from navigating because once they had proceeded for a while, they did not know where they were in the navigation hierarchy (Paper III). The users will need descriptive and consistent pairs of navigation links and page titles as well as an easily detectable scroll bar and access back, forward, safe havens and exit within the service.

In the evaluations of WAP services we had quite simple WAP phones (e.g. Nokia 7110) as user devices, but in the evaluations of location-aware services we already had media phones (e.g. Nokia 7650). As expected, the larger screen eased on-screen navigation. This is in line with the findings by Chae and Kim (2004), who found in their studies of mobile Internet usage that as screen sizes are reduced, the horizontal depths of the navigation paths have considerable negative effects on usability. However – typical of mobile Internet services (Kaikkonen and Roto, 2003; Ramsay and Nielsen, 2000) – also on the larger screen the users had problems with backward navigation when they had to navigate several steps backward, for instance when trying to get from 'search results' back to 'make a new search'.

As described in Chapter 7.1.3, the users valued seamless service entities. However, the integrated approach may cause additional on-screen navigation problems if the user navigates to another service without noticing. We identified these kinds of usability problems when evaluating the location-aware, integrated directory service (Paper VI). The users experienced navigation problems especially when trying to get back from a service that they had accessed without noticing. In a similar way, navigation problems may be encountered because the services are increasingly accessed via search engines.

To conclude, the following things support fluent navigation in mobile services:

- The most personally and situationally relevant items should be first on the page.
- A single scrollable page is good for browsing through information; separate pages are better for navigation.
- The users need ways to browse quickly through less interesting information: for instance, an adaptive scroll speed and an illustrative scroll bar are useful.
- The user needs clear feedback on which service and where in it (s)he currently is.
- The service should provide descriptive and consistent link/page header pairs for back, forward, exit, home and other safe heavens within the service.

7.2.3 Smooth user interaction with the service

In addition to navigation, the user may interact with the service using different user interface elements. These elements include text input fields, selection boxes, radio buttons and whole forms. In all our studies text input was a major effort for the users. This was especially true with the location-aware services with which the usage often took place while moving. For instance, inputting street names turned out to be difficult since text prediction was not available; the names were long and often the users were unsure about the spelling. Even if text input could be minimised with ready-made selection lists, inputting a key word may still be less tedious than using complicated navigation paths, as suggested by Kaikkonen and Roto (2003). Our studies indicate that instead of simply minimising text input, a more comprehensive view of effortless user input will be needed. Ready-made selection lists are especially useful when the user is getting acquainted with the service, whereas experienced users may prefer text input. Preferably, both alternatives should be available. Text input should be predicted and interpreted to suggest corrections to possible misspellings. Location-awareness as such could also be utilised in text input, as suggested by Ancona et al. (2001). For instance, nearby streets or the most popular search terms at a certain location could be suggested to the user. In the first place, the user should not be obliged to input information that is not absolutely necessary.

In the WAP evaluations (Paper III) the users had two alternative implementations of forms. The first alternative showed the whole form at a time and let the user select the fields that (s)he wanted to edit. The other alternative was a wizard-type solution showing each field on a separate screen and guiding the user through the whole form. In the first alternative, selecting fields to be edited included additional effort and the users had problems in realising how to select a field for editing and how to get away from the editing field. In the second alternative, the users thought that they had to fill in something in each input field of the form and that also turned out to be tedious. The implementation of forms is device-dependent and the users will probably get accustomed to the way forms are implemented in their device. Still, the intuitiveness of the implementations can be improved.

Many kinds of information are traditionally visualised as tables, e.g. public transport timetables, TV schedules and price lists. Converting tables for mobile access turned out to be difficult (Paper III). Even if the mobile browser supported tables, accessing the table was difficult, as the user did not see the whole table at a time.

To conclude, smooth user interaction with a mobile service was mainly hindered by the user interface restrictions of the devices. More usable ways to illustrate and access forms and tables are needed. Text input is often tedious but it can be eased by providing the users with alternative input methods, text prediction and spell checking. User input should not be avoided in the services but the user should not be obliged to input information that is not absolutely necessary.

7.2.4 Personally relevant services and information without expending effort on personalisation set-up

Current approaches to personalisation are mainly manual and based on a conception of design-oriented users. Unfortunately, the users are not willing to do much regarding personalisation, although they would appreciate the personalised solution, as described in Chapter 3.4. Often the users do not come to think about personalisation until they are already using the service. Based on the results of earlier research (Nielsen, 1998; Manber et al., 2000), we expected that personalisation would have been too big an effort for our users as well. In our first WAP field trial (Paper IV), we tailored personal WAP front pages for each test user based on their expressed interests. We updated the pages during the trial as the users expressed their needs or as we found interesting and convertible sites that might be of interest to them. The users could, for instance, ask us to put onto the service "things to do with my hobby" or "jokes". The users were happy with this approach but they often expressed the need to take care of the personalisation themselves. In the second trial (Paper IV), the users could use a tool, Link Editor, to personalise their WAP front pages themselves either on the Web or in WAP.

In general, the Link Editor tool was simple to use and problems were few. It was clearly beneficial that it was available both on the Web and in WAP. However, only a small proportion of the users did actively personalise their front page. The main problem with personalisation was that the users did not find the time and the motivation to perform this task. On the Web, the personalisation could have been done easily but there was no motivation: the users simply did not think about their mobile phones when they were on the Web. On the other hand, when in WAP, the users were usually too busy with their current task to think about personalisation. In the interviews the users said that they would have preferred a personalisation approach where they could have defined themes that they were interested in rather than exact Web addresses. The users were also willing to be informed about new, potentially interesting services automatically. Our findings indicate that personalisation should be voluntary, and strongly supported in the beginning. This is in line with the findings by Fano (1998) and Espinoza et al. (2001), who point out that the user should have the alternative to use the system in an explorative way: having a look around without any predefined goals. Later on, especially if they start using the service regularly, they may get interested in personalisation options.

In the future, in addition to our manual personalisation, profile-driven and learning personalisation could be effective as proposed by Norros et al. (2003). The tailoring approach that we used in the first user trial could be automated by offering to the users ready-made service packages to choose from. The offerings could be based on the user profile and sent as a push service if the user accepts it. These kinds of automatic offerings would be especially useful in giving the user an overview about locally available services.

User profiles are usually defined separately for each individual service. As the selection of services grows, defining profiles separately for each individual service might get quite tedious for the users. There will be a need to copy the profile from one service to another. On the other hand, transferring profiles and possibly also user identity may cause privacy problems.

To conclude, the following issues may ease personalisation in mobile services:

- Users could be provided with ready-made service package alternatives or guided personalisation services.
- New service offerings could be sent automatically based on user profiles if the user accepts that.
- The user should be able to refine the personalisation with his/her mobile device on the fly.
- New approaches such as profile-driven and learning personalisation may ease personalisation.

7.2.5 Easy access to situationally relevant information and services

In mobile environments, all the elements of the context of use may vary a lot. Users are different and they may use the services for many different tasks, even for tasks that were not anticipated in the design. The variety of mobile devices is growing and the users expect to be able to use the same or the same kind of services on the different devices. The technical and service infrastructure may differ and they may even change in the middle of a usage session, e.g. the network or the positioning system may change when the user moves from one place to another. Similarly, the service infrastructure, i.e. the available services and applications, may change. The physical context may vary a lot in terms of illumination, background noise, temperature and weather. The use of the device may affect the social situation in which the user finds him/herself or the social situation may affect the way the user uses the system.

In our studies of location-aware services we found that in services targeted at a limited area such as travel guides, service catalogues and event guides, the context of use could be predicted quite well according to user location and time. This gives possibilities for different context-aware features in the services, easing their use and giving the users personalised access to them.

In the user interviews regarding location-aware services, most users thought they would not mind having services or information pushed to them as long as the provided service or information was really what they needed in the situation. However, it has to be kept in mind that these users had not tried these kinds of push services in practice. Their opinions may soon change, depending on how successful the first service implementations are.

Discovery services are a new paradigm for service provision to mobile users. The idea is that mobile users and terminals can retrieve and interact with services that are available in their network locality. Services are thus more readily available exactly where they are most relevant. These kinds of locally available services could include tourist guides, event information, shopping guides, etc. The concepts for letting the user explore the services in the environment will require further studies.

Context-awareness can be complemented with personalisation to adapt to user preferences that in different contexts vary from one individual to another. In manual personalisation the user should define these preferences separately for each context. Easier personalisation approaches are needed, as indicated by the fact that the people in our studies felt that the current personalisation dialogues were already too tedious.

7.2.6 Facilitating momentary usage sessions on the move

In our field trials of the mobile Internet services (Paper IV), most usage was immobile use. The users were quite often filling in empty moments by using the services while e.g. waiting for a friend, sitting on the bus, or taking a break at work. With the location-aware services the situation was different. The services were not just mobile in the sense that they could be carried with the user but they were often really used while on the move. The users could devote only part of their attention to using the service while their main attention was on their main task of moving.

Usage sessions with location-aware services were typically non-continuous. The users activated the services and started using them but occasionally they had to put the device aside and do something else. Later on they might return to the service. In the field trials we found that neither the devices nor the services supported this kind of task resumability optimally. For instance, it was not possible to lock the input device in the middle of a usage session. These findings are in line with those of Fithian et al. (2003), who pointed out that the soft keyboard of a PDA is not very suitable for task resumability. Mobile phones are often better in this sense as they have traditionally been designed for on-the-move use, as Väänänen-Vainio-Mattila and Ruuska (2000) point out.

Cheverst et al. (2000) as well as Islam and Fayad (2003) point out the need to be able to use the mobile system both on- and offline. In our evaluations we identified similar needs – the connection may not always be on, but nevertheless the user should be provided with as much functionality as possible. Later offline use can be supported, for example, by letting the users save search results for later reference, and pick up and store Web page content.

Momentary usage sessions emphasise the need for a clear overview of the service, fluent on-screen navigation, smooth user interaction as well as personally and situationally relevant content, which were described in the previous chapters 7.2.1–7.2.5.

7.2.7 Context–aware multimedia access

In the mobile Internet trial (Paper IV) the conversion proxy converted images to a format and size that suited each client device. Our users did not rate images as being important in the services. They thought that the strength of WAP was its clarity. They did not want to sacrifice any of the speed in favour of images. In the WAP trials we did not yet have media phones with large colour screens but the display characteristics of the WAP phones were quite modest. We converted the images according to the mobile device and put the images behind links so that the users could load them only if they felt that, after seeing the textual description of the image, they wanted to view the image itself. As users get used to media phones, their expectations for multimedia content may increase (Kaikkonen and Roto, 2003).

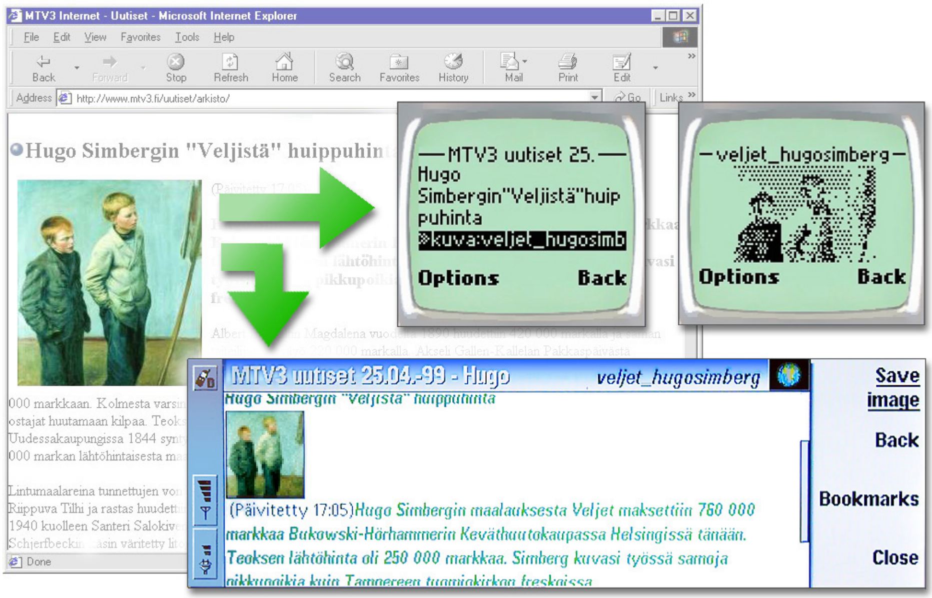


Figure 25. Images were converted and put behind links on WAP phones.

The experiences that we had with images can be utilised with multimedia content in general. Multimedia may be heavy content for future mobile devices as well, meaning that the users should get enough information to decide whether to start downloading the multimedia element or not. Poorer quality and shorter files are adequate when used only to identify the document. In our interviews, the service providers had a lot of expectations especially for audio content in mobile Internet. Mobile devices have traditionally been based on audio and this particular media format has good potential also in the information content for the mobile Internet. Currently, audio-based content may still be a little undermined in mobile Internet services.

Our studies of location-aware services included using mobile maps and route guidance in the form of 3D city models and even video (Paper VI). These user interface elements connected the service with the actual physical environment of the user, allowing the user to orient in and interact with his/her environment. These kinds of location-aware services can be seen as first steps towards future ambient intelligence applications where the user will have enhanced possibilities to interact with his/her physical environment. Overly realistic models of the environment, such as video were not very suitable for orientation. The more detailed the model was, the more correct the users expected it to be in detail. Small flaws in the model could cause misinterpretations.

Brown and Perry (2002) claim that maps on mobile devices could allow a variety of representational capabilities, tuning the spatial representation to the user's needs at a particular moment. As examples, Brown and Perry name overlaying information onto objects in the environment and integrating user-generated information into places. In our studies maps were mostly implemented as representations of traditional paper maps on small screens. Maps had good potential as successful user interface elements, since they allowed connection of functions or information to physical objects in the environment. On the other hand, the two-fold role of being both information and user interface caused usability problems. For instance, when the user was touching the map on the screen (s)he might be showing the map to a friend, asking for additional information of the object on the map, or just wanting to move the map in that direction. Making a difference between these tasks in the user interface turned out to be difficult (Paper VI).

Context-aware adaptation of maps has been studied by Chalmers et al. (2004), who propose a contextual medication approach to meet user needs for map content and presentation based on needs and limitations arising from the context. Our field studies with maps (Paper VI) also indicated needs for personalised and context-aware map presentations, so that the user would see on the map situationally relevant information. In our field trials all map material was in bitmap format. Future devices and maps will provide vector graphic formats, allowing more versatile scaling and adaptation. Different map elements are needed in different usage situations and ideally the user should be able to choose the level of detail and the type of information to be shown on the map.

To conclude, the following issues support context-aware multimedia access in mobile services:

- Multimedia content needs to be adapted according to the client device.
- The user should be able to skip multimedia content.
- The user may have use for light presentations of the multimedia to identify the material.
- Multimedia content such as maps can be used as tools to orient in and interact with the environment.
- Personalised and context-aware map presentations can improve usability by making contextually relevant information easier available.
- The user may need to complement maps and other multimedia content with their own data.

7.2.8 Design for device and network variety

One of the main challenges in designing mobile services is the growing variety of mobile devices, networks and other infrastructures. Preferably, the service should adapt to them like our conversion and adaptation proxy server (Papers III and IV) did. The Design for All approach (EDeAN, 2005) with regard to mobile services also requires taking into account all kinds of devices, not just the most advanced ones. In mobile services, a critical issue is making the service available for users of older or less common devices.

In their study of the first commercial WAP services, Ramsay and Nielsen (2000) point out how failing to take account of the different capabilities of the WAP devices and the different interpretations of WML code on the browsers caused serious usability problems. We faced similar problems, for example, when the Openwave browsers forced the focus to the first input field on the page, or when some services did not include separate Back links that Nokia phones would have required (Paper III). In addition, without device adaptation, it was not possible to utilise the advanced features of devices with larger screens and better resolution. In fact, without device adaptation the WAP services looked least attractive on the most advanced devices.

A good starting point is a simple service, suitable for any device. In this way no devices will be excluded. Then the usability and the attractiveness of the service can be improved by utilising the unique features of each device and browser. In our WAP trials (Paper III), the conversion proxy automatically took account of the capabilities of the client device to adapt the content accordingly and to be able to utilise the device to its full potential.

Style sheets can be used to define how the service should look on different devices. It will still be important to include alternative content for different user interface elements, and meta data describing the content to support the adaptation taking place in client browsers or proxy servers.

In our mobile Internet trials (Paper IV), a number of users utilised extensively both Web and WAP search services. In this way, they were able to find exactly the information that they were interested in without getting lost in the navigation hierarchies. The users were able to find almost anything on the converted Web, but the conversion result was not always good. In WAP they were less likely to find what they needed, but if they did find it, the service was usually in a more usable form. There seems to be needs for adaptive search services that would not only look for particular content, but also take into account the current client device. The search results could be prioritised according to how suitable the content is for the device and network that the user is currently using.

7.2.9 Summary

Ease of use of mobile services has been studied a lot, and some of the design implications presented in this dissertation are already familiar from other studies. However, personalisation, context-awareness and on-the-move use still pose new usability challenges. Key issues regarding ease of use of mobile services include:

- A clear overview of the service entity, fluent navigation on a small screen and smooth user interaction with the service are crucial, and their importance is emphasised with occasional usage needs and on-the-move use.
- Users appreciate personally and contextually relevant information and services but they do not want to expend effort in setting up the personalisation. Context-awareness can be utilised in providing the users with easy access to situationally relevant information and services.
- The services should be designed to facilitate momentary usage sessions on the move.
- The services should support and adapt to the growing variety of devices, networks and technical infrastructures.

7.3 Trust

Trust is an indicator of a positive belief about the perceived reliability of, dependability of, and confidence in a person, object or process (Fogg and Tseng, 1999). When consumers are using mobile services that are provided to them via complex mobile service networks, they may not know the identity of the service providers that they are transacting with. As mobile services increasingly collect and use information about the usage environment and the user him/herself, ethical issues need more attention, especially ensuring the privacy of the user. When users increasingly rely on mobile services in their everyday lives, the reliability of the technology and conveying information about reliability to the users become more important.

User trust in mobile services includes perceived reliability of the technology and the information and functions provided, reliance on the service in planned usage situations, and the user's confidence that (s)he can keep the service under control and that the service will not misuse his/her personal data. Based on the results of the case studies, I will describe in the following the identified design principles that affect user trust in mobile services.

7.3.1 The user should be able to rely on the service in intended contexts of use

Many findings of usability evaluations are related to identified technical problems or implementation errors. These issues are often easy to pass by, reasoning that the system evaluated was a prototype and that these problems will be corrected in the "final version" of the system. Severe technical problems may affect the user evaluation results, as the users may well stop using a system that does not work as expected.

Connection problems were common in the mobile Internet trial with the first WAP-enabled mobile phones (Paper IV). The users sometimes could not access the services at all or the connection was cut in the middle of a usage session. The conversion proxy did not succeed in converting all the Web content to WAP, and could not always inform the users about that. Text input included many problems that made the system lose user input. Negative usage experiences caused by the malfunctions lowered the usage rate, but did not make the users totally stop using the services.

In one of the field studies with location-aware services (Paper VI) different services and databases were integrated into a comprehensive directory service. This complex service architecture caused problems in error situations because it was often not possible to analyse and indicate to the user the reason for the error. The users were left quite helpless since they did not know whether the problems were in the mobile device, in the network or in some of the integrated services. At least, they would have liked to know whether the problem was persistent or just occasional. At times the service did not work in an occasional situation where it would have been very useful to the user, e.g. when searching for nearby hotels after a car breakdown in the middle of the night. These kinds of unsuccessful usage experiences discouraged the users from using the service again.

Although we logged usage data in the field trials, it was not technically possible to log error situations. Thus we could not assess the error rates accurately. Practically the only source of information was the observations and interviews of the test users and the experiences of the evaluators themselves.

The study described in Paper I points out that as the technical solutions get more complex, also the error situations may get more complicated and more difficult for the users to understand. The errors may be user errors, errors of the technology or unexpected events in the environment. Ideally, in error situations, the user should be told where the error is and whether and how it can be overcome. However, within a complex service entity, it may prove difficult to give clues about the reason for the malfunction or guidance as to what the user is supposed to do. The analysis of possible mistakes (by the system or by the user) and their effect on the service should be an essential part of the design. The design should include action plans for both the user and the service in each identified error situation. In the event of the user losing the connection to the service, it should be assured that no harm will be done.

Additional devices such as GPS increase the power consumption of mobile devices. Power consumption also increases considerably as the temperature decreases. The users in the field evaluations were often missing feedback on the power still available and estimates of the sufficiency of batteries with different combinations of add-on devices and functions (Kaasinen, 2003). A user on the move may need to make decisions regarding which combination of functions (s)he can afford to keep on in order to avoid exhausting the battery power totally.

When mobile services get more common, people will tend to get more and more dependent on the technology. The users should be aware of the risks in using the product and they should be provided with information about the reliability of the service so that they can assess whether they can rely on the service in the planned usage situations.

Preventing error situations and giving the user enough support if the error occurs anyway are important issues, because repeated malfunctions that the user cannot understand or solve are a major source of bad usage experiences. User errors should be prevented by all means, e.g. by trying to interpret, correct or complete user input.

Error situations and how to log and analyse them are not a very common research theme. More efforts should be made in designing how to identify, log and analyse error situations during evaluations. The design should set target values for tolerable error rates. Errors cannot always be prevented but the design should have strategies for error situations: how to recognise or forecast error situations, how to overcome these situations and how to keep the user informed so that (s)he knows what is happening and what (s)he should do.

7.3.2 Measurement without estimated accuracy is of no use

When evaluating location-aware guides with cell-based positioning, the location was indicated to the user as numeric co-ordinates and/or a symbol on a map. The users would have needed information on the accuracy of the location as well, and the location accuracy should have been taken into account when showing the map. When using cell-based positioning, the actual position of the user could be even hundreds of meters away from the position suggested by the system. As it turned out, sometimes the inaccuracy caused situations where the cell location was shown on the map but the actual position of the user was outside the map area shown on the small display. In systems without a map, the user had no possibility to check which position the information on nearest services was based on. Because many users had very indefinite conceptions of the order of magnitude of the available positioning accuracy (Paper VI), they often interpreted the consequences of inaccurate positioning as "the system is not working".

Regarding location and orientation data, misunderstandings took place as the users did not get clear feedback on the timeliness of the data provided.

Sometimes the user did not notice that (s)he did not have the GPS switched on and sometimes the last position fix had taken place a while ago since the user was in an urban canyon (Paper V). Orientation information was often difficult to interpret since the GPS devices in our evaluations could measure orientation only when the user was moving.

Regarding the positioning systems, some kind of a conception of the principles of the positioning technology would have helped the users in coping with limited positioning accuracy and temporary malfunctions. In our field trials we noticed that simplified descriptions helped a lot in understanding the basics of the positioning systems. For instance, we could describe cell-based positioning as a system that "locates the base station that can be best heard from here". The GPS system could be described as a system that "needs to see at least three satellites in the sky". It seems that understanding these basic principles also made the users more understanding in error situations, mainly because they could find a rational explanation for the malfunction and therefore did not feel so helpless.

When accessing location-based information, the users were often missing quality attributes for the data provided, such as the freshness of the data, the provider of the data, and the data entity from where the data had been retrieved (Paper V). These attributes would have helped them in assessing how to regard the information.

As mobile services are using and providing the user with increasing amounts of different measurement data, the accuracy of the data needs to be analysed in the design. Does the user have the right conception of the measurement accuracy? Is the accuracy sufficient for the kinds of tasks for which the user will be using the service? Does the user get feedback on the freshness of the data and its accuracy, especially if these vary according to the usage situation?

7.3.3 The privacy of the user must be protected even if the user would not require it

As information technology gets embedded almost everywhere in our environment, information about us and our behaviour may be collected, integrated and used in different ways. Privacy protection in mobile services is related to the right to collect personal data, connect the data with user identity, use the data, store the data and forward the data. In our studies privacy issues were raised in connection with locating people (Paper V). The legislation in most countries requires the user's permission before (s)he can be located. This is the basis for privacy protection, but social regulation can also create rules and norms for different situations in which location-aware services are used (Ackerman et al. 2001). As Pitkänen et al. (2002) point out, too strict privacy protection by legislation may lead to unintended results as some useful services are not developed. In practice, trade-offs between privacy protection and effortless use need to be resolved. The more seamless and easy the use of a location-aware application becomes, the fewer the cues to remind users that their location is being monitored (Barkhuus and Dey 2003). The user should be reminded that (s)he is being located but this should not require excessive user effort, e.g. in the form of acceptance requests.

In our studies with location-aware services, getting information about other people's location was often considered intrusive (Paper VI), similar to previous studies (Cheverst et al. 2000, Fithian et al. 2003). It seems that in many cases it is better to convey location information indirectly, in the form of a user-defined description of the current context (e.g. downtown, at work, shopping or partying) rather than exact co-ordinates.

In our interviews (Paper V), people were often worried about their privacy and the "big brother" phenomenon when considering services enabling people to be located. However, the interviewees were not worried about privacy issues with location-aware services. It did not occur to most of the interviewees that they could be located while using the services. If they did understand this, they seemed to put a great deal of trust in the current telecom operators. It was also commented that there would be regulations and legislation protecting those who use location-aware services.

The users recognised situations where they would like to convey their location to service providers. For instance, when evaluating the location-aware directory service with roadside assistance volunteers and professionals, the users pointed out that with a roadside assistance service it would be useful if the driver in trouble could easily forward his/her location to the roadside assistant. Allowing the user to control his/her own location data and e.g. forward this data might be useful with other services as well. The users will need ways to manage their location data themselves and to decide to whom they want to transfer the data either once or continuously. One obstacle here turned out to be the possession of the location data. If the telecom operator has generated the location data, the operator possesses it and does not have to give the data to the user or a third party service provider, at least not for free. In our evaluations many service models were rejected by the service providers because they expected that getting location data from the telecom operator would be too expensive.

Even though in our interviews the users were not worried about possible misuse of their location data (Paper V), it should not be interpreted so that privacy protection would not be required for user acceptance. The attitudes of our test users were based on their trust in their current telecom operators, and they assumed that also in the future their familiar operators would provide the mobile services. User location is very personal information and it should be protected even when the user him/herself would be trusting enough not to require it. The user should be provided with easy mechanisms for giving permission to use the data for a predefined purpose. The user should get clear feedback on whether the positioning is on or off. Location histories should not be stored purposelessly and without user consent. When location data is conveyed to others, it is worth considering whether they will need the exact location co-ordinates or more descriptive but less intrusive description defined by the person to be located. It should also be considered whether it is necessary to connect the location to the user identity.

Paper I suggests the following privacy protection principles regarding the use of user profiles:

- The user should be allowed to inspect and alter his/her user model in the system.
- The user must be aware of where the user model is stored and how it is being used.
- The system must have rules governing what information it can and cannot communicate to third parties.
- The user should have easy access to the information, which has been filtered on the basis of the user model.
- Especially learning systems may draw conclusions that cannot be guaranteed to be right. The user will benefit if he/she can be provided with an estimation of the probability of the results.

Legislation often tends to lag behind technical developments. When studying and developing location-aware services, many questions concerning future legislation were left open. Finally, in September 2004 the legislation (Finnish legislation, 2004) concerning privacy protection in electronic communication (516/2004) became effective in Finland as a national interpretation of European directive 2002/58 (European Union, 2002). The legislation includes separate sections for location data generated by a telecom operator. The legislation aims at achieving a balance between user control over his/her personal data and effortless introduction of new services for the service providers. The telecom operator can collect and handle location data if the subscriber has not forbidden that. The subscriber should have the possibility to forbid the collection and handling of location data easily and any time. The subscriber should have access to information regarding the accuracy of the location data, the purposes the location data will be used for, and the third parties to whom the information will be redirected. The telecom operator can give location data to third parties if the telecom operator has evidence that the third party has the necessary authorisation from the subscriber. Third party service providers do not need to ask for permission for location data from the subscriber if it is evident that the service in question includes using location data. This section is quite challenging, because as our field studies pointed out, the users may not think about themselves being located if they use location-aware services. Thus the service providers should be careful in deciding what is "self-evident".

In future services, it can be expected that in addition to user location, a lot of other personal data may be collected. This may include health-related measurements, shopping behaviour; services used and so on. The same principles as with location are to great extent valid also with this data.

7.3.4 The user needs to feel and really be in control

The more complicated the mobile services and the service networks behind them get, the less possibilities the user has to understand what is happening in the service. Basically, automated functions are intended to minimise the user's need to intervene. However, to be able to intervene when necessary, the user needs to know the reasoning behind the behaviour of the system. In our studies users often could not distinguish which functions and data took place on their mobile device and which on external servers. For instance, when using cell-based positioning, the users were saying that "the phone knows where I am". The services need to be somewhat seamless to ensure effortless use but on the other hand, some issues need to be clearly differentiated so as to ensure that the user understands what is going on. Seamless services may hide details from users when aiming to provide ease of use. This may hinder the user from understanding what is happening "behind the scenes" (Höök, 2004).

User control was studied in Paper I. The rules for user/agent teamwork can be applied also to other automated systems. The main rule is that the user has to feel and really be in control. To be able to be in control, the user needs to know the system's capabilities and rules of reasoning. The user also has to have a clear overall picture of the system. He/she has to understand what is going on and why, even if it is unnecessary to understand all the details. The user should be able to control the degree of automation and intrusiveness.

Barkhuus and Dey (2003) claim that the more useful the user considers the system, the more (s)he may be willing to give away control. That may be true but giving away user control should be voluntary and definitely not “the price of usefulness”. In the user interviews (Paper V) a common fear by the users was a predestined or over-controlled environment that aims at leading the user, and seems to know better than the user him/herself what the user needs or should do. The user should feel and be in control, and be able to override the recommendations of the system, as suggested by Cheverst et al. (2000).

Push services are a good example of seamless behaviour. Location-aware push services provide the users with information automatically, activated by the location of the user. Users may accept push services because of the effortless use (Cheverst et al. 2002). In the field trials of the mobile Internet (Paper IV), the users indicated that there are situations where they would like to be alerted of new or updated content. For instance, they would have liked to be alerted of new messages on certain message boards or of event information. However, the users expressed a number of reservations about commercial push services. It is a totally different thing to receive a push message that you have been expecting than to receive unexpected messages. The studies by Teo and Pok (2003) in a survey with 1000 users of mobile Internet in Singapore present somewhat different results. Their study revealed that 60% of the respondents would not reject mobile advertisements pushed to them, provided that the advertisements were personalised and that accessing the advertisements would subsidise their phone bill.

The attitudes of the users were quite positive towards push services in our field studies of location-aware services as well (Paper VI). The users expected that push services would provide them with personally and contextually relevant information. In practice the user should get clear feedback on the push services: which services (s)he is registered to and how probable the messages are. The user should be able to easily define which topics (s)he is sufficiently interested in to be willing to receive push messages. The user should be able to fine-tune or cancel the push feature easily, ideally as (s)he receives a push message.

7.3.5 Summary

When mobile services get increasingly involved in the personal life of people and as these services increasingly collect, analyse and store personal data, user trust becomes an important acceptability criterion. The following issues should be considered in the design:

- It should be ensured that the user feels and really is in control. This requires that the user has a clear conception of the functionality of the service even if (s)he does not need to know all the details. Push features should be introduced with care, and the user should be able to fine-tune or cancel the push feature easily.
- The user needs information on the reliability and accuracy of the service to be able to assess in which usage contexts (s)he can rely the service.
- The design should include strategies for preventing, predicting and identifying error situations, informing the user about them, and providing him/her with guidance on how to overcome the problem. Tolerable error rates should de defined as design targets, and error rates should be studied in user evaluations.
- The privacy of the user should be protected even if the user would not explicitly require it. The user should get clear feedback on which personal data is collected, where it is stored and who is using it and how.

7.4 Perceived ease of adoption

Perceived ease of adoption is related to taking the services into use. As mobile services will increasingly be available from different sources and in complex service networks, it becomes important to ensure that the users get reliable information about available services and get the necessary guidance when taking the services into use. It would be a pity if a good mobile service was left unused only because of problems in the marketing messages and the out-of-the-box experience. Taking the services into use is often dismissed in user evaluations, as installation support has not yet been designed into the prototype services being tested. This also happened in our WAP case studies. However, in user interviews we repeatedly heard users describing their difficulties in taking commercial mobile services into use. The problem with introducing new services into use was recognised in the NAVI programme, and in our project we had a separate activity aimed at defining and evaluating a Trade Description Model for personal navigation services. The results of those studies are published in separate publications (Kaasinen et al., 2002; Ikonen et al., 2002a). The results presented in section 7.4.1 are based on those publications, whereas the results in sections 7.4.2 and 7.4.3 are based on the case studies presented in Papers II–VI. In this chapter I will describe the identified design principles that may help users in adopting new mobile services.

7.4.1 Real values of the services need to be emphasised in marketing

Studies on mobile Internet and mobile multimedia messaging have identified several issues where the users had missing or wrong information on the new service (Kolari et al., 2002; Kasesniemi et al., 2003). Many users had a quite blurred understanding of the features that were available on their own mobile phones. Some users had a misconception of the services behind acronyms (e.g. MMS). The possibilities and limitations of different network technologies were unclear to many users (e.g. Bluetooth). Not all of the users understood which features and services were readily available on their phone and which features they had to install separately. Neither was it clear to them where to find the instructions for installation. And finally, the users did not always understand the installation instructions. All this lack of information constitutes quite an obstacle for taking mobile services into use. (Kolari et al, 2002; Kasesniemi et al., 2003)

In the first market failure with WAP, a wrong marketing message was a major reason. The services were marketed as "net in the pocket", giving birth to overoptimistic expectations among the users. Had the first services been marketed more modestly as an easier way to use SMS services, the acceptance might have been different. However, it has to be kept in mind that also high pricing, technical immaturity and unreliability contributed to the failure. Still, the problems with wrong marketing messages seem to be repeated with other technologies, and the introduction of new technologies seems to repeatedly follow the hype circle by Gartner (Linden and Fenn, 2003). The first introduction of a new technology is often overoptimistic and emphasises the future possibilities instead of the features actually available. The second-wave introduction usually tones down the promises remarkably, focusing on concrete and existing features.

The user should get reliable information on new products to be able to assess whether they would be suitable for him/her and his/her needs. In addition to technical specifications and lists of features the users would need information on usage possibilities as concrete descriptions of how the product would look and feel in their everyday life. As a part of our studies on personal navigation services, we constructed a trade description model that helps in the introduction of personal navigation services or products to the users (Kaasinen et al., 2002). The model is based on questions that could be asked by a buyer or a first-time user of a product. Table 4 illustrates the main questions included in the model.

Table 4. Trade description model for personal navigation products and services (Kaasinen et al., 2002).

Classification	*Trade description*
User	**Is this product/service suitable for me?** • Targeted specially at a certain user group • Targeted only at a certain group • Accessibility for disabled users
User goal	**What can I do with this product / service?** • Locate myself • Be located by other people • Locate other people • Track my property • Get route guidance • Find and use nearby services • Get help in emergency situations • Have fun
Environment	**Where can/cannot I use this product/service?**
Equipment	**What do I need to know about the technology?** • What kind of technology do I need to be able to use the service? • How compatible is this product/service with other products/services? • How accurate is the positioning? • To what extent can I rely on this product?
Service characteristics	**What specific features does this service include, what is the added value of this product compared with competing products or current ways to act?**

On the one hand, the trade description model helps consumers to compare different products and, on the other hand, it helps service providers to describe their products in a consistent way. It also turned out that the model helped in defining common terms for new concepts. Additionally, the trade description model could be used as a checklist of issues to be covered when writing "Getting started" manuals. Although the model was designed for personal navigation services, it is general enough to be adopted for the description of other mobile services as well.

We briefly evaluated the model with both designers and end users of personal navigation technology (Kaasinen et al., 2002). The trade description model seems to work as such but it is essential that the answers to the proposed questions are targeted according to the end-user group. Furthermore, the answers as such give quite a fragmented impression of the product. Most answers could be integrated inside a story that describes the product in real use and in situations familiar to the target end-user group (Ikonen et al., 2002a). In this way, the description is easier to piece together and the user can better assess if (s)he could utilise the product in his/her everyday life.

7.4.2 Disposable services for occasional needs

Consumer products are targeted at a wide market. Written instructions are often out of the question because people expect that consumer products can be taken into use without them. The initial experience a user has in taking a new product out of the box and setting it up, in preparation for use, creates a lasting impression and constitutes an important aspect of the total user experience. IBM has issued guidelines on how to design out-of-box experiences that are productive and satisfying for users. They also provide suggestions for effective evaluation and testing of the out-of-box experience. (IBM, 2005)

When WAP phones were first released onto the market, the users had to make the installations manually. This required up to twenty or more steps, which most likely prevented most users from taking the services into use. Nokia originally developed a technique by which configuration data could be sent to the phone as a smart message (over the air, OTA). When receiving the message, all the user has to do is to reply "yes" to the question whether the configuration included in the message should be taken into use. Other phone manufacturers have adopted this technique as well and it clearly eases the installations. However, the user still has to find out which message to send and where, and this still seems to be an obstacle to taking services into use.

In the evaluations of location-aware guides, many users who already owned WAP-enabled phones had not installed the mobile Internet on it (Kaasinen, 2003). A common reason for this was that the users did not need mobile Internet enough to bother finding out how to perform the installation. The users wondered why the settings could not have been made to their phone at the store where they bought the handset. Certainly, this kind of support would have encouraged them to try out the new services.

If the user's motivation to take the service into use is only moderate, (s)he may not be ready to put much effort into installation. Ideally, the services should be installed on the user device at the point of sale, and the user should at the same time get personal usage guidance, but presumably this will be possible with only few services. One of the factors affecting the success of i-mode was the ready-made installations, where the users got the phone and personally selected services all in the same ready-to-use package (Funk, 2004).

According to the results of our field studies (Papers IV, V and VI), users are not willing to spend their time on something that they do not get immediate benefit from. In general, the users preferred to utilise the services without any pre-made plans. The users could not take WAP into use spontaneously as they had to install the system using a dedicated Web service on a PC. The users of maps often needed to load the maps onto the mobile device using a PC. If they forgot to do this or did not have time for it, the mobile service was of no use on the move.

In the user evaluations of location-aware services (Papers V and VI), most users thought that they would use location-aware services occasionally and mainly in unfamiliar environments or in emergency situations. These needs indicate that the services should be easily available when the spontaneous need for them arises. If the service requires personalisation, the motivation for this in the case of occasionally used services seems quite low. Taking these kinds of services into use could be eased by letting the users have the same profile in several services, e.g. in different city guides.

In the future, context-aware services will pose additional challenges for taking new services into use. The services may be available only locally or in certain contexts. The user should be able to identify, understand and take into use these services easily while on the move. As the selection of available services grows, it will also become increasingly important to get rid of unnecessary services easily.

7.4.3 The service has to support existing and evolving usage cultures

The products shape the usage but also the usage should shape the products. The technology should support the natural usage patterns of the users. To succeed in this, existing and evolving usage cultures should be studied in parallel with the technology development. The design should fit in with the social, technical and environmental contexts of use, and it should support existing usage cultures. In addition the users should be provided with possibilities to create their own, new usage cultures. Ideally, the technology should provide the users with possibilities that they can utilise in their own way, rather than forcing certain usage models fixed in the design. (Norros et al., 2003)

In the field trials of location-aware services (Paper VI) we installed both WAP and the service to be tested on the user's phone. We gave the users a step-by-step printed guide that described the basic ways of using the service as well as typical error situations and how to overcome those errors. Installation support was clearly needed but during the evaluations we found that the step-by-step guide also limited the usage. The users preferred to use the service exactly as was described in the guide and they were shy of trying alternative ways of utilising it. For instance, the guide described first how to ask the system to locate the user and then search for nearby services. Another alternative would have been to input the location as a street address. This alternative was described later in the manual but it turned out that most users stuck to the first alternative. Even if the users would have needed information of services at their destination, many of them did not come to think that they could have got that information by inputting the destination manually.

Although the users will benefit from clear usage guidance, they should also be encouraged to find out and innovate their own ways to utilise new products. Our user studies pointed out that the users could be supported in innovating their own ways to use a product, as they were given example descriptions of other people using the system even in different application fields (Ikonen et al., 2002a).

7.4.4 Summary

To conclude, ease of adoption of the mobile service requires the following:

- The user should be given a clear description of the possibilities that the service would facilitate in his/her everyday life.
- The installation should be effortless, even on the move.
- Services should increasingly be provided as disposable packages: easy to find, take into use, use and get rid of when no longer needed.
- The services should fit into existing usage cultures but they should also provide a basis for user innovations and new usage cultures.

7.5 Summary of design implications

The Technology Acceptance Model for Mobile Services was presented in Chapter 6.2. The model defines four main user acceptance factors: perceived value, perceived ease of use, trust and perceived ease of adoption. How these factors should be taken into account in the design of individual mobile services depends on the service in question. However, according to the case studies presented in Papers I–VI, which focused on mobile Internet and location-aware information services, there are many attributes of the acceptance factors that repeat from one service to another. These attributes form a set of design implications that can be used in the design of mobile services. The design implications can additionally be used in designing user acceptance evaluations to define the issues to be studied in the evaluation. In the following, I sum up the user acceptance factors and related design implications.

Perceived value of mobile services is not based on featurism – having more and more features in the service. Instead, the key values that the service is supposed to bring to the end users, service providers or other stakeholders should be identified. The following list gives some ideas where the value can be found:

- Mobile devices are above all personal communication devices. Key values include personally relevant and interesting content and communication. The communication value can be related to communication-based services such as discussion groups but it may also be related to the possibility to spice up the service with user-generated content.
- Mobile services need to provide the users with topical information. If the information is not topical, the user can have it elsewhere and at other times.

- Mobile services need to provide the user with enough information. If the information is not comprehensive, the user has to get the rest of the information elsewhere, and soon (s)he learns to go elsewhere in the first place.
- Seamless service entities support the user throughout an activity, even from one service to another and from one device to another.
- The usage needs are often occasional, even if the service would be very useful in those occasional usage situations.

Perceived ease of use of mobile services has been studied a lot, and some of the ease of use design implications presented in this dissertation are already familiar from other studies. However, personalisation, context-awareness and on-the-move use still pose new usability challenges. Key issues regarding ease of use of mobile services include:

- A clear overview of the service entity, fluent navigation on a small screen and smooth user interaction with the service are crucial, and their importance is emphasised with occasional usage needs and on-the-move use.
- Users appreciate personally and contextually relevant information and services but they do not want to expend effort on setting up the personalisation. Context-awareness can be utilised in providing the users with easy access to situationally relevant information and services.
- The services should be designed to facilitate momentary usage sessions on the move.
- The services should support and adapt to the growing variety of devices, networks and technical infrastructures.

Trust becomes increasingly important as mobile services get more and more involved in the personal life of people and as these services increasingly collect, analyse and store personal data:

- It should be ensured that the user feels and really is in control. This requires that the user has a clear conception of the functionality of the service even if (s)he does not need to know all the details.
- The user needs information on the reliability and accuracy of the service to be able to assess in which usage contexts (s)he can rely the service.
- The design should include strategies for preventing, predicting and identifying error situations, informing the user about them and providing him/her guidance on how to overcome the problems.
- The privacy of the user should be protected even if the user would not require it. The user should get clear feedback on which personal data is collected, where it is stored and who is using it and for what purposes.

Perceived ease of adoption becomes increasingly important as mobile services may be available only locally or in certain contexts. The user should be able to easily identify, understand and take into use mobile services even while on the move:

- The user should be given a clear description of the possibilities that the service would facilitate in his/her everyday life.
- The installation should be effortless, even on the move. Services should increasingly be provided as disposable packages: easy to find, take into use, use and get rid of when no longer needed.
- The services should fit into existing usage cultures but they should also provide a basis for user innovations and new usage cultures.

8. Service providers' point of view

To reach the market, mobile services should be acceptable to the end users and a profitable business for the organisations involved in the business network required implementing and maintaining the service. Key issues are the effort required to implement and maintain the service vis-à-vis the expected number of users, usage frequency and willingness of the users to pay for the service.

When new mobile services are being developed, device manufacturers, application developers and network operators are obvious participants. The companies, organisations and individuals that will actually provide the content for the services should also have roles in the development process. Service and content providers are the users of the technology being developed: they use the technical solutions to better serve their clients. Typically, information technology is not their special field of expertise, so they will need reliable guidance – not just marketing hype – to be able to consider how to utilise the possibilities of mobile services in their business. In the case studies we included content and service providers in our human-centred design process. We interviewed them and evaluated the proposed solutions with them to ensure that the mobile services would be acceptable also from their point of view.

8.1 The hype and flaw of mobile Internet

The development of mobile Internet in Europe has been dominated by technology push. Technology has been pushed both to end users and service providers. Both actors have been quite unaware of the possibilities of the new technology. For instance, WAP was introduced as "Web in the pocket", giving an overly optimistic view of the possibilities of this technology.

In the case study of mobile Internet services (Paper IV), we interviewed content and service providers in two phases. The first interviews were carried out in spring 2000 with 25 interviewees, and the second one in spring 2001 with 11 interviewees. The interviewees represented organisations that were seen as potential content and service providers for mobile Internet services, and they were responsible for their organisations' Web and mobile service production. The organisations in the interviews included small-, medium- and large-size companies as well as public-sector organisations and non-profit organisations. All the organisations were providing services for consumers.

In the year 2000, half of the organisations were providing some kind of mobile services, typically light Web services for mobile laptop users and SMS-based services for mobile phone users. Some organisations were already providing their services also on the WAP portals of different operators. Mobile hype was still strongly in the air and the service providers thought that mobile Internet would soon be a significant business area that they needed to be involved in. The interviewees thought that within a couple of years their organisations would have strong mobile Internet services in parallel with their Web services, as a response to market demands.

However, the interviewees saw many open questions that they wanted answered before they would start real development work. They could not commit to the walled garden portals of single operators since they had to serve all their customers, regardless of their operator. The operator portal was also seen as an obstacle that prevented direct contacts to clients and made their brands less visible. The interviewees had not thought how to comply with the variety of mobile devices. This was seen as a problem that device manufactures would solve by standardisation. Our conversion and adaptation proxy was seen as an interesting solution, but too indefinite for commercial services.

Within one year the mobile hype had faded away and most interviewees in the spring of 2001 were quite reserved about WAP solutions. They said that they would be ready to provide mobile Internet services as soon as they could see realistic business models in the field. Before that, new networks and terminal devices were needed. By this, they were referring especially to 2.5G networks that were not yet available at that time. Many service providers were developing SMS-based services because they saw that this technology would, after all, live longer than expected. SMS services were available for all mobile phone users, the technology was mature and the business model was working. This time the interviewees were quite interested in our conversion proxy approach. They saw that many of their services would be such that they could not charge for them. That is why the easy solution of providing the services via our conversion proxy was appealing, compared with setting up separate mobile services.

In the later interviews in 2001, the service providers still thought that they did not have to solve the problems with the variety of mobile client devices; rather they thought that device manufacturers would soon agree on suitable standards. Meanwhile, the service providers thought that an adequate solution was to concentrate on the most common devices on the target market.

After the inflated expectations of mobile Internet, the service providers did not expect the market of mobile services to be very big. They thought that many services would be such that the users would not be willing to pay for them. That is why service providers had to consider solutions that would require a minimum amount of development and maintenance efforts. Unfortunately there is a clear contradiction between these requirements and the user acceptance factors described in Chapter 7. The users would require comprehensive information, frequent updates and optimal adaptation for each client device.

8.2 Integrating business networks for location-aware services

The experiences with location-aware services were quite similar to WAP, although the hype with location-aware services was not as high. In 2001 service providers widely thought that location-aware services would be a significant consumer business in the near future. Within one year, in 2002 the focus was shifting more towards business use because the business models for consumer services were found to be too challenging after all, and the consumer market was not seen as promising as before.

As mobile services become more versatile, they will require a growing number of different business actors. As an example, Figure 26 presents the general architecture of one of the services that was evaluated in Paper VI, a location-aware integrated directory service, NAVISearch. NAVISearch aims to provide comprehensive content by integrating several directory services and databases. Local service providers can import information to NAVISearch, thus providing the users with topical and personal information about local services. With the NAVISearch service entity a mobile user can locate him/herself, search for nearby points of interest (POIs), such as shops, services and events, and get information about these POIs and see them on a map.

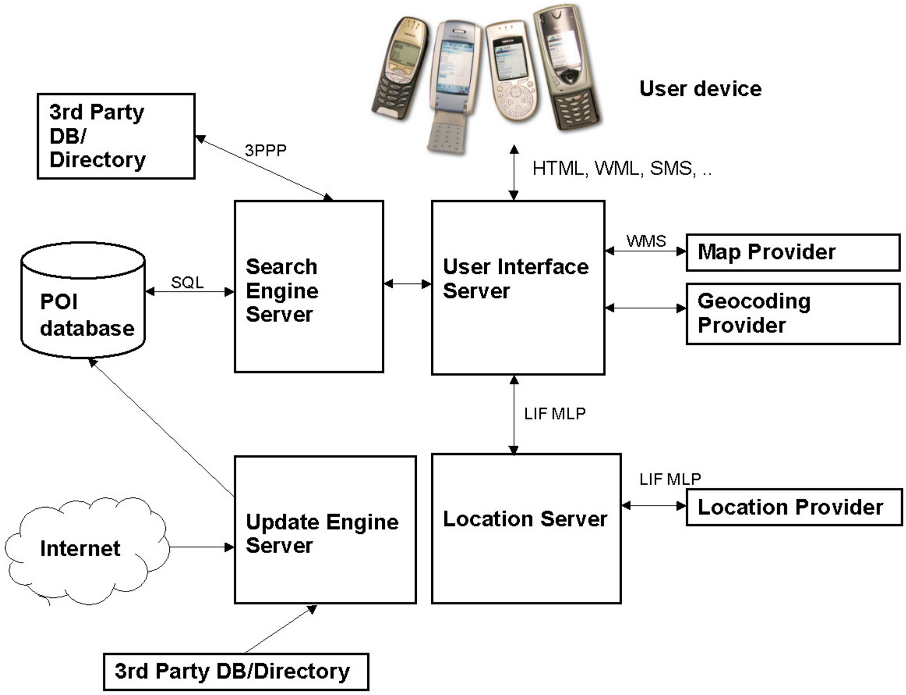


Figure 26. Architecture of NAVISearch location-aware integrated directory service.

The objective of NAVISearch was to implement a generic architecture that would combine a multitude of information sources, act as a platform for a variety of services, support different user terminals and interface location, geocoding and map services. The NAVISearch approach differed from many earlier approaches for providing location-aware services in the sense that NAVISearch truly integrated both independent services and individual databases. Related studies have addressed services where the service provider had full control over all of the content provided (Lee et al., 2002). NAVISearch provided a more versatile service but required quite a complicated business network.

Figure 26 illustrates the architecture of the NAVISearch service entity. The User Interface Server adapts the NAVISearch user interface to the client device in HTML, WML or as SMS, thus supporting various client devices. The Search Engine Server redirects the search simultaneously both to the POI database of NAVISearch and to third party directory services outside NAVISearch. The Location Server fetches the location of the user from location providers, thus supporting different positioning techniques. Maps are requested from Map Providers including e.g. location and map size as attributes. The geocoding provider converts street addresses into co-ordinates as needed. The NAVISearch Update Engine Server provides tools with which service providers can import information from Web pages or databases to the NAVISearch POI database. The tools can extract geotags from Web pages but if geotags are not available, the tools can convert textual address information on the Web pages into location co-ordinates.

Table 5 describes the business actors that would be involved in the business network of the NAVISearch location-aware integrated service directory. The classification is built on the one proposed by Kallio (2004) and presented in Chapter 4.3.2.

Table 5. The business roles in NAVISearch location-aware integrated service directory.

Business role	*Role in location-aware integrated service directory*
Application provider	Develops and implements the service software
Application service provider	Runs and maintains the service
Content aggregator	Markets the service and collects content from local companies such as shops and restaurants
Content integrator	Integrates the content collected by the content aggregator
Content provider	Local company that provides information to the POI database
Infrastructure provider	Location provider, Geocoding provider, Map provider or 3rd Party DB/Directory provider
Network provider	Telecom operator(s)
Service provider	Updates the service with information from the content integrator, acts as the contact point for the end users
Terminal manufacturer	Provides terminal devices
End user	Individual or company that uses the service

The results of the user evaluation of NAVISearch indicate that the users appreciated the comprehensive content enabled by the integration of several individual services and individual databases. NAVISearch would require a business network that integrates several directory service providers, content providers, positioning services, map services, geocoding services and route services. The situation gets even more complicated when several alternative actors provide the same basic service. For instance, the user might like to choose his/her primary map provider according to the quality and price of the offered service.

We evaluated the NAVISearch service with two kinds of service providers: a roadside assistance company and the tourist information office in the city of Turku. The roadside assistants said that they do not have many regular customers but people need their services occasionally and on the road. Their clients find them by calling some phone-based directory inquiry service and asking for "the nearest breakdown recovery service". Good visibility in these service directories is fundamental for their business, and they have to make sure that their contact information is included in all major directory services. For the breakdown recovery entrepreneurs an integrated service would be most welcome because they would not have to worry about having their service available in each individual directory service.

The city of Turku provided services for citizens and tourists on the Web but also some experimental services in the mobile Internet. The employees responsible for these Web services appreciated the NAVISearch solution that facilitated automatic importation of their Web content to the mobile service and even equipped the data with location information. This feature would be especially useful in location-aware event guides for the different events in the city throughout the year. Also, a service with minimal extra maintenance costs was considered very attractive for the representatives of the city because they saw that they could not charge for their services.

Small service and content providers would benefit from participating in an integrated service entity. Larger service providers may want to stick with their existing services, such as the 3rd party directory services in NAVISearch. Still, both actors will need to co-operate with a multitude of other business actors to set up a reasonable service entity: telecom operators, location providers, map providers, content aggregators and so on. To facilitate service entities such as NAVISearch, an integrator, i.e. the business actor who is responsible for the service entity and acts as the contact point for the end users, is in a central role.

8.3 Future challenges for mobile business models

After the unsuccessful introduction of mobile Internet, the renaissance of mobile services has not been easy. The service providers still have to cope with the bad reputation of mobile Internet. Today, the focus has turned towards business applications for mobile Internet and location-aware services (Alahuhta et al., 2005). Regarding consumer services, many service providers have returned to SMS-based services. These services have a high potential client population, the services can be implemented easily and they use an already mature technology.

As the NAVISearch example in Chapter 8.2 demonstrated, the business models for future mobile services will be quite complex. Recent interviews with Finnish service providers (Alahuhta et al., 2005; Kallio, 2004) point out that service providers see the complex business models as the main problem in implementing mobile services. The Japanese i-mode service is a good example of a success, gained by designing the business model in parallel with the technology and service development. The i-mode business model is very hierarchical and the operator controls practically everything. The open business models that European mobile markets are committed to have to construct co-operative and open business networks, allowing all actors to participate. These kinds of business networks will be much more difficult to construct and, in particular, to control.

Old business models cannot be transferred as such to new technology environments. Mobile business networks will include several actors, and new services should provide reasonable business to each of them, at the same time keeping the services affordable for end users. Peer-to-peer communication will be an important part of the business in the future as the experiences with mobile Internet have demonstrated in Japan, and the business models should take into account users as content providers.

When designing new services, it will be necessary to study what kinds of future business and value networks there will be, and what kind of combination of actors will be required to provide the services in practice. New services may generate new business networks that will appear and work side by side with existing ones. Business models and the pricing models should be developed and assessed in parallel with the design process. A multidisciplinary design approach is needed to involve experts in human sciences, technology and business.

9. Implications for usability design and evaluation methods

In our case studies we adopted human-centred design as defined by ISO 13407:1999 (ISO, 1999). We modified the approach so that it better supports the software development process as described in Chapter 6.1. Our case studies were carried out within research projects that produced technical infrastructures and built on them demonstrator applications. We applied the Technology Acceptance Model to study the factors that affect user acceptance of the foreseen applications of the technology. We involved end users, application field experts, content providers, service providers, terminal manufacturers, application developers and usability experts in the human-centred design process. In this way we were able to integrate different points of view into the design, and fit together user needs, technical possibilities and expectations of the service providers.

This chapter gives an overview of implications that our experiences indicate regarding the human-centred design process and related design and evaluation methods.

9.1 Wider views of the human-centred design process

New services may require increasingly complex service networks. Many actors in the service networks can be seen as users of new technical possibilities. The human-centred design approach should include the viewpoint of all these actors. Our case studies revealed that, like end users, service and content providers were often unaware of the technical possibilities and mislead by the marketing messages and the hype. All members of the service network should be included in the service development so that new technologies are developed, taking into account from the very beginning how the technology will be utilised in real applications and services and how these services will finally be provided to the users.

It is important to adopt the human-centred design approach in the development of individual services. In addition, new technologies need to be assessed from a wider view. This is especially important in technology research projects that should look further than just individual experimental applications. Scenario-based methods can be used to illustrate to potential end users and service providers the possibilities of new technologies as stories of everyday life, as we did in evaluating location-aware services in focus groups (Paper V). The end users and the service providers can give early feedback on the perceived value of the proposed concepts and they can suggest new service concepts.

In addition to including different business actors in the development of individual services, it is even more important to establish co-operation between the different actors of the future service networks, even before it has been decided what kind of services the network will start designing. The studies described in Papers V and VI were carried out within the Finnish Personal Navigation (NAVI) research and development programme. These kinds of programmes facilitate co-operation between different actors involved in developing and utilising evolving new technologies. Early co-operation facilitates fitting together technical possibilities, user needs and business opportunities instead of pushing through ready-made technical solutions. Horizontal activities, such as our studies on usability and ethical issues, facilitate collecting user expectations in general and early user experiences of first pilot services. By analysing and synthesising these results, common design guidelines can be drawn up to support the next waves of service design. Within the NAVI programme there were also horizontal research activities concerning technical solutions, legal issues and user needs. Co-operation with these activities constituted a firm basis for our studies on usability and ethical issues.

9.2 Towards value-centred design

Human-centred design should be enhanced towards value-centred design, as proposed by Boehm (2003) and Cockton (2004a,b). When collecting feedback from many different test users, designers may get lost in featurism and end up adding more and more features to the service. The emphasis of the design should be on deciding what are the targeted values to be provided to the end users and other stakeholders, and then focusing on designing and evaluating those values into the service. However, the design should be open enough to refine and change the targeted values according to user feedback.

Technology shapes the usage but also usage should shape the technology. A continuous dialogue will be needed between users and technology developers. The studies of evolving usage cultures provide an insight into how users adopt new technology. The technology should not just provide ready-made solutions but also give the users the freedom to innovate new ways of utilising the technology. Evaluation methods that facilitate a user-designer dialogue during evaluations are promising concepts to identify and utilise user innovations instead of treating users as passive research objects. These kinds of methods have been suggested e.g. by Hulkko et al. (2004), Iacucci and Kuutti (2002) and Isomursu et al. (2004).

As mobile services are more and more involved in the personal lives of the users and the users are increasingly dependent on these services in their lives, user trust becomes an important design goal. The foreseen contexts of use should be defined together with the users. Then the usage situations should be analysed regarding expected trust in the service. User expectations should be in balance with the actual trustworthiness of the solutions, including technical reliability, accuracy, privacy protection and user possibilities to keep the service under control.

9.3 Adapting to faster development processes

In today's fast design processes user feedback is often received too late and cannot be taken into account fully. More attention should be paid to user studies before the actual design process to understand more clearly the physical, social and technical environments in which the new technology is targeted.

With consumer products the heterogeneous user groups and the varying contexts of use force the designers to choose key user groups and key contexts of use in the design, because all possible contexts with all possible users cannot be studied. The decisions on the target user groups and contexts have to be made early in the design process and essential user groups may be overlooked. Thorough user studies should precede the actual design process so that design decisions are based on a good understanding of the users, their everyday lives and the existing and evolving usage cultures.

Initial user requirements for the WAP conversion proxy server were defined on the basis of literature studies, studies of existing Web services and an analysis of current and future mobile devices (Paper II). The initial requirements were fine-tuned in design walkthrough meetings. It was possible to identify many important usability requirements already during these phases. For instance, by analysing the existing and future mobile devices in our studies, we identified device-adaptation of the services as a major challenge. By analysing existing Web services, we saw that all Web content cannot be converted because of the varying quality of the content. We noticed that user feedback in the form of warnings about non-convertible content would be important. The large amount of information on many Web sites indicated a need for ways to browse through or skip excess information. These activities during the early phases of the design process did not require much effort and they were efficient ways to define the design targets more clearly.

In the studies of location-aware services early user feedback was collected by scenario evaluations in focus groups (Paper V), where we illustrated different future possibilities of location-aware services to the users in the form of scenarios. Scenario evaluations enabled us to include a variety of service concepts and a variety of user groups, thus preventing us from fixing targeted users too early.

9.4 From laboratory to the field

Our studies with location-aware services pointed out the trend that mobile products and services are increasingly related to their usage contexts, and may even only work in certain contexts. With these products traditional laboratory evaluations did not give much feedback, and in practice often the prototype evaluations had to be carried out in the field, where the users had the necessary infrastructure around them. In the NAVI programme, we had our own technical test environments in real usage environments and they turned out to be very useful is setting up field evaluations quickly.

Laboratory studies are important in the early phases of the design process. They give early feedback on the ease of use and understandability of the service. Value can also be studied in laboratory evaluations by interviewing the users and recording their comments during the test tasks. However, reliable feedback on user acceptance can be collected only by letting the users use the services on their own for a relatively long period. Also short-term field evaluations are useful for facilitating more realistic usage situations. With location-aware mobile services, we often had to have the laboratory evaluations outdoors because the services did not work technically indoors (e.g. GPS) or because the usage situation would have been pointless inside (e.g. route guidance).

When laboratory evaluations are brought outdoors, the evaluators need to be prepared for setbacks such as bad weather, test users not showing up, test personnel not available, device or service malfunctions or breakdowns and test equipment breakdowns. Spare batteries are needed for both the devices under evaluation and the test equipment. It is also worth considering whether the test users will need insurance against possible accidents during the evaluation.

Taking the service into use can often be tested in laboratory evaluations. However, it should be considered how complicated and long an operation it will be to take the service into use. It may require a separate evaluation session.

When conducting field trials, our main information source was usage logs. Log files gave precise data on how each individual user had been using the services, and also provided a basis for statistical analysis of the amount and type of usage. The log data was available to us throughout the field trial and we utilised it in planning interviews during and after the trial by identifying phenomena that we wanted to discuss with the test users. However, we could not log error situations, which often caused problems since we could not analyse the error situations described by the users. Further studies will be needed to implement logs with which we could analyse the reliability of the service.

The appropriate design of log functions for the evaluations is important and it should be done in parallel with the technical development because it may not be easy to add the log functionality to a ready-made system. It should be considered which data is needed and how the data could be automatically collected and converted into such a form that it is easy to analyse already during the evaluation. In our case studies it turned out extremely important to design the data analysis in parallel with the design of the logs. Manual analysis soon becomes impossible as the amount of data grows, and being able to export log data directly into analysis software saves a lot of time.

Key issues when planning the logs are:

- Identifying individual users
- Identifying individual usage sessions and activities during them
- Identifying error situations whenever possible
- Logging user location and other contexts if necessary and possible
- Informing the users about the log and getting written permission from them to collect and use log data
- Explaining clearly to the users how these logs will be collected, analysed and used.

When planning and carrying out field trials it must be ensured that the service is regularly maintained with topical and comprehensive content. Otherwise the users will have no motivation to use the service. The service has to be reliable enough to put it out to a field trial. Maintenance and usage support should be planned before the trial. It may not be possible to provide a 24/7 support service, but the users should be informed about the availability and the form of usage support. In the WAP trials the users were most active in the evenings and during the weekends, and that is why usage support only during office hours was not enough. When planning the usage support it should be remembered that each error situation demotivates the user and makes him/her more reluctant to use the service again. The evaluators should be prepared for delays – the service has to be tested thoroughly for technical problems before starting the field trial. It is much better to postpone the trial than to start it with an unstable and unreliable service.

Continuous contact with the users during the field trial helps to identify problems and gives the evaluators a clear understanding of how the trial is going. The log files help in this, for instance identifying passive users to find out whether low usage is due to technical problems. The users tend not to be active in reporting these problems; instead they just stop using the service.

Ideally, the users should be allowed to use the service totally freely during the test period. However, according to our experiences this often leads to a situation where many users do not use the service at all. That is why we agreed on rules for the field test with the users. For instance, in the WAP trials we asked the users to use the service at least twice a week. The usage over and above the agreed minimum gave us feedback on real usage needs.

Supportive information collected by the users, e.g. in the form of a usage diary, is often useful. In our field studies of location-aware services, photo diaries were easy to maintain for the users with their media phones. This gave concrete illustrations of the real usage situations and helped the users to remember them when later interviewed.

Setting up the test service for the field trials is often time-consuming and expensive and that is why the set-up should be utilised to its full potential. It would be a pity to close it down right after the trial. In the WAP field studies (Paper IV) we kept the service on continuously after the first field trial. We could then easily organise small or larger scale field trials any time during the project.

10. Conclusions and discussion

Mobile Internet services have been emerging since 1999 but big commercial breakthroughs are still absent with rare exceptions such as Japanese mobile services, NTT DoCoMo's i-mode in the forefront. The experiences to date point out that a mobile phone is first and foremost a personal communication device. Services that complement and enhance person-to-person communications have been the most successful ones. These kinds of services include e-mail, chat, message boards and messaging services that enable user interaction with services on other media such as TV (cross-media services). Traditional Web information and transaction services transferred to mobile phones have achieved quite modest usage figures. A major reason for this is that these services are needed only occasionally on the move. Services that provide topical fast-changing information while on the move and that take advantage of user mobility are more promising. Location-aware guidance services are good examples of these kinds of services. Location is one of the first context attributes that can be measured quite accurately and easily. As more context attributes such as attributes of the physical or social environment are made technically available, new kinds of mobile services can be provided to the users and existing services can be provided in personally and contextually relevant forms.

Current research on mobile services has mainly concentrated on the usability of alternative user interface implementations. Small mobile devices pose significant usability challenges and the usability of the services is still worth studying. However, more attention should be paid to user acceptance of the planned services. The reason for many commercial failures can be traced back to the wrongly assessed value of the services to the users. It is not easy to find a single mobile service that consumers would value highly enough to bother taking mobile Internet into use. In Japan the selection of mobile services was wide from the very beginning, and email boosted all the other services. The occasional usage needs identified for many mobile services indicate that the success factor of mobile services may be a wide selection of services rather than any individual service. The case studies presented in this dissertation indicate that users will value personally and contextually relevant, topical and comprehensive information and services that are easy to access. In addition, service adoption needs to be made much easier than it is today, and the users should easily get information about available services while on the move.

10.1 Back to the research questions

The main research question of this dissertation was:

1. *What are the key attributes that affect user acceptance of mobile services, focusing on mobile Internet services and location-aware information services targeted at consumers?*

Based on the results of the case studies, I propose the Technology Acceptance Model for Mobile Services. The model is a modification of the original Technology Acceptance Model by Davis (1989). According to the Technology Acceptance Model for Mobile Services, user acceptance of mobile services is built on three factors: perceived value of the service, perceived ease of use and trust. These three factors affect the intention to use a mobile service. With mobile services targeted at consumers, taking the services into use is often a major obstacle to the user. A fourth user acceptance factor, perceived ease of adoption becomes more prominent as the user proceeds from an intention to use to actual usage.

Instead of a collection of useful features, mobile services should provide value to the user. The value of the service can be based on utility, communication or fun provided to the user. A successful service will offer comprehensive, topical and familiar information, and will also include personal and user-generated content. The users value seamless service entities whereby the user is supported throughout the whole usage situation, even from one terminal device to another.

Ease of use requires a clear overview of the service entity, fluent navigation on a small screen and smooth user interaction with the service. The users should get personally relevant services and information without the need to expend effort on personalisation, and they should have easy access to situationally relevant information and services. The services should be designed to be adaptive to the wide variety of devices and networks.

As individual users increasingly use mobile services in their everyday lives and the services increasingly collect, analyse and redirect personal data, user trust in the services becomes more and more important. The user should be able to assess whether (s)he can rely on the service being available when needed, and whether the service provides accurate enough information and functions in the intended contexts of use. The user needs to feel and really be in control, and the privacy of the user must be protected.

Occasional usage and momentary usage sessions on the move are typical of mobile services. In addition, services are increasingly available only locally or in certain contexts of use. This indicates a need for disposable services: services that are easy to find, take into use, use and get rid of when no longer needed. The user needs realistic information about the actual values of the services, so that (s)he can realise how to utilise the service in his/her everyday life and innovate new usage possibilities.

The proposed Technology Acceptance Model for Mobile Services provides a framework for future user acceptance studies. The framework helps designers of mobile services identify key issues on which the design should be focused in order to ensure user acceptance. Thus the motivation of the model is different than the motivation of the original TAM (Davis, 1989) and the extensions to it that were built to explain user acceptance and its underlying forces for existing technical solutions. The proposed model can also be extended by identifying the underlying forces that influence the user acceptance factors. As the original TAM suggests, these underlying forces are related to the characteristics of the user and his/her social environment. To identify these forces and their relation to the user acceptance factors, user surveys will be needed.

The case studies (Papers I–VI) that the model is based on deal with mobile Internet services and location-aware information services targeted for consumers. Entertainment and communication were in minor roles. For these kinds of services the model can be applied as a starting point, but it may need modifications.

The Technology Acceptance Model for Mobile Services was constituted by analysing and synthesising the results of a series of case studies. When user acceptance studies are carried out in connection to technology research as we did, the studies generate knowledge that can be utilised in future commercial service development. It is important to crystallise the key findings to design implications that direct the designers’ focus on key user acceptance factors. As design processes are getting faster and the services and technical infrastructures are getting more complex, the Technology Acceptance Model for Mobile Services together with the design implications provide efficient tools to communicate previous user acceptance findings to the design.

The supplementary research questions were:

2. *What factors affect the acceptance of service providers to utilise mobile Internet and positioning in their services, and what kinds of future service concepts does this indicate?*

The main challenges for providing comprehensive mobile services are the complex business models needed. The operator-centric business model that facilitated the success of i-mode in Japan cannot be copied into European business cultures. Instead, working co-operation models of the business actors are needed both in service development and maintenance. Services that can provide the users with personally and contextually relevant information or functions are promising concepts. Mobile services also have good potential in complementing services available in other media. Mobile services may provide already existing customers with added value, help in reaching new customers, and offer good alternatives in providing non-commercial services. Niche services targeted at well-defined user groups such as hobbyists or tourists at a particular resort are also promising service concepts. In Japan, email boosted all the other mobile Internet services. As usage needs for individual mobile services are quite occasional, market breakthrough and profitable services will require service entities consisting of several services and will have to be targeted at a wide user population.

3. *What are the implications of research questions 1 and 2 for the human-centred design process of mobile services and related design and evaluation methods?*

Human-centred design should involve all the actors that are needed in the business networks to provide mobile services. Service and content providers are not necessarily experts on information technology but they are the best experts to innovate how they can serve their customers better by utilising new mobile technology. The wider view of human-centred design enables the assessment of new technologies from three points of view: what are the possibilities of the technology, what do the end users value, and how can the service providers utilise mobile technologies in their existing or future business. Featurism, where the service gets equipped with more and more features, will not be the right approach for designing mobile services. Value-centred design is a promising approach that aims at identifying the key targeted value(s) for end users, service providers and other stakeholders, and then focuses on these values in the design.

Human-centred design has to move increasingly from the laboratory to the field to be able to study user acceptance of mobile services in real contexts of use. At the same time, the design and evaluation methods need to adapt to the fast and concurrent service development processes.

10.2 What could have been done otherwise

In the WAP studies we used the original Technology Acceptance Model (Davis, 1989) as the research framework. User evaluations were focused on ease of use and utility, but we were also open to other kinds of findings. In the evaluations of location-aware services ethical issues and the introduction of the services to consumers were set as additional research topics. The numerous user evaluation findings on user trust in the field trials of both WAP and location-aware services, and the repeated test user comments of their difficulties in taking commercial services into use, indicated a need to refine our research framework. The construction of the Technology Acceptance Model for Mobile Services was thus based on the field study findings. In the future, as the Technology Acceptance Model for Mobile Services is used as the framework of the evaluation from the very beginning, we will probably be more effective in identifying issues related to the new acceptance factors of trust and perceived ease of adoption. However, the evaluation still needs to be open to additional findings, outside the actual framework.

The evaluations with the service providers tended to be loosely coupled to the other parts of the work. In practice, we had separate interviews with service providers or included representatives of service providers as normal test users in the evaluations. In the future, service providers need to be integrated into the human-centred design process more tightly. Methods that facilitate communication and the exchange of ideas between technology experts, end users and service providers would certainly be fruitful for identifying new ways of utilising mobile technologies.

The Innovation Diffusion Theory (Rogers, 1995) could have been integrated more tightly into our research, even though the technology research projects took place well before the estimated start of the diffusion of mobile services. In most case studies we had relatively small test user groups. However, in the evaluations of the mobile Internet services we had two long-term field trials with 40 users in each. We selected the users so that they represented different groups that could utilise mobile Internet services in group communication in addition to utilising the individual services. We included in the test user groups users of both genders, different ages, professions and backgrounds. It would have been useful to position the users into the different user groups of the Innovation Diffusion Model. In this way we could have analysed the evaluation results in connection with the foreseen innovation diffusion in different user groups.

In future research, we should critically analyse new technology in relation to the hype cycle (Linden and Fenn, 2003). We should analyse the phase of the technology in the cycle and proportion user feedback to the possible hype effect. In our WAP studies we experienced both positive and negative hype during the field trials and the interviews of the service providers. Even though we recognised that the hype affected some of the attitudes, we did not systematically analyse the effect of it.

The research and development network of the Personal Navigation (NAVI) programme, in connection with which we did the studies of location-aware services, managed to bring together many actors in the application field. The programme turned out to be a good co-operation model to raise topics for research and to jointly study issues such as key technologies, business models, user needs, usability issues, ethical issues and legal issues. However, in the NAVI programme the service and content providers were not as well represented as the device and network providers. In the future, in similar activities, service and content providers should be better represented. These actors are the best experts to assess how new mobile technologies can be utilised in serving their customers.

10.3 Future directions

The studies in this dissertation were related to mobile Internet services and location-aware information services targeted for consumer use. The identified user acceptance factors can be utilised in designing these kinds of services, but they can also be applied when designing other kinds of mobile services. In future visions, mobile services are increasingly interacting with their environment and are transforming into tools with which the user can orient in and interact with the environment. As the user moves from one environment to another the available services will change accordingly. These kinds of services will require extreme ease of adoption, and, as the services will increasingly deal with personal data, the user's trust in the services will become an even more important user acceptance factor.

The Technology Acceptance Model for Mobile Services defines a framework for user acceptance factors of mobile services. Further studies will be needed to study the mutual relations of the four user acceptance factors. As with the original TAM, the model can be enhanced by studying key forces underlying the judgements of perceived value, perceived ease of use, trust and perceived ease of adoption.

The results of this dissertation indicate that in technology research projects human-centred design needs to be expanded from individual services to service entities and key technologies. The technologies should be assessed with end users, service providers and technology providers in order to innovate the best usages for new technical possibilities. Horizontal usability studies enable the identification of key usability challenges and best practices in selected application areas. Studies on business models enable the study of the business potential of suggested new services and the setting up of business networks for the most promising service concepts.

Our methodology of analysing and synthesising the results of parallel research and development projects can be utilised in many kinds of other research as well. Detailed usability evaluation results of individual demonstrator applications may not as such be applicable to future product development. By synthesising and generalising the results of parallel research activities, key user acceptance factors and design implications for future service development can be identified.

We are currently applying value-centred design in connection with the research and development of sensor and tag technologies. These are the key technologies to enhance mobile phones into tools with which the user can orient in and interact with his/her environment. This concept can be utilised in different application fields to implement ubiquitous computing services in such a way that the user feels and really is in control. The value-centred design process brings together technical experts, business actors from different application fields, application field experts and end users. The objective of our work is to identify the most promising application concepts in which the new technology could be utilised, and to define user acceptance criteria for those concepts. As the technical research is focused on the basic components and technical infrastructures of ubiquitous computing, we have several degrees of freedom in studying with users and application field experts where those components could be best utilised.

We are also continuing the horizontal usability studies with enhanced focus. We have set up a horizontal research project, where ten research groups from different universities and research institutes in Finland are bringing together their case studies of ubiquitous computing. By analysing and synthesising the results together, we expect to be able to identify criteria for acceptable ubiquitous computing services in different application fields. In addition we will combine our experiences to identify and develop best practices of design and evaluation methods.

The hype cycle of technologies is a fascinating phenomenon. We should increasingly analyse the phase of the technologies on the cycle, and interpret our evaluation results accordingly. In particular, we should critically analyse our own dissemination activities, to make sure that irrespective of the hype, our research remains impartial, and in particular that we are not the ones adding fuel to the hype.

References

Ackerman, M., Darrel, T. and Weitzner, D. J. 2001. Privacy in context. Human-Computer Interaction, 16, pp. 167–176.

Ajzen, I. and Fishbein, M. 1980. Understanding attitudes and predicting social behaviour. Prentice Hall.

Alahuhta, P., Jurvansuu, M. and Pentikäinen, H. 2004. Roadmap for network technologies and services. Tekes, Helsinki, Finland, Technology Review 162/2004. ISBN 952-457-176-5.

Alahuhta, P., Ahola, J. and Hakala, H. 2005. Mobilizing business applications. A survey about the opportunities and challenges of mobile business applications and services in Finland. Tekes, Helsinki, Finland, Technology Review 167/2005. ISBN 952-457-184-6.

Ancona, M., Locati, S. and Romagnoli, A. 2001. Context and location aware textual data input. SAC 2001, Las Vegas. ACM 2001. Pp. 425–428.

Antifakos, S., Schwaninger, A. and Schiele, B. 2004. Evaluating the effects of displaying uncertainty in context-aware applications. In: Davies, N., Mynatt, E. and Siio, I. (eds.). Proceedings of Ubicomp 2004: Ubiquitous Computing 6th International Conference. Springer-Verlag. Pp. 54–69.

Barkhuus, L. and Dey, A. 2003. Is context-aware computing taking control away from the user? Proceedings of UbiComp 2003 Conference. Pp. 150–156.

Barnes, S. J. and Huff, S. L. 2003. Rising Sun: imode and the wireless internet. Communications of the ACM, Vol. 46, No. 11, pp. 79–84.

Battarbee, K. 2004. Co-experience. Understanding user experience in social interaction. Academic dissertation. University of Arts and Design, Helsinki. ISBN 951-558-158-3.

Bellavista, P., Corradi, A. and Stefanelli, C. 2002. The Ubiquitous provisioning of Internet services to portable devices. IEEE Pervasive Computing, Vol. 1, No. 3, July–September, pp. 81–83.

Bergman, E. and Norman, D. A. 2000. Making technology invisible: A conversation with Don Norman. In: Bergman, E. (ed.). Information Appliances and Beyond. Morgan Kaufmann Publishers. Pp. 10–26.

Beyer, H. and Holzblatt, K. 1998. Contextual design: Defining customer-centered systems. Morgan Kaufmann, San Francisco.

Bieber, G. and Giersich, M. 2001. Personal mobile navigation systems – design considerations and experiences. Computers & Graphics, 25, pp. 563–570.

Billsus, D., Brunk, C. A., Evans, C., Gladish, B. and Pazzani, M. 2002. Adaptive interfaces for ubiquitous Web access. Communications of the ACM, Vol. 45, No. 5, pp. 34–38.

Blom, J. 2000. Personalisation – a taxonomy. Extended abstracts of CHI'00. ACM, New York. Pp. 313–314.

Boehm, B. 2003. Value-based software engineering. Software Engineering Notes, Vol. 28, No. 2, March 2003, pp. 1–12. ACM SIGSOFT.

Brown, B. and Perry, M. 2002. Of maps and guidebooks. Designing geographical technologies. DIS2002 Conference, London, UK, 25–28 June 2002. ACM Press. Pp. 246–254.

Buchanan, G., Farrant, S., Jones, M., Thimbleby, H., Marsden, G. and Pazzani, M. 2001. Improving mobile Internet usability. Proceedings of 10th World Wide Web Conference, WWW10. ACM. Pp. 673–679.

Buranatrived, J. and Vickers, P. 2004. A study of application and device effects between a WAP phone and a Palm PDA. In: Brewster, S. and Dunlop, M. (eds.). Mobile Human-Computer Interaction – Mobile HCI 2004. Proceedings. Springer-Verlag. Lecture Notes in Computer Science 3160. Pp. 192–203.

Chae, M. and Kim, J. 2004. Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard Web phones. Behaviour & Information Technology, Vol. 23, No. 3, pp. 165–185.

Chalmers, D., Dulay, N. and Sloman, M. 2004. A framework for contextual mediation in mobile and ubiquitous computing applied to the context-aware adaptation of maps. Personal and Ubiquitous Computing, 8, pp. 1–18.

Chávez, E., Ide, R. and Kirste, T. 1999. Interactive applications of personal situation-aware assistants. Computers & Graphics, 23, pp. 903–915.

Chen, L., Gillenson, M. L. and Sherell, D. 2004. Consumer acceptance of virtual stores: a theoretical model and critical success factors for virtual stores. ACM SIGMIS Database archive, Volume 35 , Issue 2, Spring, pp. 8–31.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C., 2000, Developing a context-aware electronic tourist guide: some issues and experiences. CHI 2000 Conference Proceedings. Pp. 17–24.

Cheverst, K., Mitchell, K. and Davies, N. 2002. Exploring context-aware information push. Personal and Ubiquitous Computing, 6, pp. 276–281.

Chitarro, L. and Dal Chin, P. 2002. Evaluating interface design choices on WAP phones: Navigation and Selection. Personal and Ubiquitous Computing, 6, pp. 237–244.

Ciavarella, C. and Paterno, F. 2004. The design of a handheld, location-aware guide for indoor environments. Personal and Ubiquitous Computing, 8, pp. 82–91.

Cockton, G. 2004a. From Quality in Use to Value in the World. Proceedings of CHI2004. ACM. Pp. 1287–1290.

Cockton, G. 2004b. Value-Centred HCI. Proceedings of the Third Nordic Conference on Human-Computer Interaction, pp. 149–160.

Concejero, P., Ramos, R., Castellanos, M. A. and de Herrero, C. G. 2000. Usability evaluation tutorial. In: Concejero, P. (ed.). Usability in the Information Society. European Commission. ISBN 84-89900-23-X. (CD)

Davis F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quartely, 13/1989, pp. 319–339.

Davis, F. D. and Venkatesh, V. 2004. Toward preprototype user acceptance testing of new information systems: implications for software project management. IEEE Transactions on Engineering Management, Vol. 51, No 1.

Dey, A. K. 2001. Understanding and using context. Personal and Ubiquitous Computing, 5, pp. 20–24.

Dey A. K., Abowd G. D. and Salber D. 2001. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. International Journal on Human-Computer Interaction, 16:2, pp. 97–166.

Dix, A., Finlay, J., Abowd, G. and Beale, R. 1997. Human-Computer Interaction. Second edition. Prentice Hall.

Dunlop, M. and Brewster, S. 2002. The challenge of mobile devices for human-computer interaction. Personal and Ubiquitous Computing, 6, pp. 235–236. Springer-Verlag.

EDeAN. 2005. European Design for All e-Accessibility Network. Homepage. www.e-accessibility.org/. [online, cited 4.1.2005]

Espinoza, F., Persson, P., Sandin, A., Nyström, H., Cacciatore, E. and Bylund, M. 2001. GeoNotes: Social and navigational aspects of location-based information systems. Lecture Notes in Computer Science, Vol. 2201, pp. 2–17.

European Union. 2002. Directive 2002/58. Privacy and Electronic Communications.

Fano, A. E. 1998. Shopper's Eye: Using location-based filtering for a shopping agent in the physical world. Autonomous Agents '98 Conference.

Finnish legislation. 2004. Sähköisen viestinnän tietosuojalaki 516/2004. (Legislation for privacy protection in electronic communication). In Finnish.

Fishbein, M. and Ajzen, I. 1975. Belief, Attitude, Intention and Behaviour. An Introduction to Theory and Research. Addison–Wesley, Reading, MA.

Fithian, R., Iachell, G., Moghazy, J., Pousman, Z. and Statsko, J. 2003. The design and evaluation of a mobile location-aware handheld event planner.

Mobile HCI 2003 Conference, Udine, Italy, 8–11 September 2003. Springer. Pp. 145–160.

Fogg, B. J. and Tseng, H. 1999. The elements of computer credibility. Proceedings of CHI 99 Conference. Pp. 80–87.

Funk, J. L. 2004. Mobile disruption. The technologies and applications driving the mobile Internet. Wiley-Interscience. ISBN 0-471-51122-6.

Gartner Research. 2004. Hype Cycle for mobile and wireless hardware and software. Strategic Analysis Report GOO120921. 9 June 2004.

Gefen, D. 2000. E-Commerce: The role of Familiarity and Trust. Omega: the international journal of management science, Vol. 28, pp. 725–737.

Gefen, D. and Devine, P. 2001. Customer loyalty to an on-line store: The meaning of online service quality. Proceedings of the 22nd international conference on information systems. Pp. 613–617.

Gefen, D., Karahanna, E. and Straub, D. W. 2003. Inexperience and experience with online stores: the importance of TAM and Trust. IEEE Transactions on Engineering Management, Vol. 50, No. 3, August 2003, pp. 307–321.

Gulliksen, J., Blomkvist, S. and Göransson, B. 2003. Engineering the HCI profession or softening development processes. In: Jacko, J. and Stephanidis, C. (eds.). Human-Computer Interaction. Theory and Practice (Part I). Volume 1 of the Proceedings of HCI International 2003, 10th International Conference on Human-Computer Interaction. Pp. 118–122.

Henderson, A. 2005. Design: The innovation pipeline: design collaborations between design and development. ACM Interactions, Vol. 12, No. 1, pp. 24–29.

Hjelm, J. 2000. Designing Wireless Information Services. John Wiley & Sons, Inc.

Hollensberg, J. and Vind Nielsen, H. 2000. Survey of usability on Danish WAP portals – How the Danish telecom operators entered the WAP world. Mobilethink as. [online, cited 29 October 2001] http://www.mbusiness-insight.de/deutsch/download.html

Holmqvist, L. 1999. Will baby faces ever grow up? In: Bullinger, H.-J. and Ziegler, J. (eds.). Human-Computer Interaction. Ergonomics and User Interfaces. Volume 1 of the Proceedings of the 8th International Conference on Human-Computer Interaction, 1999. Pp. 706–709.

Hulkko, S., Mattelmäki, T., Virtanen, K. and Keinonen, T. 2004. Mobile probes. Proceedings of the Third Nordic Conference on Human-Computer Interaction. ACM Press. Pp. 43–51.

Höök, K. 2004. Active co-construction of meaningful experiences: but what is the designer's role? Proceedings of the Third Nordic Conference on Human-Computer Interaction. ACM Press. Pp. 1–2.

Iacucci, G., Kela, J. and Pehkonen, P. 2004. Computational support to record and re-experience visits. Personal and Ubiquitous Computing, 8, pp. 100–109.

Iacucci, G. and Kuutti, K. 2002. Everyday life as a stage in creating and performing scenarios for wireless devices. Personal and Ubiquitous Computing, 6, pp. 299–306.

IBM. 2005. Out-of-box experience. www-3.ibm.com/ibm/easy/eou_ext.nsf/publish/577 [Online, cited 10.1.2005]

Ikonen, V., Ahonen, A., Kulju, M. and Kaasinen, E. 2002a. Trade description model – helping users to make sense of the new information technology products. In: Wiszniewski, B. (ed.). Electronic Commerce: Theory and Applications. Proceedings of 2nd International Interdisciplinary Conference on Electronic Commerce. Gdansk, Poland. Pp. 57–63. ISBN 83-88617-31-1.

Ikonen, V., Anttila, V., Petäkoski-Hult, T., Sotamaa, O., Ahonen, A., Schirokoff, A. and Kaasinen, E. 2002b. Products and Services for Personal Navigation – Adaptation of Technology and Usage Cultures. Publications of the NAVI programme. www.vtt.fi/virtual/navi [online, cited 4.1.2005]

Islam, N. and Fayad, M. 2003. Toward ubiquitous acceptance of ubiquitous computing. Communications of the ACM, Vol. 46. No. 2, pp. 89–92.

ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. International Standardization Organization. Geneve.

ISO 13407:1999. Human-centred design processes for interactive systems. International standard. International Standardization Organization. Geneve.

Isomursu, M., Kuutti, K. and Väinämö, S. 2004. Experience clip: method for user participation and evaluation of mobile concepts. Proceedings of Participatory Design Conference 2004. Toronto, Canada. ACM. Pp. 83–91.

ITU 2002. Internet for a mobile generation. ITU Internet reports. Revised edition, 19 September 2002. International Telecommunication Union (ITU). ISBN 92-61-09851-7. http://www.itu.int/publications/docs/gs/intrep4ed.html [online, cited 4.1.2005]

Järvenpää, S. L., Lang, K. R., Takeda, Y. and Tuunanen V. K. 2003. Mobile commerce at crossroads. Communications of the ACM, Vol. 46. No. 12, pp. 41–44.

Kaasinen, E. (ed.) 2003. Products and Services for Personal Navigation – Usability Design. Part III. Case Studies and Usability Guidelines. Publications of the NAVI programme. www.vtt.fi/virtual/navi [online, cited 4.1.2005]

Kaasinen, E., Ikonen, V., Ahonen, A., Anttila, V., Kulju, M., Luoma, J. and Södergård, R. 2002. Products and Services for Personal Navigation – Classification from the User's Point of View. Publications of the NAVI programme. www.vtt.fi/virtual/navi [online, cited 4.1.2005]

Kaasinen, E., Kantola, K. and Olphert, W. 2000. Human-centred design process. In: Concejero, P. (ed.). Usability in the Information Society. European Commission. ISBN 84-89900-23-x. (CD)

Kaikkonen, A. and Roto, V. 2003. Navigating in mobile XHTML application. Proceedings of CHI 2003 Conference. Pp. 329–336.

Kallio, P. 2004. Emergence of Wireless Services. Business Actors and Their Roles in Networked Component-based Development. Espoo: Technical Research Centre of Finland, VTT Publications 534. 118 p. + app. 71 p. ISBN 951-38-6386-7; 951-38-6387-5.

Kallio, T. and Kekäläinen, A. 2004. Improving effectiveness of mobile application design: user-pairs testing by non-professionals. In: Brewster, S. and Dunlop, M. (eds.). Mobile Human-Computer Interaction – Mobile HCI 2004. Proceedings. Springer-Verlag. Lecture Notes in Computer Science 3160. Pp. 315–319.

Karahanna, E.; Straub, D. and Chervany, N. 1999. Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Quartely 23, pp. 183–207.

Kasesniemi E.-L., Ahonen A., Kymäläinen T. and Virtanen T. 2003. Elävän mobiilikuvan ensi tallenteet. Käyttäjien kokemuksia videoviestinnästä. (First recordings of live mobile video. User experiences). Espoo: Technical Research Centre of Finland, VTT Tiedotteita – Research Notes 2204. 95 p. In Finnish. ISBN 951-38-6158-9; 951-38-6159-7.

Ketola, P. 2002. Integrating usability with concurrent engineering in mobile phone development. Academic Dissertation. University of Tampere. Department of Computer and Information Sciences. A2002-5. ISBN 951-44-5359-x.

Kiljander, H. 2004. Evolution and usability of mobile phone interaction styles. Doctoral thesis. Helsinki University of Technology. Publications in Telecommunications Software and Multimedia. TML-A8. ISBN 951-22-7319-5. Otamedia Oy.

Kindberg, T., Sellen, A. and Geelhoed, E. 2004. Security and trust in mobile interactions – a study of users' perceptions and reasoning. In: Davies, N., Mynatt, E. and Siio, I. (eds.). Proceedings of Ubicomp 2004: Ubiquitous Computing 6th International Conference. Springer-Verlag. Pp. 196–213.

Kohrs, A. and Merialdo, B. 2001. Creating user-adapted Web sites by the use of collaborative filtering. Interacting with Computers, 13, pp. 695–716.

Kolari J., Laakko T., Kaasinen E., Aaltonen M., Hiltunen T., Kasesniemi, E.-L., Kulju, M. and Suihkonen, R. 2002. Net in Pocket? Personal mobile access to Web services. Espoo: Technical Research Centre of Finland, VTT Publications 464. 135 p. + app. 6 p.

Lee, D. L., Xu, J., Zheng, B. and Lee, W.-C. 2002. Data management in location-dependent information services. IEEE Pervasive Computing, Vol. 1, No. 3, July–September, pp. 65–72.

Legris, P., Ingham, J. and Collerette, P. 2003. Why do people use information technology? A critical review of the Technology Acceptance Model. Information & Management, 40, pp. 191–204.

Liew, E., Misuyama, N. and Bourlias, C. 2004. Knowing Customers is Key for Mobile Data Services in Japan. Research Brief. 1 December 2004. Gartner Research.

Linden, A. and Fenn, J. 2003. Understanding Gartner's Hype Cycles. Strategic Analysis Report R-20-1971. 30 May 2003. Gartner Research.

MacKay, B., Watters, C. and Duffy, J. 2004. Web page transformation when switching devices. In: Brewster, S. and Dunlop, M. (eds.). Mobile Human-Computer Interaction – Mobile HCI 2004. Proceedings. Springer-Verlag. Lecture Notes in Computer Science 3160. Pp. 228–239.

Manber, U., Patel, A. and Robinson, J. 2000. Experiences with personalisation of Yahoo! Communications of the ACM, Vol. 43, Issue 8, pp. 35–39.

Marcus, A. and Chen, E. 2002. Designing the PDA of the future. ACM Interactions, 9:1, pp. 34–44.

Mathieson, K., Peacock, E. and Chin, W. W. 2001. Extending the Technology Acceptance Model: the influence of perceived user resources. The DATA BASE for Advances in Information Systems, Vol. 32, No. 3, pp. 86–112.

Mitsuyama, N. 2003. NTT DoCoMo: i-mode Wireless Internet Services. Operational Management Report. Gartner Research. DPRO-96595.

Moore, G. A. 1999. Crossing the Chasm. Second Edition. Capstone Publishing, Oxford.

Moran, T. P. and Carrol, J. M. 1996. Overview of Design Rationale. In: Moran, T. P., Carrol, J. M. (eds.). Design Rationale. Concepts, Techniques and Use. Lawrence Erlbaum Associates, Inc. ISBN 0-8058-1566-X.

Nah, F. F.-H., Siau, K. and Sheng, H. 2005. The value of mobile applications: a utility company study. Communications of the ACM, Vol. 48, No. 2 (February), pp. 85–90.

Nielsen, J. 1993. Usability Engineering. Academic Press, Inc. ISBN 0-12-518405-0.

Nielsen, J. 1998. Personalization is Over-Rated, Jakob Nielsen's Alertbox, October 4, 1998. [online, cited 29 October 2001] http://www.useit.com/alertbox/981004.html

Nokia. 2004. Tips and tricks for designing XHTML content. Version 1.1. www.forum.nokia.com [online, cited 4.1.2005]

Norman, D. A. 1998. The Invisible Computer. MIT Press, Cambridge, MA.

Norros, L., Kaasinen, E., Plomp, J. and Rämä, P. 2003. Human-Technology Interaction. Research and Design. VTT Roadmap. Espoo: Technical Research Centre of Finland, VTT Research Notes 2220. 118 p. + app. 11 p.

NTT DoCoMo, Inc. 2002. i-mode service guideline. Version 1.2.0. 4 March 2002. www.nttdocomo.com/corebiz/imode/why/guideline/guideline.html [online, cited 4.1.2005]

Perkowitz, M. and Etzioni, O. 2000. Adaptive Web sites. Communications of the ACM, Vol. 43, No. 8, pp. 152–158.

Pitkänen, O., Mäntylä, M., Välimäki, M. and Kemppinen, J. 2002. Assessing legal challenges on the mobile Internet. HIIT Technical Report 2002-4. www.hiit.fi

Pittet, S. 2004. User Survey: Mobile Phones and Services, Western Europe, 2004 (Executive Summary). 15 November 2004. Gartner Research.

Plomp J., Ahola J., Alahuhta P., Kaasinen E., Korhonen I., Laikari, A., Lappalainen, V., Pakanen, J., Rentto, K. and Virtanen, A. 2002. Smart Human Environments. In: Sipilä, M. (ed.). Communications Technologies. The VTT Roadmaps. Espoo: Technical Research Centre of Finland, VTT Research Notes 2146. Pp. 61–81.

Po, S., Howard, S., Vetere, F. and Skov, M. B. 2004. Heuristic evaluation and mobile usability: Bridging the realism gap. In: Brewster, S. and Dunlop, M. (eds.). Mobile Human-Computer Interaction – Mobile HCI 2004. Proceedings. Springer-Verlag. Lecture Notes in Computer Science 3160. Pp. 49–60.

Pousman, Z., Iachello, G., Fithian, R., Moghazy, J. and Stasko, J. 2004. Design iterations for a location-aware event planner. Personal and Ubiquitous Computing, 8, pp. 117–125.

Rainio, A. 2003. Henkilökohtainen navigointi, NAVI-ohjelma. Loppuraportti. (Personal Navigation NAVI programme. Final report) Liikenne- ja viestintäministeriön julkaisuja 11/2003. Edita Prima Oy, Helsinki. ISBN 951-723-824-X. In Finnish.

Ramsay, M. and Nielsen, J. 2000. WAP Usability. Déjà Vu: 1994 All Over Again. Report from a Field Study in London, Fall 2000. Nielsen Norman Group.

Rao, B. and Minakakis, L. 2003. Evolution of mobile location-based systems. Communications of the ACM, Vol. 46. No. 12, pp. 61–65.

Rogers, E. M. 1995. The diffusion of innovations. Fourth edition. New York. Free Press.

Sarker, S. and Wells, J. D. 2003. Understanding mobile handheld device use and adoption. Communications of the ACM, Vol. 46. No. 12, pp. 35–40.

Shneiderman, B. 2000. Designing trust into online experiences. Communications of the ACM, Vol. 43, No. 12, pp. 57–59.

Sony Ericsson. 2002. How to make a Web site mobile. Extending Web services with a successful mobile service. White Paper EN/LZTBP 150 105 R1A. December 2002. www.ericsson.com/mobilityworld/ [Online, cited 4.1.2005]

Sorvari, A., Jalkanen, J., Jokela, R., Black, A., Koli, K., Moberg, M. and Keinonen, T. 2004. Usability issues in utilizing context metadata in content management of mobile devices. Proceedings of the third Nordic Conference on Human-Computer Interaction. ACM Press. Pp. 357–363.

Svanæs, D. 2001. Context-aware technology: A phenomenological perspective. Human-Computer Interaction, Vol. 16, pp. 379–400.

Tamminen, S., Oulasvirta, A. Toiskallio, K. and Kankainen, A. 2004. Understanding mobile contexts. Personal and Ubiquitous Computing, 8, pp. 135–143.

Teo, T. S. H. and Pok, S. H. 2003. Adoption of the Internet and WAP-enabled phones in Singapore. Behaviour & Information Technology, Vol. 22, No. 4, pp. 281–289.

Ueda, H., Tsukamoto, M. and Nishio, S. 2000. W-MAIL: An electronic mail system for wearable computing environments. Proceedings of the 6th Annual Conference on Mobile Computing and Networking (MOBICOM 2000).

Venkatesh, V. and Davis, F. D. 2000. Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. Management Science, 46: 2, pp. 186–204.

Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. 2003. User acceptance of information technology: Toward a unified view. MIS Quarterly, Vol. 27. No. 3, September, pp. 425–478.

Väänänen-Vainio-Mattila, K. and Ruuska, S. 2000. Designing mobile phones and communicators for consumer's needs at Nokia. In: Bergman, E. (ed.). Information appliances and beyond. Interaction design for consumer products. Morgan Kaufmann Publishers. Pp. 169–204. ISBN 1-55860-600-9.

Weiss, S. 2002. Handheld Usability. Wiley. ISBN 0-470-84446-9.

Ziefle, M. and Bay, S. 2004. Mental models of a cellular phone menu. Comparing older and younger novice users. In: Brewster, S. and Dunlop, M. (eds.). Mobile Human-Computer Interaction – Mobile HCI 2004. Proceedings. Springer-Verlag. Lecture Notes in Computer Science 3160. Pp. 25–37.

Appendix III of this publication is not included in the PDF version.
Please order the printed version to get the complete publication
(http://www.vtt.fi/inf/pdf/)

Paper I

Kaasinen, E. 1999. Usability Challenges in Agent-based Services. In: Zuidweg, H., Campolargo, M., Delgado, J. and Mullery, A. (Eds.) Intelligence in Services and Networks. Lecture Notes in Computer Science 1597. Berlin Heidelberg: Springer-Verlag: Pp. 131–142.

Usability Challenges in Agent-based Services

Eija Kaasinen

VTT Information Technology
P.O.Box 1206, FIN-33101
Tampere, Finland
eija.kaasinen@vtt.fi

Abstract. Agent technology includes many possibilities to improve the usability of network services. However, agent technology also introduces new kinds of usability challenges and even threats. There is not yet very much experience of the human factor issues in agent technologies. This paper surveys current research results and analyses where agents can improve usability and where they may cause usability threats. The results are summarised as usability design recommendations for agent-based services.

1 Introduction

Our current computer environment is getting more and more complex. At the same time the applications and services are increasingly targeted at the general public. The users are a heterogeneous group, in which the abilities and level of the knowledge varies a great deal. The number of tasks to take care of and the number of issues to keep track of is growing continuously. These issues can be coped with by making the software more personalised and by letting the software take the initiative in helping us. Software agent technology provides solutions to make complex systems more usable. The agents provide the users with extra eyes, ears, hands or even brains [4] [7] [12] [16].

On the other hand, several researchers have drawn attention to the possible usability threats that agent technologies may cause. The users may have difficulties in understanding the complex functionality of the agents. The users may lose control over their autonomous agents and their mobile agents may escape from them. Who is responsible if the user loses control of her/his agent? We have to be aware of the human factor threats that the adaptation of agent technologies may include [3] [4] [5] [6] [7] [9] [11] [18].

This study has been made by the USINACTS project of the European ACTS Telecommunications Research Programme [1]. The aim of the USINACTS project is to support and train other projects in usability issues. An important part of the training is to introduce usability design guidelines to the designers of new services. USINACTS has collected and analysed information on usability issues in new technologies and paradigms like virtual reality, speech interfaces, mobile applications and so on.

In this paper I present usability design recommendations for agent-based network services. Agents are also common in user interfaces and in automation systems, but these fields have been left outside the scope of this study. I hope that this study will inspire the designers of agent-based services to give us feedback and further practical experiences of the human factor issues in agent-based services.

This paper first discusses the application areas where agents can be utilised. Then I describe some key properties of agents and describe the usability threats that these properties may cause. In many services, the agents collaborate with human beings. Agents are also often used to personalise services. These two issues are discussed in the next chapters. Error situations in agent applications are often complicated and difficult for the users to deal with. A separate chapter is devoted to error situations and how to handle them. Finally, the paper summarises the results as usability design recommendations for agent-based network services.

2 Where Can You Find Agents?

Agents are already quite common on the Web. Web agents can search for news that the user might be interested in or help the user to navigate the web [11]. Web agents can act as a filter in the retrieval of relevant information or as a matchmaker between people with common interests. A growing application area is negotiators, where agents buy and sell things according to instructions from their owners [2]. Agents can negotiate with each other, e.g. scheduling a meeting for a user group [23].

The current web applications are typically such that the mistakes, which the agent makes, are not too harmful. For instance, a newsagent may miss a piece of news that the user would have been interested in. As long as the success rate remains moderate, the user probably won't mind. However, when agents are used, for instance, in electronic commerce services, any mistakes may cause serious problems.

Different kinds of agent applications can be found in automation systems. Agents can be utilised in medical systems, e.g. monitoring different sensors, analysing the results and presenting conclusions to the medical staff. Agents can also be found in power plants, aircraft auto pilots and flight controls, network management systems and military control systems.

In these automation systems, the agents must be very reliable, especially if the agent is not only monitoring the system but also controlling the system or parts of it.

This study concentrates on agent-based network services. Automation systems as well as pure user interface agents have been left outside the scope of this study.

3 Properties and Usability Threats

Agents as a concept have raised a lot of fear, horror scenarios and extravagant claims [4] [5]. The developers of agent based services often seem to undervalue the threats and tend to emphasise the advantages of the agents from the users point of view [4] [12]. The arguments are mostly based on assumptions because there is still very little empirical evidence on how well people can understand and control their agents. Will ordinary consumers be able to program their agents in such a way that the agents would really reduce the user's work and information overload? Far more research is needed on how people experience agents [4] [12] [15].

Franklin and Graesser [8] define the properties of agents as presented in tables 1 and 2. Table 1 lists the properties that are common to all agents. Table 2 lists optional properties, which produce extensions to basic agents. I have added to the tables an extra column, where I identify the usability threat that this property may pose.

3.1 Constant Properties of Agents

Property	Meaning	Usability threat
Reactive (sensing and acting)	Responds in a timely fashion to changes in the environment.	Will it spy on me?
Autonomous	Exercises control over its own actions.	What if it won't obey me? Who is responsible?
Goal-oriented	Does not simply act in response to the environment.	How do I know how the agent will proceed?
Temporally Continuous	Is a continuously running process.	What if I lose control?

Table 1 Fixed properties of agents [8]

A software agent is *reactive*: it is able to take the initiative, e.g. to tell you something that you need to know but did not know to ask. To be able to be reactive, the agent has to be able to sense its environment and act accordingly. As well as the environment, the agent can sense the actions of the user and thus collect information on the user. The usability threat in this respect is where will the information be stored and how will it be used.

An agent is *autonomous*: it does not only obey instructions but it makes its own decisions. The usability threat here is the limit of its autonomy. The agent may decide that it knows better than the user what to do. Who will be responsible for the possible consequences?

An agent is *goal-oriented* while traditional software is more task-oriented. The usability challenge here is to keep the user informed about how the agent is proceeding. The user needs to know how things are going but is not necessarily interested in excess details.

Software agents *keep running continuously* while the user is doing other things. The user may totally forget that his/her agent is working on a certain issue. It has to be assured that the agent cannot do any harm when out of contact with its owner.

3.2 Optional Properties of Agents

Optional property	Meaning	Usability threat
Communicative	Communicates with its user and other agents.	Will it gossip about me?
Learning	Changes its behaviour based on its previous experience.	How do I know what it is able to do?
Mobile	Is able to transfer itself from one machine to another.	What if it escapes from me?
Flexible	Actions are not scripted.	Thus also unpredictable!
Character	Has a believable personality and emotional state.	Will it affect people emotionally?

Table 2 Optional properties of agents [8]

If the agent can communicate with the user, it can better adapt to the user's needs. However, as it can communicate with its owner, it may also be able to communicate with other people or other agents. It has to be assured that the agent knows what information it can and cannot communicate to others.

A learning agent learns by analysing the user's operations, receiving feedback from the user, receiving instructions from the user and asking advice from other agents [15]. In principle this improves the usability of the system. However, it may also cause usability problems if the user cannot understand the changes in the behaviour of the agent.

Mobile agents are able to travel in the network and execute their tasks in different computers while they travel. The user may lose track of his/her agent. It has to be assured that the agent cannot do any harm when out of contact with its user.

When the user population is very heterogeneous, flexibility makes it possible to adapt the user interface and the functionality to the different needs. The agent can also adapt to different environments, user terminals, networks and network conditions. As with a learning agent, this property can improve the usability of a system but it may also cause disorientation if the user cannot understand the adapted behaviour.

Agents may be personified or anthropomorphised, but this is not necessary or even common. Personified agents may be used to affect people emotionally or even to mislead the users to believe that there is a real person behind them. This kind of system violates ethical design principles.

4 User-Agent Collaboration

The user and the agent have a collaborative relationship in many applications. To successfully collaborate with an agent (or anybody else), you have to have a common goal and an agreed sequence of operations to accomplish the common goal. In addition each must be capable of performing their assigned actions and each must be committed to the overall success of the collaboration, not only their own part.

As illustrated in Figure 1, the agent has an impact on a system on three levels. The agent assists the user in the user's tasks. The agent has its own, autonomous tasks. The communication between the user and the agent adds new elements to the user interface.

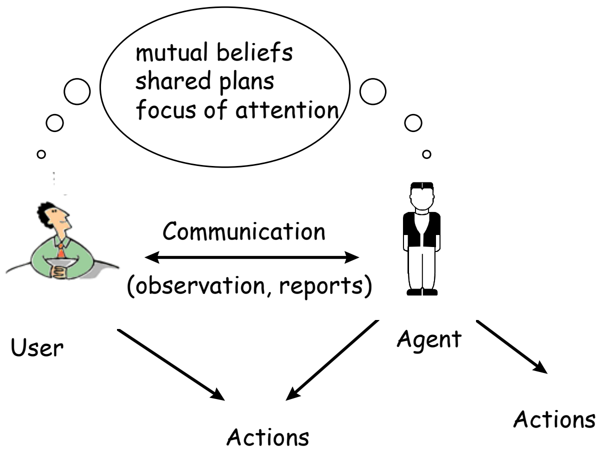


Figure 1 Collaborating with an agent [19]

The agent should be able to pick up tasks from the human when asked, and to volunteer help when appropriate. Advice should be available when the problem first appears, but the user should not be forced to interrupt her/his flow of thinking. Roth et al. [20] describe an agent as a team player:

- a team player is a good follower and supports the human in his/her role as team leader
- a team player is reliable and consistent
- a team player is a good co-ordinator
- a team player is a good communicator

In a human-agent team the roles of both the user and the agent have to be defined clearly. The user must be the team leader: s/he has the responsibility but also the power to make decisions. A proper analysis of users, their tasks and needs is essential to be able to implement agents that really meet user expectations.

The true challenge lies in the design of the right user-agent interface. The user-agent collaboration can only be successful if the user can understand what the agent does and what its limitations are and if the user can trust the agent. Although the user hands over control of some tasks to the agent, the user must feel in control or have control whenever s/he wants it. More attention should be paid to the design of the supervision of the agents and to user interfaces for programming agents [7].

4.1 User Interface of the Agent

Personified agents can be motivated by the fact that people tend to personify their computers anyway. However, personified agents strengthen the impression that the system will be able to take responsibility for its actions and that the system will act rationally, similar to another human being [3].

Agents need not be personified or anthropomorphised. Personified agents are useful in some applications but they must not confuse users and lead them to assume that there is a real person behind them. Pattie Maes and Ben Shneiderman claim that in the most successful interfaces the agents are pretty much invisible [7].

Whenever the agent gets a personality, design decisions on its characteristics and appearance should be made. Should it be a dominant expert or a humble servant? Should it resemble a human being or be more like a cartoon character? Should the characteristics and the appearance be localised to different cultural environments? Is it necessary to make the personality of the agent adaptable according to the preferences of different users?

A benefit of a personified agent is that it is able to display the status of the system by simple mimicry. For example, cartoon figures are able to express many different feelings by very simple changes in expression or appearance.

Especially in professional applications the question of whether the agents should be personified deserves careful consideration. Does the personified agent improve the usability of the system compared to a more invisible solution? Is the personality of the agent such that the users will accept it as their team-mate?

5 Personalisation of Services

Personalisation is a good way to improve the usability of complex applications. The user interface, the contents and the functionality can be adapted according to the user, his/her task, terminal, network, network conditions and so on.

Learning agents detect patterns in user actions and learn from other agents. User-programmable agents are difficult to set up and artificial intelligence based agents suffer from a lack of sufficient common-sense knowledge. These approaches can, however, be used to complement a learning agent. Pattie Maes proposes the following guidelines [2]:

- the model of the user must be inspectable by the user
- the user needs to understand the method of operation
- users need to be able to control the degree of automation and intrusiveness
- gradual automation works best

The challenge of designing adaptive services lies in finding the best way of extracting people's preferences and behaviour, and using this as a basis for adaptation. Often a dialogue between the user and the system is needed to fine-tune the user model.

Iterative design is essential when designing adaptive applications. First, we must identify a problem that can be solved by adaptivity. Then, we have to identify user characteristics related to the problem. Then, we have to identify an appropriate adaptation technique. Users' profiles, tasks and adaptations should be identified and implemented incrementally [3] [12].

5.1 User Model

The user should be allowed to inspect and alter her/his user model in the system. However, the user model is not always in such a form that the user could interpret it from various parameters and rules. Also in this case, the user should be able to predict the effects of the adaptivity, and if s/he is not able to fine-tune the model, s/he should at least be able to turn the model off [12].

”Don't diagnose what you cannot treat” is a good principle. The user model should only include the user characteristics that have an effect on the adaptation. If we can find ways by which the user can control and alter the explanations provided by the agent so that they fit in with their knowledge, this would probably be a better approach than making the system guess these aspects of user characteristics and be adaptive [12].

Most models of user knowledge are unreliable because in real life there are so many variations in user behaviour. If the adaptive system assumes too rigid and static a model of the user's plans and goals, it cannot capture the continuous improvisation that people are involved in.

Modelling and adapting to the user's preferences seems to be the most successful approach so far. Even if preferences are as hard to identify from user behaviour as the user's knowledge and habits, we can easily ask the user about her/his preferences. It is more difficult for the users to describe their knowledge or their habits [12].

An ideal user profile should include information on how the user looks, how s/he sounds, what s/he knows, what s/he is interested in and what her/his habits are [18]. However, the few commercial successes have either had a very simple knowledge of the user or have created their adaptations based on what other users do, rather than on complex models of individual users [12] [14].

6 Agents Make Mistakes – and so Do Users

It is very hard to make agents that always come up with the right answers and always do the right things. Learning agents make their decisions based on examples given by the user or by trying to analyse the behaviour of the user. It is clear that the results cannot be 100% right. In certain kinds of applications the users don't find this to be a problem; for instance, the user may find partially right search results of a web search agent to be very useful.

6.1 Who is Responsible?

Users may start to overestimate their skills when they forget that they are working with a system and not alone. Users must be aware of all limitations and be able to recognise wrong or irrelevant suggestions. Rejecting a proposal requires more knowledge than uncritically accepting it. People who do have not enough knowledge tend to accept incorrect advice [22].

When using an agent, the user gives up some control but at the same time s/he is aware of what to expect from the agent as a result. S/he does not and needs not know in detail how the agent actually carries out its tasks. The user has to be in charge, so some responsibility for reacting to warnings or questions from the agent has to be borne by the user.

6.2 User Autonomy

You cannot require the user to take responsibility for the system unless you assure her/his autonomy. Software agents can both promote and undermine user autonomy. Friedman and Nissenbaum [9] identify different aspects that affect user autonomy:

- Agent capability - user autonomy can be undermined when the software agent does not provide the user with the necessary technological capability to realise her/his goals. Usually the reason is that the user is not able to accurately represent her/his intentions with the user interface provided.
- Agent complexity - this is typical of end-user programmable agents that require a lot of insight, understanding and effort from the end-user. The user has to recognise the opportunity to employ a software agent, take the initiative to create an agent, endow the agent with explicit knowledge and maintain the agent's rules over time. Thus problems of agent complexity rise from a mismatch between the user's capabilities and the capabilities assumed by the agent.
- Knowledge about the agent - sometimes in order to use the services of an agent as desired, the user has to know how the agent goes about its task. The user needs explanations from the agent about the reasoning. The explanations must be given in the user's own language, e.g. in terms of past examples.
- Misrepresentation of the agent - users can experience a loss of autonomy when they get inaccurate information about the agent. For instance, a web agent may represent itself to the user as another human being and the user may be led to engage socially as with a real person. Once discovered, this kind of experience casts doubt on future interactions.
- Agent fluidity - the agents need to be able to support and easily accommodate changes in the user's goals. Most filtering agents don't provide access to the information that was filtered out. However, such information is essential if users are to recognise when their agents need to be reprogrammed to meet evolving goals. The users must be able to review and fine-tune their agents as their goals change.

6.3 Trust

Since the agent takes over some tasks from the user, the user has to trust the agent. If the agent asks for advice from other agents, the user should be able to trust these agents, too [12].

If an adaptive system gives wrong advice just once, the user's trust in the system will go down drastically. However, recent research results show that when the users are taken into the design process from the very beginning, they begin to understand the nature of learning agents.

Gustafson et al. [10] report experiences of the on-line production staff of a web newspaper. Despite the agent's imperfections, it quickly gained the confidence of the users. Because of their intense involvement in the software design, the users felt strong ownership of the designed system. This contributed to the satisfaction of the users and the ease with which the users adapted to the agent.

St. Amant and Dulberg [21] report on a study where the users considered even bad advice better than no advice. The advice had either a high-probability or low-probability guarantee with it. The user could immediately see how much s/he could rely on the advice.

Trust may also be a question of culture. Once we get used to having adaptive systems around us, we will gradually build mental models of how they work and how far they can be trusted.

In the design of agent-based applications, decisions on how to handle mistakes are crucial. If the application is such that mistakes can be tolerated, the agent could learn from them. If so, the user should be able to point out the mistakes to the agent so as to improve the success rate of the agent in the future.

In each case, however, the analysis of possible mistakes and their effect on the system should be an essential part of the design. If the user makes a mistake, the system should be able to prevent erroneous actions or help the user in recovering the system. If the agent encounters its limits and cannot succeed in its task anymore, it should be able to provide all the possible advice to the user so that s/he can take over the task.

7 Usability Design Recommendations for Agent Appications

The design recommendations can be classified under four main themes: agent appearance and communication, teamwork between the user and the agent, modelling the user, and error situations.

7.1 Agent Appearance and Communication

- The aim of the agent is to make things easier for the user. The agent should be kept simple so that understanding and controlling the agent does not compromise the initial aim.

- The agent should keep the user informed of how it is proceeding with its task. In the design it has to be decided how much and how detailed information is needed and whether this information should be adapted according to the user.
- Personified agents can make the system easier to understand and interact with. However, sometimes the match may get too close: the user may get the impression of a real person and expect too much from the agent.
- Consider whether the personified agent really improves the usability of the system. Agents who work seamlessly in the background are often more effective and usable than personified agents.
- Consider whether the personality of the agent is such that the user will accept it as a team-mate
- Assure that the user will understand the changes in the behaviour of the agent.
- Consider whether the characteristics and appearance of the agent needs to be localised to different cultural environments.

7.2 Teamwork between the User and the Agent

- The user has to feel and be in control, even though some of her/his tasks have been delegated to an agent.
- The user must have a clear picture of the agent's capabilities and rules of reasoning.
- The user should always have a clear picture of the whole system. He/she has to understand what is going on and why, even if it is unnecessary to understand the details of the agent and in the normal situation the user doesn't have to interfere.
- The user must be able to stop the agent safely, even though undo and redo may not be possible.
- The tasks of both the user and the agent should be defined clearly, based on a thorough task analysis.
- The user needs to understand the method of operation.
- The user needs to be able to control the degree of automation and intrusiveness.

7.3 Modelling the User

- The user should be allowed to inspect and alter her/his user model in the system.
- The user must be aware of where the user model is stored and how it is being used.
- The agent must have rules governing what information it can and cannot communicate to others.
- The user should have easy access to the information, which has been filtered away by the agent based on the user model.
- Especially learning agents are often deliberately designed so that they make conclusions that cannot be guaranteed to be right. The user will benefit if he/she can be provided with an estimation of the probability of the results.
- Personalisation can significantly improve the efficiency of use. Personalisation has to be optional and the user should be able to use the service without personalising it first.

- The aim of personalisation is to simplify the use of the application. Complicated personalisation dialogues may contradict this aim.

7.4 Error Situations

Agents work continuously and autonomously on their task. Thus the number and diversity of possible error situations is higher than in direct manipulation applications. The error situations may also be more complicated and more difficult to understand. The errors may be user errors, agent errors or unexpected events in the environment.

- The analysis of possible mistakes (by the agent or user) and their effect on the system should be an essential part of the design
- The design should include action plans both for the user and the agent for each identified error situation.
- It has to be assured that the agent cannot do any harm when out of contact with the user.

8 Conclusions

Agents provide many possibilities to improve the usability of systems by carrying out tasks on behalf of the users, searching out and filtering information for them, providing on-line assistance and so on. On the other hand, current research has revealed some threats posed by agent technologies, which the designers should be aware of.

This study suggests usability design recommendations for agent-based services. The rules deal with the appearance and communication of the agent, user-agent cooperation as a team, user modelling and error situations.

The design recommendations are not complete but I hope that they will give a good starting point for the designers of agent-based services. I also hope that this study will inspire the designers of agent-based services to give us feedback and further practical experiences of the human factor issues in agent-based services.

References

1. Advanced Communications Technology and Services - ACTS Information Window. Http://www.infowin.org/ACTS/
2. Bevan, N. HCI'96 Report. Imperial College, London. 1996. Http://info.lut.ac.uk/research/husat/inuse/hci96.html
3. Boy, G.A. Knowledge Elicitation for the Design of Software Agents. In Helander, M., Landauer, T.K. and Prabhu, p. (eds.). Handbook of Human-Computer Interaction. 1997. Pp 1203 - 1234.
4. Chester, J. (Ed.) Towards a Human Information Society. People issues in the implementation of the EU Framework V Programme. USINACTS project. 1998.

5. Bradshaw, J.M. (Ed.) Software Agents. AAAI Press / The MIT Press. 1997.
6. Clarke, A. et al. Further Human Factor Issues in ACTS. USINACTS project public report. 1997. Available at Http://www.hhi.de/usinacts.html
7. Excerpts from debates at IUI 97 and CHI 97. Ben Schneiderman and Pattie Maes : Direct Manipulation vs. Interface Agents. ACM Interactions. November + December 1997. Volume IV.6. 1997. Pp 42 - 61.
8. Franklin, S. and Graesser A. Is it an Agent or just a Program?: A Taxonomy for Autonomous Agents. Proceedings of the Third International Workshop on Agent Theories, Architectures and Languages. Springer-Verlag. 1996.
9. Friedman, B. and Nissenbaum, H. Software Agents and User Autonomy. ACM Agents '97 Conference Proceedings. 1997.
10. Gustafson, T., Schafer, J. B. and Konstan, J. Agents in their Midst: Evaluating User Adaptation to Agent-Assisted Interfaces. Proceedings of ACM Conference IUI'98. San Francisco. 1998.
11. Helmers, S., Hoffmann, U. and Stamos-Kaschke, J. (How) Can Software Agents Become Good Net Citizens? CMC Magazine. February 1997. Available at Http://www.december.com/cmc/mag/1997/feb/heldefs.html
12. Höök, K. Steps to take before Intelligent User Interfaces become real. Designing and Evaluating Intelligent User Interface. Tutorial of ACM Conference IUI'98. San Francisco. 1998. Pp 108 - 126.
13. Kanfer, A. Sweet, J. and Schlosser, A. Humanizing the Net: Social Navigation with a "Know-Who" Email Agent. Proceeding of Web Conference 1997. Available at
14. Jennings, N. R. and Wooldridge, M. J. (Eds.) Agent Technology Foundations, Applications, and Markets. Springer 1998.
15. Maes, P. Agents that Reduce Work and Information Overload. Communications of the ACM, Vol. 37, No.7, July 1994.
16. Magedanz, T. (Ed.) Proceedings of ACTS Agent Workshop. Brussels. 1998.
17. Nielsen, J. Ten Usability Heuristics. 1994. Available at Http://www.useit.com/papers/heuristic/heuristic_list.html
18. Orwant, J. For Want of a Bit the User was Lost: Cheap User Modeling. IBM Systems Journal, 35(3-4). 1996.
19. Rich, C. and Sidner C. L. COLLAGEN: When Agents Collaborate with People. Proceedings of First International Conference on Autonomous Agents. Marina del Rey, CA, 1997.
20. Roth, E.M., Malin, J.T. and Schreckenghost D.L. Paradigms for Intelligent Interface Design. In Helander, M., Landauer, T.K. and Prabhu, p. (eds.). Handbook of Human-Computer Interaction. 1997. Pp 1177 - 1201.
21. St. Amant, R. and Dulberg, M. S. An Experiment with Navigational and Intelligent Assistance. Proceedings of ACM Conference IUI'98. San Francisco. 1998.
22. Wærn, Y. and Hägglund, S. User Aspects of Knowledge-Based Systems. In Helander, M., Landauer, T.K. and Prabhu, p. (eds.). Handbook of Human-Computer Interaction. 1997. Pp 1159 - 1176.
23. What's an Agent? Http://www.ai.mit.edu/people/mhcoen/agents/
24. Foundation for Intelligent Physical Agents. Http://drogo.cselt.it/fipa/

Paper II

Kaasinen, E., Aaltonen, M. and Laakko, T. 1999. Defining User Requirements for WAP Services. In: Bullinger, H.-J. and Ziegler, J. (Eds.) Human-Computer Interaction. Communication, Cooperation and Applications Design. Volume 2 of the Proceedings of the 8th International Conference on Human-Computer Interaction. Lawrence Erlbaum Associates. Pp. 33–37.

Defining User Requirements for WAP Services

Eija Kaasinen, Matti Aaltonen, Timo Laakko
VTT Information Technology

1 Introduction

The rapid growth of web services has led to a situation where companies and individuals are increasingly reliant on material that is available on the Internet or intranets. From this it follows that people need access to the web even when they are away from home or office. Access itself will probably be the killer application for Mobile Internet (Kylänpää and Laakko 1999).

Wireless Application Protocol (WAP) and Wireless Markup Language (WML) will together form a framework in which web services can be brought to small mobile terminals. WAP will constitute an open architecture, enabling the user to access the same services from a number of different terminals via mobile networks (WAP Forum 1999). Although the international specification work on WAP is still going on (April 1999), the first WAP-compliant devices were introduced in March 1999.

This paper describes our approach to define user requirements for two kinds of WAP services. Our user group is not limited solely to the mobile users of the services. We have also considered the requirements of service providers, terminal manufacturers and application developers, because they are all needed to bring WAP services to end users.

2 Wireless Application Protocol

The key idea of WAP is that it enables the service providers to offer services for their mobile customers with various kinds of terminal devices. The WAP specification defines a framework for describing and transmitting information about the client device. The information includes device and browser software characteristics together with user preferences. In the WAP specification this information about the client is called the User Agent Profile (WAP Forum 1999).

WML provides the basic means for document formatting and user interaction, but presupposes little about how they are actually implemented. Each client device implements the interactions in its own way, depending on the device type and the WML browser software (WAP Forum 1999).

The first WAP-compliant devices have been mobile phones, but in the future WAP devices will vary a lot in size, input/output facilities, accepted content types, network connections etc.

In our project, parallel to the ongoing international specification work of WAP, we have studied in practice how the future WAP services can be designed and implemented. On the one hand, we are studying how to build mobile aware services using WML. As a case study we have developed a business card search service. On the other hand, we have developed a WAP Proxy server, which is able to convert HTML pages automatically to WML, thus providing mobile transparent services.

3 The Users

Mobile users require easy and efficient access to WAP services. Terminals may have small screen sizes, limited memory capacities, restricted input mechanisms and slow network connections. The users want the services to be able to adapt to various terminal devices in the best possible way (Clarke et al. 1997).

The service providers have to decide if they want to make a specific WAP application or to rely just on HTML/WML conversion. The conversion approach may be sufficient when mobile users are not an essential user group of the service or when the structure of the service is very straightforward.

The technical support group of our project includes not only designers and usability experts but also members from the telecommunications industry. In the support group we have a device manufacturer (Ericsson), whose interest is to assure that their future WAP devices will provide a good platform for WAP services. We also have a network operator (Radiolinja), who will provide their own and third party WAP services to their mobile clients. The third member is a software company (Teamware), which is interested in studying how to design and implement WAP-compliant services for their clients in the future.

4 The Design Process

We have adopted an iterative human-centred design approach based on the draft ISO standard ISO/DIS 13407 (ISO/DIS 1997). Our usability framework includes effectiveness, efficiency and satisfaction as defined in draft ISO standard ISO/FDIS 9241-11 (ISO/FDIS 1997). In our project we will evaluate both usefulness and usability. These two attributes predict the user acceptance of the new services (Davis et al. 1989).

We have been defining user requirements from three different perspectives: the technology itself, the mobile user and the service provider. It is essential to understand the technical possibilities and limitations that affect user requirements. The mobile users need easy and quick access to the mobile services. The service providers want to be able to offer their services to different clients at minimum additional cost.

We started our design process by defining initial user requirements. These were based on literature studies, studies of corresponding products and analyses of current mobile devices and predictions about future ones. Our design process allows us to identify new user requirements and to feed them back into the process throughout the entire life cycle of the project.

We reported the user requirements as concrete, numbered statements. In this way we got a user requirements guideline for our project. This guideline will be maintained throughout the project as new user requirements are identified and old requirements are refined as the result of the evaluation. Figure 1 illustrates our initial user requirements for the HTML/WML conversion.

7.1.5 The user should be informed about problems in conversions

- the user should get information about elements that could not be converted
- the user should get information on web pages that could not be converted

7.1.6 The user should be able to avoid excessive amounts of unwanted information

- direct search is more effective than “browse and look”
- the user would benefit from a description of content he/she is about to access.

Figure 1 A view of initial user requirements for the HTML/WML conversion

Evaluation has been a continuous activity during the requirements definition and design phases of the project. The main evaluation method has been the pluralistic walkthrough. This means meetings at which the users, application field experts, designers and usability experts together evaluate design solutions to get feedback and to generate new ideas. Towards the end of our project we will concentrate on user trials.

In the software requirements definition phase we defined the basic functionality of our services as use cases and evaluated them together with our technical support group. Figure 2 illustrates a part of a use case. We also made scenarios of HTML/WML conversion by using common web services (news, schedules, reservations etc.) as examples. The use cases and scenarios helped us in defining which of the described features and functions were to become a part of our services. The initial user requirements defined individual features of the services, whereas the use cases and scenarios described complete user tasks.

In a recent meeting Paavo has met a lady who works at Radiolinja. Unfortunately Paavo does not remember her name.

- Paavo asks for a search form from the BCSS
- He fills in the form “Org: Radiolinja”, “Date of birth: 1.1.60 or later”
- he presses “search”
- BCSS finds many ladies with these attributes
- Paavo has plenty of time and he decides to browse through the cards

Figure 2 A view of a use case of the Business Card Search Service

In the iterative design phase we have implemented prototypes of our services on WAP emulators by Unwired Planet (UP.Simulator 3.1, www.unwiredplanet.com), Nokia (WAP SDK version 1.01, www.nokia.com) and Dynamical Systems Research Ltd. (Wireless Application Reader version 1.1, www.wap.net). Unwired Planet and Nokia offer cellular phone type client emulators and Dynamical Systems Research’s emulator has in addition a palmtop type client emulator. We have evaluated the prototypes in pluralistic walkthrough meetings using the use cases as test tasks.

Figure 3 Business Card Search Service on Nokia and Unwired Planet platforms

In practice it has not been easy to make an application look good and be easy to use on different types of devices. The platforms vary by the screen size, the handling of soft buttons and the way links and document elements are visualised. If you try to improve the usability of the application on one target platform, you may generate usability problems on another platform.

5 Conclusions

In defining the user requirements for WAP services, we have considered the points of view of mobile end users, service providers, device manufacturers and application developers. We started our design process by defining the initial user requirements and use cases. In this kind of continuously changing environment it is essential to be able to refine the user requirements throughout the design process. It is also important to record the design rationale behind the design decisions. In this way we can easily return to rejected requirements, for instance when a new WAP device enables us to implement new features in our services.

So far we have only been able to test our services on simulated WAP terminals. It will be interesting to see how our services behave on the first real WAP terminals.

6 References

Clarke, A., Carter, C., Kaasinen, E., Kantola, K., Chester, J., Concejero, P., Buss, R., Mühlbach, L. and Ruschin, D. Further Human Factors Issues in ACTS. USINACTS project public report. 1997. Pp 20-27. Available at http://www.hhi.de/usinacts.html

Davis, F., Bagozzi, R. and Warshaw, P. User Acceptance of Computer Technology; a Comparison of Two Theoretical Models. Management Science. Volume 35. No. 8. August 1989. Pp 982-1003.

ISO/DIS 13407. Human-centred Design Processes for Interactive Systems. Draft international standard. 1997.

ISO/FDIS 9241-11. Ergonomic Requirements for Office Work with Visual Display Terminals - Guidance on Usability. Draft international standard. 1997.

Kylänpää, M. and Laakko, T. Adapting Content To Mobile Terminals: Examining Two Approaches, Third Generation Mobile Systems in Europe, London, 25.-27.1.1999.

WAP Forum. Wireless Application Protocol. Wireless Application Environment Overview. http://www.wapforum.org/. 1999.

Paper IV

Kaasinen, E., Kasesniemi, E.-L, Kolari J., Suihkonen, R. and Laakko, T. 2001. Mobile-Transparent Access to Web Services – Acceptance of Users and Service Providers. Proceedings of International Symposium on Human Factors in Telecommunication. Bergen, Norway, 5.–7.11.2001.

Mobile-Transparent Access to Web Services – Acceptance of Users and Service Providers

Eija Kaasinen, Eija-Liisa Kasesniemi, Juha Kolari, Raisa Suihkonen and Timo Laakko

VTT Information Technology, P.O. Box 1206, FIN-33101 Tampere, Finland

eija.kaasinen@vtt.fi

Abstract

The number and variety of mobile Internet devices is growing rapidly. It is no longer enough to design web services for generic mobile access; the services also have to adapt to different contexts of use. We have studied an approach where the users of different WAP phones can transparently access the same web services as they use on their PCs. We have evaluated the user acceptance of this solution in field trials with users and in interviews with service providers. The results suggest that there is a need for cross-media services that the users can access with different devices depending on their current context of use. Mobile devices are above all personal communication devices. More attention should be paid to providing personally selected contents and new kinds of interactive services where the users can participate and communicate with other people.

Key words: web, WAP, mobile access, usability, user evaluation

1 Mobile Access to the Internet

Today, mobile Internet devices range from phones to communicators, palmtop computers and portable computers. The main problems with the mobile Internet include slow connections, small screen, limited input techniques and limited functionality of the browser. Current approaches to provide markup language based services to small mobile devices are based on markup languages specially designed for this purpose. Some approaches like i-mode provide limited access to the web whereas Wireless Application Protocol (WAP) services are designed as separate applications and little if any attention has been paid to their co-existence with a corresponding web service (Hjelm, 2000).

From the user's point of view, the mobile Internet cannot be an entity in its own right. Many web services are such that the users will need to access them anywhere and with whatever device they have available. These services should be transparently available to the mobile users (Kaasinen et al., 2001). The need for device-transparent services has been recognised in the World Wide Web Consortium (W3C). In spring 2001, W3C started Device Independence Activity that is working towards seamless web access and authoring (W3C, 2001). The W3C specification for modular XHTML (eXtensible Hypertext Markup Language) is a concrete step towards a device-transparent web.

In his Alertbox, Jakob Nielsen (1999) gives a list of principles that should be observed when offering the same service with multiple interfaces. It is important that the different interfaces feel like variations of a single system, even though they have different designs.

Consistency is a commonly accepted but sometimes problematic design goal. In the usability design of multi-platform web services the question of whether the site should look the same on all platforms or whether it should be changed to make use of the idiosyncrasies of a particular environment must be considered. (Ketola et al., 2000)

2 The WAP Proxy Server for Mobile-Transparent Web Access

The WAP services are implemented using the XML-based Wireless Markup Language (WML). WML is designed for low-end devices and for slow, unreliable networks. WML provides the basic means for document formatting and user interaction, but presupposes little about how they are actually implemented (WAP Forum, 1999). There are several different WAP browsers on the market, each with its own user interface philosophy and usability guidelines (Nokia, 2001) (Openwave, 2001).

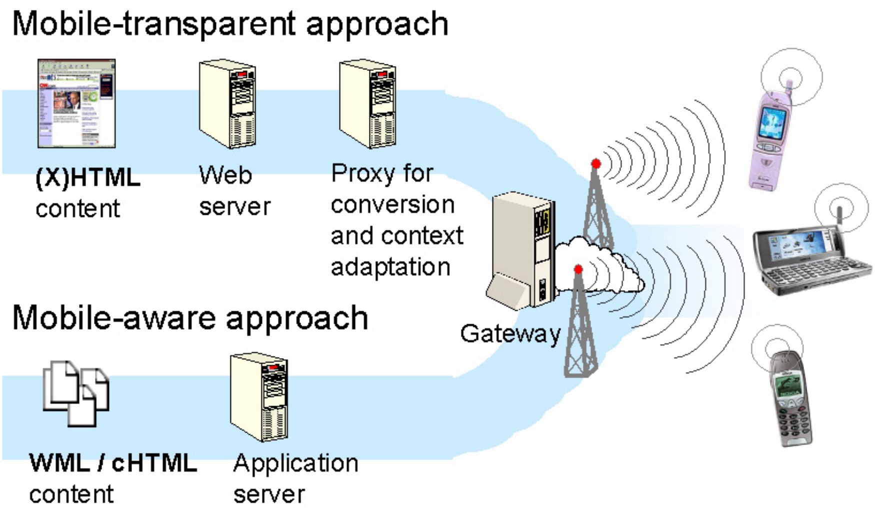


Figure 1. Mobile-transparent access to the web via the (X)HTML/WML conversion proxy

Our WAP conversion proxy server facilitates mobile-transparent access to (X)HTML-based web services for different WAP devices, as illustrated in Figure 1. The conversion proxy converts HTML content to WML decks. The proxy server also aims to format the service according to the capabilities and preferences of the mobile client, as defined by the User Agent Profile (UAProf) of the device. Images are converted to a form that the client device supports and image sizes are scaled according to the client. The proxy can also perform state management on behalf of the client, enabling access to web sites that require user login with cookies. One HTML document typically produces several WML decks, depending on the memory limits of the client device (Kaasinen et al., 2000).

Mobile Access Guidelines by W3C (1999) can be utilised in designing convertible sites. Our WAP conversion proxy could correctly convert most of the sites that were designed according to these guidelines (Kaasinen et al., 2001).

3 User Acceptance

We evaluated mobile-transparent access to the web in two field trials, the first one in summer 2000 and the second one in summer 2001. Both trials lasted for two months and included 40 users, aged 14-79 years and representing evenly both sexes. The experience of the users with

GSM phones, the Internet and computers varied from novice to professional. In the 2000 trial we had three types of WAP phones as user devices. In the 2001 trial the user devices included ten types of WAP phones as well as PDAs and communicators.

In the first trial we built a tailor-made WAP front page for each test user according to his/her interests. From this front page the user had access to both mobile-aware WAP services (Radiolinja WAP portal) and converted web services. In addition to the personally selected web links, each user had access to the web portal of MTV3 media company and the teletext services by two Finnish TV companies (MTV3 and YLE). MTV3 web portal was redesigned for the field trial to better support the WAP conversion. The main change was to provide a separate, simplified front page for mobile users and paying more attention to the selection of link texts. Most sub pages did not require any modification for mobile use.

Even though the users were happy with the tailor-made front pages in the 2000 trial, some of them said that they would prefer to personalise the front page themselves. In the 2001 trial all the users had a similar front page with a link to their personal page. They could freely personalise the personal page with a tool, Link editor, which was available both in the web and in the WAP. In addition to the personally selected web links, the users could write their own text notes and link them to the personal page. The front page of each user was accessed with a unique URL, so that the users did not have to log in to the page.

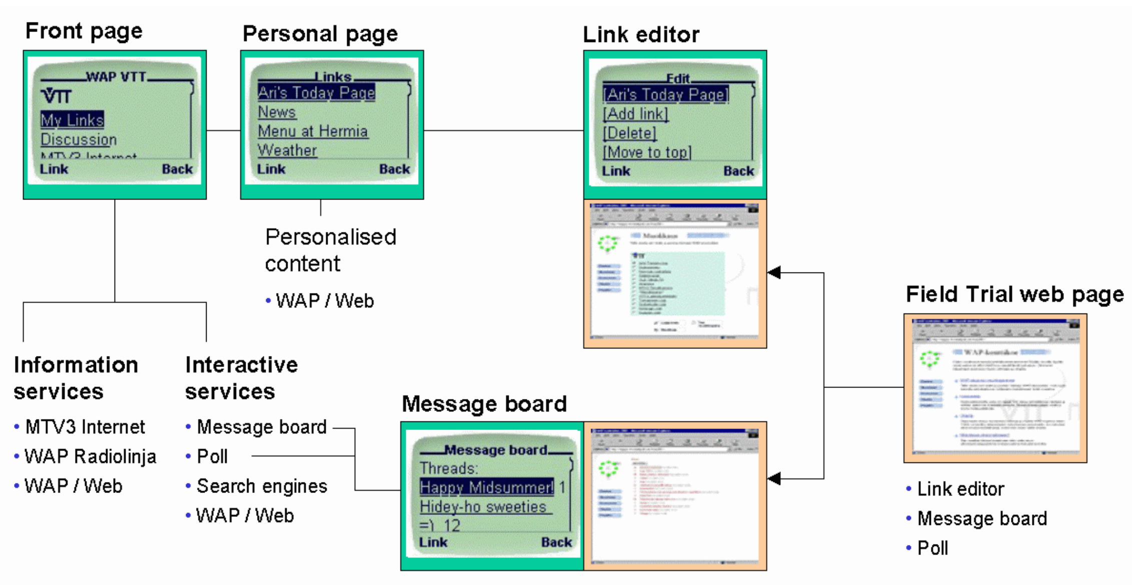


Figure 2 The set-up of field trial 2001 as seen by the user (screen views translated, original service in Finnish)

In the second trial we studied group communication with a Message board service implemented by Teamware Interactive. The Message board service was set up for four different user groups: bird watchers, choir members, teens and expecting mothers. Each group was made up of five members. In addition to the four small groups, all the test users could access a common chat group "Pulinapalju" ("Chatterbox"). The other common services in the 2001 trial included converted MTV3 portal, a poll service, web search engine (Google), WAP search engine (Evreka) and Radiolinja WAP portal (Figure 2). In both trials, the test users were asked to visit their WAP page at least twice a week. Otherwise the users could use the

WAP freely, up to two hours a day. Data collection took the form of thematic interviews, observation and log files.

Most usage in the field trial was immobile use. The users were quite often filling in empty moments with WAP services, e.g., while waiting for a friend, sitting on the bus, or taking a break at work. It is also notable that a lot of the usage took place at home. The users could have got the same information from their PCs or teletext, but the WAP phone was more readily at hand. By using the mobile service, the user did not have to move him/herself.

3.1 Access to Information

The usability problems with the converted content were very similar to the problems with mobile-aware WAP services. The main problems were related to navigation, scrolling on the page, slow response times as well as device and network hang-ups. The users felt they needed an overview of the available information. They could read long articles if they found them interesting, but a quick way to browse through less interesting information was lacking.

Generic WAP services by the teleoperator had the fewest usability problems and they were also technically the most reliable ones. The users often thought that the content of these services was too limited. Some users, however, found favourites among these services and started to use them regularly. Typical favourite services were news and weather forecasts.

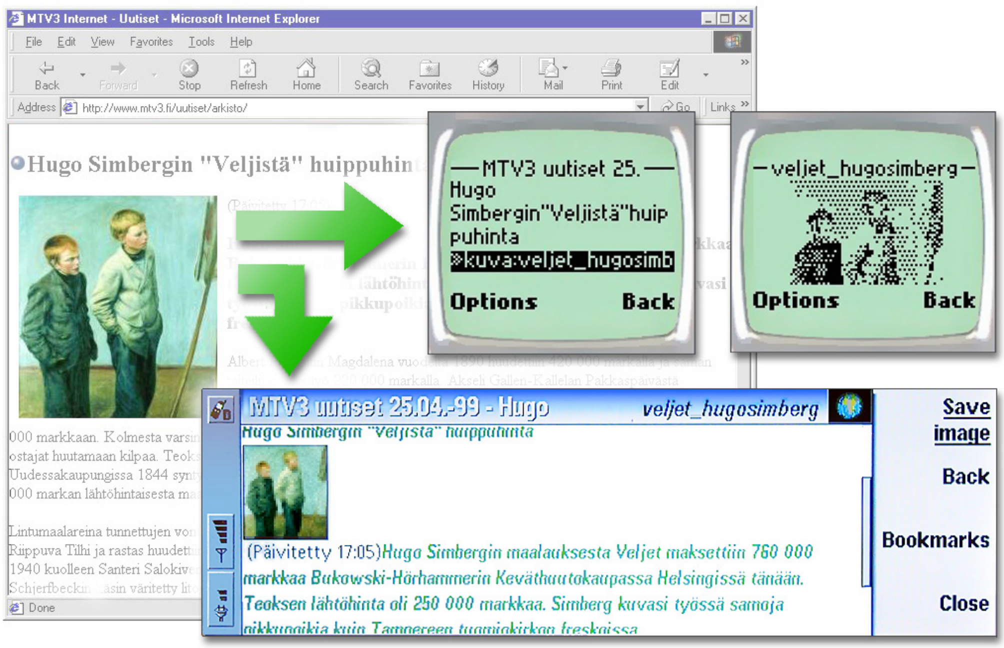


Figure 3 The conversion proxy adapted images according to the client device

MTV3 services were quite familiar to most of the users. WAP access worked very well and some of the test users even stated that “the WAP version is much easier to use than the version on the web”. The portal with its sub-pages was the most popular service in both trials. During the 2001 trial, the MTV3 main page was accessed over 3800 times. The most popular sub-pages included news (541), TV schedule (267), page of a popular TV show "Far Out" (206) and sports (110). The users commented positively on the converted images on the pages but they seldom saw the images as being absolutely necessary (Figure 3).

Teletext services were familiar also to the users who were not familiar with the web. The teletext content worked well on WAP and the users liked the wide range of material available. Familiarity with the service seemed to ease navigation.

In addition to MTV3 news, other news services were also popular. During the 2001 trial, the newspaper Helsingin Sanomat was accessed 551 times, the afternoon tabloid Iltalehti 180 times and the YLE's teletext service 210 times. It seems that the users did not want just any news, but the news from their favourite service provider.

Several users were excited about the personal links. This kind of content could include home pages of the family members, web pages of hobby groups, web pages of the user's own university and similar personal things.

In the 2001 trial, Google(web) and Evreka(WAP) search services were put to good use by quite a few users. The users were surprised to see their WAP searches gain lots of items. The traditional WAP portals had hidden the fact that a surprisingly large number of WAP services are already available. Unfortunately, quite a few WAP services failed to define the title of the page, thus giving the searcher no clue of the contents of the page.

3.2 Personal Interactive Services

Some users actively updated their personal pages. Most users, however, found using the Link editor too laborious or did not have enough motivation to do the personalisation. In the interviews, the users said that they would prefer giving themes that they are interested in, rather than web URLs. If they then could get a list of alternative links, they could select from these. Also, they would have liked to be informed of new and interesting services, based on their expressed interests.

Personal communication clearly activated the users in a way that was not possible with mere information services. Even novice users associated e-mail strongly with WAP. For the trial, we tried to ensure that all the users who wanted to have e-mail access could get it.

The discussion group of Vauva magazine – a magazine for the parents of newborn babies - was quite popular (accessed 774 times), taking into account that it was probably mainly accessed by the small group of expecting mothers. Some services that were used only by a single user were surprisingly high on the usage log. An e-mail service that turned out to convert quite nicely was accessed 254 times by a single user, and a game called "Planetarion" was accessed 191 times.

The five most active users in the 2001 trial accessed their front pages 445, 422, 332, 253 and 219 times. Most users accessed their front page around 100 times, i.e. about once a day.

With the Message board service, long user input caused technical problems in quite many WAP devices. The browsers easily hang up when the user was in the middle of writing a message. Usually the user then lost the message that (s)he had been writing. In some devices, the input fields were too short on the screen. The user had to finish the input seeing only the last few characters of the message. It seems that active user input has not been adequately considered when designing WAP browsers.

The problems with long input fields clearly reduced the use of the Message board. The teens were the only group that used the Message board actively during the trial in spite of the technical problems. The teen group soon became differentiated into communication pairs and they wanted to have new discussion groups accordingly. One of the new groups – two girls – wrote over 1200 messages during the trial.

The birdwatchers would have liked to report their observations to a wider audience. They would also have liked to get an alert when new messages arrived. The choir had somewhat similar needs – the service would have been more useful if all the members of the choir could have participated. However, mobile access to the web page of the choir, where the members could see forthcoming events, was commented on positively.

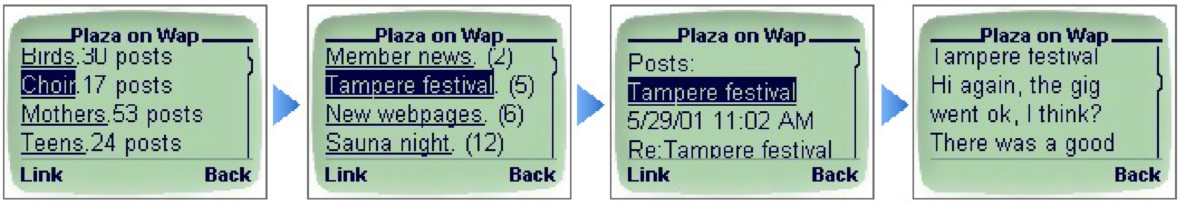


Figure 4 The Message board on a WAP phone (translated, original service in Finnish)

The experiences with the teen group suggest that for a small user group the Message board service can be truly personal – enabling the users to keep in touch with their friends and not only to retrieve information but also to participate actively. Such services support the mobile devices as personal communication devices.

Existing groups often need better access to the services that they already have in use, rather than totally new services. The whole group should be able to use the service, and each one should be able to access the service with any device that is currently available to him/her.

Some of the test users utilised the personal notes widely. The notes were e.g., memos, to-do lists, shopping lists or even self-written poems. This kind of usage is a starting point for using the mobile device as an access and storage point for all kinds of personal data. The user should be able to access this kind of personal "data chest" with any web/WAP capable device.

4 The Point of View of the Service Providers

4.1 Interviews in Spring 2000

We studied the acceptance of service providers by interviewing the representatives of current Internet service providers and probable future service providers, altogether 25 organisations in Finland. The organisations in the interview study included five public-sector organisations, three non-profit organisations, nine large enterprises and eight small or medium-size enterprises. We interviewed the managers responsible for the WWW-services in qualitative non-structured theme interviews in spring 2000. The aim was to study how the organisations plan to serve their mobile customers in the future and how the WAP conversion proxy would fit in with their plans.

All the companies were providing web-services for consumers, and half of the companies mobile services as well. The first WAP services were provided on the portals of the teleoperators in a "walled garden". This was still mostly the situation, although some alternative WAP portals had come on the market. The service providers said that they could not commit to the portal of a single teleoperator since their customers use different teleoperators. They also thought that having the teleoperator between them and their customers would hinder their contacts with the customers and make their brand less visible.

The majority of the service providers had not thought about how to adapt their services to different mobile devices. They were relying on device manufacturers to agree on suitable standards. Most service providers thought that the markets would quite soon demand mobile services. They considered the concept of an automatic conversion and adaptation proxy server useful, but mainly in free services. For commercial services the approach was seen as being too indefinite.

4.2 Interviews in Spring 2001

In the second round of service provider interviews we demonstrated the conversion proxy and interviewed 11 service providers in spring 2001. The interviewees included four public-sector organisations, five companies and two public utility associations. The interviewed people were responsible for their own organisation's web and mobile service production. All but one of the organisations had already participated in the first round of interviews.

This time the attitudes of the interviewees towards mobile services were a bit more reserved. The representatives of the companies said that they would be ready to provide mobile services as soon as they could see a realistic business model in the field. Before that new network technologies and devices will be needed on the market.

All the interviewees were curious about the conversion proxy approach. The small enterprises, public-sector organisations and public utility organisations regarded the conversion proxy approach as quite attractive, because they thought that they would not be able to provide separate mobile services in the near future. The other interviewees thought that they could provide some of their services via the conversion proxy, mainly services that they cannot charge for. Some interviewees thought that the conversion proxy could be useful in providing Intranet services for their mobile employees. Seven organisations said that they might consider redesigning some of their services to better support the conversion.

5 Conclusions

The results of this study highlight the fact that mobile devices are above all personal communication devices. The WAP Proxy conversion brought WAP to a more personal level. With more options, users began to shift towards services that were personal and/or interesting to them. Also with general services like news, people used the possibility to select their service provider. More attention should be paid in providing users with easy-to-use and effective personalisation tools.

The implementation of WAP communication services was a vitalising factor in our trial. Personal communication services seem to be somewhat underrated, considering the way in which WAP devices and services are currently designed and marketed. Self-generated content and peer-to-peer communication should be seen as a more integral part of mobile devices.

Mobile devices are usually secondary access points to the Internet for the users. That is why access to existing services may be more useful than new services designed specifically for the mobile environment. Future services will be designed for device independence from the very beginning. Meanwhile, the service providers regarded the conversion proxy as a good approach to providing access to the web services also for mobile customers and employees.

6 Acknowledgements

Our research work was carried out with the financial support of the National Technology Agency of Finland (Tekes). The companies Alma Media, Teamware Interactive, Nokia and Radiolinja also supported our work both financially and technically. The authors would like to thank these partners for the fruitful co-operation.

References

Hjelm, J. 2000. Designing Wireless Information Services. John Wiley & Sons, Inc.

Kaasinen, E., Aaltonen, M., Kolari, J., Melakoski, S. and Laakko, T. 2000. Two Approaches to Bringing Internet Services to WAP Devices, Computer Networks 33 (2000), p. 231-246.

Kaasinen, E., Kolari, J. and Laakko, T. 2001. Mobile-Transparent Access to Web Services. In Hirose, M. (ed.) Human-Computer Interaction Interact '01. IOS Press. p. 719-720.

Ketola P., Hjelmeroos H. and Räihä K.-J. 2000. Coping with Consistency under Multiple Design Constraints: The Case of the Nokia 9000 WWW Browser. Personal Technologies, Volume 4 (2&3), June 2000, p. 86-95.

Nielsen, J. 1999. Graceful Degradation of Scalable Internet Services, Alertbox, October 31, 1999. [online, cited 6 September 2001] www.useit.com/alertbox/991031.html

Nokia Mobile Phones. 2001. WAP Service Developer's Guides [several online documents for different Nokia WAP devices, cited 6 September 2001]. Forum Nokia. www.nokia.com

Openwave. 2001. Genie Application Style Guide for Openwave™, Nokia™ Model 7110™, Model 6210/6250™, anid Mitsubishi™ Trium™ WAP™ browsers. Release 1.0. February 2001. [online, cited 6 September 2001] http://developer.openwave.com/support/techlib.html#styleguides

W3C. 1999. HTML 4.0 Guidelines for Mobile Access. W3C Note, 15 March 1999, The World Wide Web Consortium. [online, cited 6 September 2001] http://www.w3.org/TR/NOTE-html40-mobile/

W3C. 2001. Device Independence Activity Working Group. The World Wide Web Consortium. [online, cited 6 September 2001] http://www.w3.org/2001/di/

WAP Forum. 1999. Wireless Application Protocol. Wireless Application Environment Overview. Version 1.3, 29-March-2000. [online, cited 6 September 2001] http://www.wapforum.org/what/technical_1_2_1.htm

Paper V

Kaasinen, E. 2003. User needs for location-aware mobile services. Journal of Personal and Ubiquitous Computing (6), pp. 70–79. Springer-Verlag London Limited.

Pers Ubiquit Comput (2003) 7: 70–79
DOI 10.1007/s00779-002-0214-7

ORIGINAL ARTICLE

Eija Kaasinen

User needs for location-aware mobile services

Received: 1 August 2002 / Accepted: 15 November 2002

Abstract Mobile contexts of use vary a lot, and may even be continuously changing during use. The context is much more than location, but its other elements are still difficult to identify or measure. Location information is becoming an integral part of different mobile devices. Current mobile services can be enhanced with location-aware features, thus providing the user with a smooth transition towards context-aware services. Potential application fields can be found in areas such as travel information, shopping, entertainment, event information and different mobile professions. This paper studies location-aware mobile services from the user's point of view. The paper draws conclusions about key issues related to user needs, based on user interviews, laboratory and field evaluations with users, and expert evaluations of location-aware services. The user needs are presented under five main themes: topical and comprehensive contents, smooth user interaction, personal and user-generated contents, seamless service entities and privacy issues.

Keywords Location-aware services · Mobile services · Usability · User evaluation · User needs

1 Introduction

In mobile environments, all the elements of the context of use may vary a lot. Users are different and they may use the services for many different tasks, even for tasks that were not anticipated in the design. The variety of mobile devices is growing and the users expect to be able to use the same or the same kind of services on the different devices. The technical and service infrastructure may differ and they may even change in the middle of a usage session, e.g. the network or the positioning system may change when the user moves from one place to another. Similarly, the service infrastructure, i.e. the available services and applications, may change. The physical context may vary a lot in terms of illumination, background noise, temperature and weather. The use of the device may affect the social situation in which the user finds him/herself or the social situation may affect the way the user uses the system.

An efficient way of improving the usability of mobile services and applications is to adapt the contents and presentation of the service to each individual user and his/her current context of use. In this way, the amount of user interaction will be minimised: the user has quick access to the information or service that (s)he needs in his/her current context of use. The information can even be provided to the user automatically.

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task [1]. The main problem with context adaptation is that the context cannot be easily identified or measured. The location of the user is an element of the context that currently can be measured more or less accurately depending on the positioning system in use. In this paper, location-aware services are defined as context-aware services that utilise the location of the user to adapt the service accordingly. Location-based services are services that are related as such or by their information contents to certain places or locations. Thus location-aware services are a special case of location-based services.

So far, context-awareness has mainly been studied from the technical point of view and the studies have concentrated on location. Different experimental systems have been set up but only a few user evaluation results from small-scale trials are available. Location-aware services are a concrete step towards context-awareness. Other aspects of context-awareness will follow as soon as the corresponding elements of the context, such as weather or the social situation, can be measured and the adaptivity needs can be identified.

Eija Kaasinen
VTT Information Technology,
P.O. Box 1206,
FIN-33101 Tampere, Finland
e-mail: eija.kaasinen@vtt.fi

This paper studies location-aware mobile services from the user's point of view. The paper draws conclusions about key issues related to user needs, based on user interviews, laboratory and field evaluations with users, and expert evaluations of location-aware services.

2 Methods of locating the user

From the point of view of the service, the simplest method of locating the user is to let him/her tell the location. From the point of view of the user, this method requires extra effort because the user needs to define his/her location and input it to the system as a part of the search.

The user can be located with different positioning systems. If the user device includes a GPS (Global Positioning System) module, the user's location can be defined very accurately (2–20 meters). A GPS cannot be used indoors and it may not work in 'urban canyons' either. The location is calculated in the user device and it has to be sent to the service provider in order to get location-aware services. The range of commercial products currently available include mobile phones with integrated GPS modules, separate GPS modules for PDAs (Personal Digital Assistant), and GPS devices with integrated mobile phone and data features.

A mobile phone can be located by the telecom operator in the network. The positioning is based on identifying the mobile network cell in which the phone is located, or on measuring distances to overlapping cells. In urban areas the accuracy can be down to 50 meters, whereas in rural areas the accuracy may be several kilometres. The advantage of the cell-based positioning method is that the user needs no extra equipment - an ordinary mobile phone will do. If the user wants to use location-aware services from other service providers, the location has to be transferred to the other service provider and the telecom operator must get permission for this from the user. The location data is possessed by the telecom operator, which may not be willing to pass it on free of charge. Possibly because of these data transaction needs, current cell-location-based services are provided mainly by telecom operators.

The user can also be identified at a service point, utilising e.g. WLAN (Wireless Local Area Network), Bluetooth™ or infrared technologies. These kinds of proximity positioning systems require a dense network of access points. The density of the network depends both on the required location accuracy and on the range of the access points. The accuracy can be down to 2 meters. The user needs special equipment, although WLAN and Bluetooth, for instance, are becoming increasingly common in current mobile devices. Because of the required infrastructure, such systems can only be used in a predefined area, e.g. a shopping centre, an exhibition area or an office building. The location of the user is available only when the user is in the service area.

3 Recent research on context- and location-aware services

Context-awareness can be implemented as an adaptation of the user interface or the contents of the service. Services can also be invoked based on the identified context.

A context-aware user interface can select the appropriate modes for service interaction. A context-aware user interface can also be implemented e.g., as context-aware text prediction [2] or a location-aware remote control for the environment [3]. A major challenge for the context-aware user interface is that the context may be continuously changing. This raises the problem of integrating changes into the user interface in such a way that the user remains in control [4]. Moving can also be seen as one mode of interaction with the system. This interaction mode is quite challenging because it is difficult to know the user's intention: is the moving really taking place in order to interact with the system [5].

In recent research, context-aware contents have been studied in different application areas, e.g. tourist guidance [6], exhibition guidance [7], e-mail [8], shopping [9], mobile network administration, medical care and office visitor information [10]. In these studies, the location of the user is the main attribute used in the context-adaptation. In well-defined application areas, it is possible to predict the other elements of the context according to the location of the user. Designing for more general user groups and wider contexts of use will be much more challenging.

4 The empirical studies

4.1 Scenario evaluations

We have carried out several empirical studies to study user attitudes, needs and preferences for location-aware services. We started with scenario evaluations in group interviews. The aim of this evaluation was to study broadly the attitudes of the potential users towards different personal navigation services. Future possibilities of personal navigation products and services were presented to the interviewees as pictured scenarios of everyday life. In addition to location-based services, the scenarios also introduced route guidance services, services for tracking property and services based on locating other people.

We had 13 evaluation groups, each with 3 to 7 people, totalling 55 persons of different ages, different backgrounds and from different parts of Finland (Table 1). The groups were selected so that they broadly represented the potential users of personal navigation services. The groups included four families, three hobby groups (football players, boaters and hunters), two youth groups, senior citizens, a group of motor-disabled people, a group of visually impaired people and students of well-being technology. The groups were somewhat

Table 1 User groups in the scenario evaluations

Group	Female	Male	Ages	Mobile phone owners (%)	Regular computer users (%)	Have been using navigation device (%)	Locality
Family 1	1	1	36–39	50	100	–	Urban, Northern Finland
Family 2	1	1	34–36	100	100	–	Rural, Southern Finland
Family 3	2	1	14–45	100	67	33	Rural, Western Finland
Family 4	1	1	15–46	50	100	–	Urban, Northern Finland
Football club	–	3	25–27	100	100	–	Urban, Western Finland
Sailing club	1	6	32–55	100	70	70	Rural, Western Finland
Hunting club	–	4	38–44	75	50	75	Rural, Northern Finland
Youth group 1	3	2	17	80	80	20	Rural, Northern Finland
Youth group 2	–	5	15–16	100	60	20	Rural, Western Finland
Senior citizens club	2	3	56–66	100	80	20	Urban, Western Finland
Disabled people at a service house	2	3	24–50	60	60	20	Urban, Western Finland
Visually impaired	2	3	27–46	100	80	20	Urban, Western Finland
Students of well-being technology	4	3	21–37	100	100	29	Urban, Northern Finland
Total	19	36	14–66	90	78	29	

male-oriented, mainly because of the selected hobby groups. We wanted to include boaters and hunters because these groups are already familiar with navigation devices and thus might be early adopters of new personal navigation services.

Each group evaluated between three and five scenarios that were selected so that they presented the different aspects of personal navigation and were targeted according to the group.

The original scenarios were written in 1999 by a multidisciplinary team of experts, as a part of setting up the Personal Navigation Research and Development Programme in Finland [11]. We modified the scenarios so that they reflected the present situation in the research and industrial fields, and so that they covered different aspects of personal navigation. We also wrote some brand new scenarios, targeted specially at different age and/or hobby groups.

The scenarios described location-aware advertising in the form of junk mail, a visit to an exhibition, different holiday and working trips, meeting friends in the evening, going to work and shopping. The scenarios were short stories of everyday life, illustrated with pictures of the context of use and imaginary mobile devices and services. Figure 1 presents a part of the scenario of location-aware spam.

The scenarios were delivered to most of the groups in advance, so that the participants could read the scenarios before the interview. In the semi-structured group interview, the scenarios were discussed one at a time. The evaluators presented the scenario to the group and started the discussion by asking the interviewees how credible they considered the scenario, and why. Figure 2 illustrates the group interview with the senior citizens.

4.2 User evaluations

We have also evaluated with users different commercial location-aware services in Finland. The aim of these evaluations was to identify good solutions in current services as well as user needs for future services. Benefon

Lauri was waiting at the traffic lights when the phone alerted. Well, again commercial messages, this time from the shoe shop near by. Just after Lauri had put the phone into his pocket, another alert was coming. This time he was tempted to have a look at a brand new internet-connected toaster at the household appliance store on the other side of the street.

Fig. 1 Part of a scenario describing location-aware spam (translated from Finnish)

Fig. 2 Group interview with senior citizens

Esc! (Fig. 4) is a mobile phone that includes an integrated GPS system. When the Benefon Esc! is used together with a Yellow Pages short message service (SMS), the user can get information on nearby services as well as their location, which the Benefon Esc! can display on the map screen. The Benefon Esc! was evaluated outdoors in the city of Tampere with six male users: three students and three middle-aged men (aged 59–62) having fishing and hunting as their hobby. We ended up with a male user group because of the hobby group selected.

In Finland, Sonera Pointer (Fig. 4) was one of the first location-aware WAP (Wireless Application Protocol) services that utilised cell-based positioning. Pointer Bensa (Gasoline) gave information on the cheapest gasoline stations in the vicinity of the user. Pointer Opas (Guide) offered information about the district around the user: sightseeing and attractions, events, activities, accommodation and tourist information. Pointer Fakta (Facts) gave statistical facts about the city or municipality in which the user was located. The Pointer services were evaluated in our usability laboratory with five users, three women and two men, aged 25–64, and with different backgrounds. Only one of the users had previous WAP experience. One of the users did not own a mobile phone himself.

The evaluations of both Benefon Esc! and Sonera Pointer were carried out in four phases. The evaluation started with filling in a background form to collect user data and information on his/her experience, e.g. with mobile phones and services. Then the user carried out test tasks that were combined with a frame story (“You are driving from Helsinki to Jyväskylä. Near Tampere you notice that you would need to fill up your car...”). After the test tasks, the user could use the system freely, if there was time left and if the user wanted. Finally, the user was interviewed about his/her general impressions of the evaluated service. One evaluation session took about two hours. The laboratory evaluations were recorded on video, whereas the outdoor evaluation was recorded on audio and as photographs of usage situations. Figure 3 illustrates a user evaluation of Benefon Esc!.

In addition, we evaluated with five users a traditional GPS device (Garmin GPS 12) and a GPS module on a PDA (Magellan GPS Companion for Palm). These devices and the map services on them may not be considered as location-aware services as such, but still they are important support services for the users of location-aware services.

4.3 Expert evaluations

We also reviewed some services in expert evaluations because these services were not available in Finland. CeBIT 2001 Mobile Fair Guide (Fig. 4) was a tradeshow guide for PDA computers. The software together with the Fair Catalogue could be downloaded from the web.

Fig. 3 Evaluating Benefon Esc! with a user

Fig. 4 Some of the evaluated devices and services: the Benefon Esc! GPS phone, Sonera Pointer on a WAP phone, Cebit2001 Guide and Vindigo on a PDA

During the fair, there was also a demonstration on using the guide together with a Bluetooth-based indoor navigation system. Pocket Streetmap (www.pocketstreetmap.com) is a map software for PDAs, which also includes some location-based services. Vindigo (Fig. 4, www.vindigo.com) is a location-aware service guide for local entertainment, shops and restaurants in major cities in the USA and in London. We reviewed the Cebit 2001 Guide at the trade fair in Hanover, whereas Pocket Streetmap and Vindigo were reviewed in London.

The group interviews were carried out in spring 2001 and were followed by the user and expert evaluations from summer 2001 to spring 2002. In the following section, the results of all the above-mentioned evaluation activities are discussed as key issues to consider when defining user needs for location-aware services.

5 User needs for location-aware services

This section presents our main findings about user needs of location-aware services. The findings are grouped under five main themes: contents, interaction, personalisation, service entities and privacy. As a starting point, Sect. 5.1 presents an overview of user attitudes identified in the group interviews and evaluations.

5.1 User attitudes

The attitudes of the users towards location-aware services were quite positive both in the group interviews and in the user evaluations. Location-aware information was expected to be especially useful in special situations, e.g. in unfamiliar environments, when looking for a specific service or in emergency situations. Location-aware information on parking lots, detours and contemporary events such as concerts and football matches was found useful. The disabled users would appreciate accessibility information, and the visually impaired pointed out the importance of speech-based systems.

However, criticism of new technology was brought up in many group interviews. A predestined and over-controlled environment was seen as dubious, and the interviewees did not accept the rational and purpose-oriented attitude to life that they identified in the scenarios. In addition, some of the scenarios, for instance proactive shopping and exhibition guides, were seen as going too far beyond the real needs of people. The Exhibition scenario presented a system that guided the visitor along a predefined route, pointed out people that the visitor should meet, and offered a possibility to make appointments to popular stands. Almost all interviewees commented that these kinds of systems might be useful for 'some businessmen' but not for them. The interviewees were mainly familiar with quite small exhibitions where the idea is to 'just wander around anyhow'. The Boater group and Family 1 also pointed out that setting up the system before the exhibition visit might become a major task. In some groups, the Exhibition scenario catalysed discussions where the interviewees identified useful applications for this kind of service. In the Senior citizens group, one lady told about using an automatic guidance system in a submarine museum in the US. The group agreed that the system should provide local information when needed but should not keep to a predefined route. The group also innovated a city guidance system based on the same idea.

Participants in the user groups mainly wanted solutions to ease their life in certain functions, but the aim of ultimate effectiveness itself was seen as bad. The interviewees thought that the location-aware reminder presented in one of the scenarios created a feeling of haste; the servant becomes a master that starts to give commands to the user. Fear of radical changes in human interaction, the usability of systems and the narrow use of new services and products were also raised in the groups.

5.2 Contents

5.2.1 Topical information

In the interviews and user evaluations, topical information turned out to be important to the users. This is the kind of information that may change while the user is on the move, in which case the information checked previously from other media (e.g. newspaper, TV or Web) may no longer be valid. Examples of such topical information are traffic information, weather forecasts, last-minute theatre ticket deals, or on-line chat. Unfortunately, in the evaluated services the only somehow topical information was included in Pointer Bensa (Gasoline), where gasoline prices at different service stations were updated once a day. In the group interviews, the Senior Citizens group as well as the Boaters group pointed out emergency situations, e.g. roadside help in a situation when the car breaks down. In this kind of situation, the traveller would need information on where (s)he could get help immediately, rather than a general service catalogue. Three Sonera Pointer test users also expressed a need for roadside services.

In addition to topical information, the users will need guidance on how to proceed in the changed situation. For instance, a train schedule as such can be obtained elsewhere but once on the move, the user will need information on delays and estimated arrival times. Furthermore, the service could suggest alternative routes in case of delays. The need for topical information is much greater than the Web is currently providing. This sets high expectations for the systems that maintain information as well as for the mechanisms by which the user will be informed and alerted.

5.2.2 Comprehensive contents

Although a mobile service on a small device can give only small portions of information to the user at a time, the user should have the possibility to access further information as exhaustively as (s)he needs. This need has been identified in other studies as well [12].

In the user evaluation of the Yellow Pages service, a user was looking for the closest restaurant with a dance floor and found that it was over 100 km away. The service did not include any of the dozens of restaurants with this facility in the city where the service was evaluated. The limited content of the services available was one of the main reasons why users considered the services not useful. The users expected to get some idea of the extensiveness of the contents as well as the geographic area that the service covers to be able to assess the usefulness of the service. Starting up a service will not be easy: on the one hand, a critical mass of service providers will be needed to get the users in, and, on the other, a critical mass of users will be needed to get the service providers in. It should be considered carefully how much content a pilot service should be included to be worth releasing.

The Cebit 2001 Guide did not have such problems because the exhibitors at a trade fair are an entity that can quite easily be managed. The earlier user evaluation results of the CeBIT2000 exhibition guide by Bieber and Giersich [7] point out that users wanted to have both more extensive information about the trade fair itself (more information about the exhibitors and more extensive search possibilities), and more information about the immediate surroundings (e.g. a shuttle plan of the fairground and overall travel information). It is not wise to restrict the available information only to the current location and time: the users may also need to plan their next activities or to return to previous activities.

Although the user groups in our evaluations were small, the results point out that information and service needs vary, not only according to the location but also according to the user and the usage situation. In one test task, the users were looking for accommodation information from the Sonera Guide. All the users missed some additional information about the hotels, but the type of information required varied. Some users would have liked to get information on price and room availability; some would have liked to get more information about the location, some about the quality and facilities. Similar variations in user preferences could be identified in making a reservation and in getting route information. The users will need comprehensive information both in breadth (number of services included) and depth (enough information on each individual service).

5.3 Interaction

Location-aware services differ from many other types of mobile services because they are not just mobile in the sense that they can be carried with the user but they are actually used on the move. The usage situations are demanding because the user can often devote only partial attention to the device, concentrating on his/her primary task of moving. In addition, the physical environment (e.g. background noise, illumination, weather) may disturb the usage situation.

5.3.1 Push or pull

In the group interviews, most users thought that they would not mind having the service or information pushed to them as long as the provided service or information was really what they needed in the situation. The attitudes of the young users (Youth groups 1 and 2; Students of well-being technology) towards location-aware ads was very positive, although the teens pointed out that they wanted the ability to select what kind of ads, from whom and when they received. The attitudes of the older interviewees were also mainly positive but more prudent. Some of the older interviewees said that they just do not like ads in general. Although the attitudes of the users were positive, location-aware push services should be designed carefully. Attitudes will soon change if the users find themselves continuously receiving information that they do not need.

In most cases, location as such is not enough to activate push but location should be complemented with personalisation. Excess content must be avoided, but at the same time the user must get an interesting, up-to-date and useful service from his/her personal point of view. User needs for personalisation are discussed in Sect. 5.4.

5.3.2 Detailed search options

When evaluating the Yellow Pages service, one person was looking for a restaurant with a dance floor, another for a vegetarian restaurant, and a third for a moderately-priced lunch restaurant. In our trial, the Yellow Pages service made no distinction between different kinds of restaurants. The probability of satisfying the user need in the search was near to chance. When getting information about gasoline stations with Sonera Bensa (Gasoline), the users in the evaluations said that in many situations they would have a greater need for information on the nearest or next gasoline station on the route than on the cheapest one provided by the service.

Vindigo provided shopping guidance by classifying the shops into women, music, sports, electronics etc. On the basis of our user interviews and evaluations, this kind of classification seems to be on too general a level. The users would benefit more from a classification based on the items that the user is looking for, e.g. tennis shoes, music scores or mobile phones.

Location-aware guidance services should provide accurate searches, both at the level of the search criteria and the classification of information. Similar needs have been identified in mobile services in general [12].

5.3.3 Planning versus spontaneity

An important issue is whether the user needs the information when (s)he is at the given location, before getting to the location, or when planning the visit. Some information or options may be needed only after the visit. Flexibility of use requires that the services should support both pre-trip planning and on-route information on occasionally found points of interest. In general, the user should be able to utilise the services without any premade plans.

Our scenarios of use presented active users who plan their visits carefully and configure their devices accordingly. While travelling around, the users are ready to fine-tune the settings and respond to all kinds of inquiries from other travellers or service providers. The same attitude could be recognised in some of the evaluated services. This is partly because of the slow mobile networks that require the user to pre-load the necessary maps and other information using faster fixed-line connections. In practice, people may not be willing to spend their time on something from which they do not get immediate benefit. In particular, people tend to be busy while on holiday; they do not want to waste a single minute of their holiday time.

In the user evaluations, most users thought that they would use location-aware services occasionally and mainly in unfamiliar environments or in emergency situations. These needs indicate that the services should be easily available when the spontaneous need for them arises. The services should be easy to find, easy to take into use and use thereafter.

The users in our interviews said that location-aware systems should not lead to a predestined and over-controlled environment. The user should feel and be in control. Similar needs have been identified in other studies as well. Cheverst et al. [6] state that the user should be able to override the recommendations of the system. Fano [9] and Espinoza et al. [13] point out that the user should have the alternative of using the system in an explorative way, i.e. having a look around without any predefined goals.

5.4 Personalisation

5.4.1 Personal options and contents

The contents of the evaluated services were quite limited and thus it is understandable that the services did not include any personalisation options. However, when the users in the evaluations described the kinds of information that they were missing from the services, it was evident that there are personal variations in preferences for both the contents and the presentation.

Also, location-awareness as such seemed to create an impression of personal entry to the service, e.g., quoting one of the test users of Sonera Pointer: ‘‘Does this phone know that we are in Tampere?’’

Personalisation in location-aware services is a good way of improving the usability of the services by providing the most essential information and the most probable options easiest available. Earlier studies on personalisation [14–16] indicate that, although the users were interested in personalised services, the motivation to actually do something for this is often quite weak. Hollensberg and Vind Nielsen [17] suggest that with WAP (Wireless Application Protocol), personalisation is a highly advanced feature, which WAP users will explore and use later, if ever. On the other hand, Ramsay and Nielsen [18] point out that personal and localised services are where WAP should be able to make an impact. It is a big design challenge to design personalised location-awareness so that it does not require too much effort on the part of the users.

The user may personalise the system according to his/her preferences but if the preferences are different in different locations, configuring the system for all these locations becomes a major task. Ways of assisting the user in the personalisation of the system are worth considering. Cheverst et al. [19] point out the need to make the user profile persistent, enabling the use of the same profile with other services as well, e.g. in the next city to be visited. Collaborative filtering can be used to identify similar users and to adapt the system according to the group profiles. The user should also be able to use the same profile with different devices.

5.4.2 User-generated contents

Far too often users are seen as passive information consumers. However, letting the users participate and provide their opinions and recommendations could enhance many services. Such contents may enrich the service, bringing in additional users and encouraging a sense of community among users. The dynamically changing information generated by other users at a particular location may be better suited to the needs of the next visitors than a more general type of information that is provided externally. The users would also benefit from the possibility to store, access and share with others their own location-based information: notes, photos, etc. [13,20,22]. The system can also identify locations frequently visited by the user and then assist the user in associating information with these places [22]. Shared contexts can encourage a sense of community among visitors [19].

In the user evaluation of the Benefon Esc!, the users appreciated the possibility to generate and store their own information related to locations that were personally important to them. In the group interviews, information provided by other visitors at the same place was deemed interesting. However, most users did not accept the idea of being contacted by strangers.

5.5 Seamless service entities

5.5.1 Consistency

With the Sonera Pointer Services, a number of usability problems arose because of inconsistencies in different Pointer services. The logic employed to navigate backwards in the services did not stay the same in different parts of the services. Some pages were organised so that links to the previous and front page were in the middle of the page. Most users did not notice that the page was continuing after those links. In one test task, the users got information on age distribution in the city of Tampere from Pointer Fakta (Facts). When reading the long list on a small screen, the users easily lost the focus and could not say which information was belonging to which age class. In this kind of lists, individual data fields should be marked consistently.

When the choice of services grows, and since the contents may come from several separate service providers, consistency becomes even more important. The user learns to use the services more easily when the information contents and structure remain similar in the different services. When travelling abroad, the user would appreciate familiar-looking services that are consistent with the services that (s)he is using at home. The internationalisation and localisation of location-based services is challenging because the local services will probably have both local users and domestic as well as foreign visiting users [21].

5.5.2 Seamless solutions support the whole user activity

In our evaluation, some of the evaluated products were designed for route guidance and some for location-aware services. On the one hand, when using route guidance services, the users often missed information about places and nearby services. On the other hand, when using location-aware services, the users expressed a need to get route guidance to the place of interest. They also wanted to have easy access to phone numbers, e.g. the possibility to call the given hotel to make a reservation. Ideally the user should see all the necessary information for a given task in a single view. The information and the options should be accessible directly from the point in the service where the need for that piece of information or option arises. The design should aim for a seamless solution whereby the user is supported throughout the whole usage situation.

Cheverst et al. [6] point out the need to be able to use the mobile system both on and off line. Connectivity affects functionality, e.g. the availability of interactive services. In our evaluations we identified similar needs – the connection may not always be on, but nevertheless the user should be provided with as much functionality as possible. For example, the users of Sonera Pointer wanted to save the search results for later reference.

5.6 Privacy

Privacy protection in location-aware services is related to the right to locate a person, use the location, store the location and forward the location. Current legislation is the basis for privacy protection but social regulation can also create rules and norms for different situations in which location-aware services are used [23]. Espinoza et al. [13] emphasise the right of the user to remain anonymous. Marmasse and Schmandt [22] suggest that problems with privacy could be avoided by performing location tracking and analysis solely on the client device. Ljungstrand [24] and Ackerman, Darrel and Weitzner [23] point out the trade-off between privacy intrusion and user benefit. They think that, if the benefit is perceived as large enough, some degree of privacy loss will probably be accepted. Continuous requests for permission may overwhelm the users and may disturb the user in his/her activity.

The World Wide Web Consortium (W3C) has defined the Platform for Privacy Preferences Specification (P3P). This specification allows web browsers to interpret the privacy practices of the Web sites and compare these with the predefined user preferences for privacy. Then the user does not have to consider privacy issues separately with each individual Web site [25]. This kind of approach could also be utilised with location-aware services.

Privacy can also be affected by the selected input/output modes of the system. Speech-based interaction may create privacy problems when used in public places [22]. Bisdikian [26] points out that using public terminals as private displays may also create privacy problems.

In our group interviews, people were worried about their privacy and the "big brother" phenomenon when considering services enabling people to be located. However, the interviewees were not worried about privacy issues with location-aware services. It did not occur to most of the interviewees that they could be located while using the service. If they did understand this, they seemed to put a great deal of faith in the current telecom operators: "The telecom operators will guard that kind of information. They already have all kinds of information about me but do not distribute it around". It was also commented that there would be regulations and legislation protecting those who use location-aware services. The danger of someone somehow abusing their knowledge of the user's location came up only occasionally in the group interviews. In the Sonera Pointer evaluation, two users spontaneously mentioned invasion of privacy as a potential threat in location-aware services.

The location-aware services should inform the users of what kind of data is collected, how is it used and who has access to it. The faith that the users have in the technology, the service providers and the policy-makers should be regarded highly. Any abuse of personal data can betray that trust and it will be hard to win it back again. The user should be able to flexibly control the release of private information such as his/her location at a given time. The user should be allowed to remain anonymous when (s)he wants.

6 Conclusions

Mobile contexts of use vary a lot and may even be continuously changing during use. The context is much more than location, but its other elements are still difficult to identify or measure. Location information is becoming an integral part of different mobile devices. Current mobile services can be enhanced with location-aware features, thus providing the user with a smooth transition towards context-aware services. Potential application fields can be found in areas such as travel and tourist information, shopping, entertainment, event information and different mobile professions.

Our interviews with potential users and the user evaluations of some of the first location-aware services point out that user expectations are high and that the users in Finland at the time of the evaluations trusted current service providers and policy-makers for issues related to privacy protection. This constitutes a good starting point for location-aware services. It did not occur to most users that they could be located when using location-aware services. This puts additional responsibility on the service providers and policy makers.

Our results highlight the need for comprehensive services, in terms of geographic coverage, breadth (number of services included) and depth (enough information on each individual service). The selection of contents and options should cover the different needs of individual users and the user should be given a realistic description of the coverage of the service. The need for topical information is high: the users can get static information from elsewhere before starting off on their journey. User needs may be related to the past, current or planned location and each user may have personal preferences about what (s)he needs in the different usage situations. The users need seamless service chains that serve them throughout their mobile activity, e.g. planning, searching services, finding the route as well as visiting and storing information.

The evaluation results point out the need for spontaneous and occasional use. The services should be easy to find, and it should be easy to get an overview of the available services as well as their coverage, and services should be easy to take into use and use thereafter.

The willingness of users to be active should not be overestimated. Although the users would benefit from a personalised service, they may not be ready to define a profile separately for each service and each context of use. On the other hand, users may want to participate in the content creation instead of being passive information consumers. Users want to create and store their own location-aware data and they may be willing to share this data with others.

Acknowledgements Our research work was carried out with the financial support of the Personal Navigation (NAVI) Programme, hosted by the Ministry of Transport and Communications of Finland. The author wishes to thank the members of the NAVI management board for their valued co-operation: Benefon, Digia, Fortum, Genimap, Hewlett-Packard, ICL Invia, Nokia Mobile Phones, Novo Group, Radiolinja, Satama Interactive, Sonera, Suunto, TietoEnator, the Ministry of Transport and Communications, the Ministry of Agriculture and Forestry, the Ministry of Education and Culture, the Ministry of Defence, the Ministry of the Interior, the Ministry of Social Affairs and Health, the Finnish National Fund for Research and Development SITRA, the National Technology Agency TEKES and the Finnish National Road Administration as well as the SME enterprises of the NAVI network. I also wish to thank the superb evaluation team at VTT: Ari Ahonen, Veikko Ikonen, Minna Kulju, Sari Lehtola and Tuula Petäkoski-Hult.

References

1. Dey AK (2001) Understanding and using context. Personal and Ubiquitous Computing 5: 20–24
2. Ancona M, Locati S, Romagnoli A (2001) Context and location aware textual data input. In: Proceedings 2001 ACM Conference on Applied Computing (SAC). Las Vegas. ACM pp 425–428
3. Fano AE (2001) What are the location's "File" and "Edit" menus? Personal and Ubiquitous Computing 5: 12–15
4. Davies N, Mitchell K, Cheverst K, Blair G (1998) Developing context sensitive tourist guide. In: Proceedings First Workshop on Human Computer Interaction with Mobile Devices.
5. Svanæs D (2001) Context-aware technology: a phenomenological perspective. Human-Computer Interaction 16: 379–400
6. Cheverst K, Davies N, Mitchell K, Friday A, Efstratiou C (2000) Developing a context-aware electronic tourist guide: some issues and experiences. In: CHI 2000 Conference Proceedings. ACM pp 17–24
7. Bieber G, Giersich M (2001) Personal mobile navigation systems – design considerations and experiences. Computers & Graphics 25: 563–570
8. Ueda H, Tsukamoto M, Nishio, S (2000) W-MAIL: An electronic mail system for wearable computing environments. In: Proceedings 6th Annual Conference on Mobile Computing and Networking (MOBICOM).
9. Fano AE (1998) Shopper's eye: using location-based diltering for a shopping agent in the physical world. Autonomous Agents '98, Minneapolis, USA
10. Chávez E, Ide R, Kirste T (1999). Interactive applications of personal situation-aware assistants. Computers & Graphics 23: 903–915
11. Rainio A (2000) (ed). Personal Navigation NAVI Programme 2000– 2002. VTT Research Notes 2038. pp 19–20
12. Billsus D, Brunk CA, Evans C, Glandish, B, Pazzani M (2002) Adaptive interfaces for ubiquitous web access. Communications of the ACM 45(5): 34–38
13. Espinoza F, Persson P, Sandin A, Nyström H, Cacciatore E, Bylund M (2001) GeoNotes: Social and navigational aspects of location-based information systems. Lecture Notes in Computer Science 2201. Springer pp 2–17
14. Nielsen J. Personalization is over-rated, Jakob Nielsen's Alertbox, October 4 1998. [online, cited 29.10.2001] http://www.useit.com/alertbox/981004.html
15. Manber U, Patel A, Robinson J (2000) Experience with personalisation of Yahoo! Communications of the ACM 43(8): 35–39
16. Kaasinen E, Kasesniemi E-L, Kolari J, Suihkonen R, Laakko T (2001) Mobile-transparent access to web services – acceptance of users and service providers. In: Proceedings 18th International Symposium on Human Factors in Telecommunication. pp 227–234

17. Hollensberg J, Vind Nielsen H (2000) Survey of usability on Danish WAP portals – How the Danish telecom operators entered the WAP world. Mobilethink as, [online, cited 29.10.2001] http://www.mbusinessinsight.de/deutsch/download.html
18. Ramsay M, Nielsen J (2000) WAP Usability. Déjà Vu: 1994 all over again. report from a field study in London, Fall 2000. Nielsen Norman Group. [online, cited 20.10.2001] www.NNgroup.com/reports/wap
19. Cheverst K, Smith G, Mitchell K, Friday A, Davies, N (2001) The role of shared context in supporting cooperation between city visitors. Computers & Graphics 25: 555–562
20. Goel A (2001) Urban Pilot: a dynamic mapping tool for personalizing the city through collective memory. In: Proceedings Fifth International Conference on Information Visualisation. pp 227–232
21. Ahonen A, Ikonen V, Kaasinen E (2002). Localisation of personal navigation products and services. In: Coronado J, Day DL and Hall B (eds) Designing for global markets 4. Proceedings Fourth International Workshop on Internationalisation of Products and Systems. pp 101–112
22. Marmasse N, Schmandt C (2000). Location-aware information delivery with ComMotion. In: Thomas P, Gellersen HW (eds) Handheld and Ubiquitous Computing. Second International Symposium, HUC 2000. Springer. pp 157–171
23. Ackerman M, Darrel T, Weitzner DJ (2001). Privacy in context. Human-Computer Interaction 16: 167–176
24. Ljungstrand P (2001). Context-awareness and mobile phones. Personal and Ubiquitous Computing 5: 58–61
25. The World Wide Web Consortium, The Platform for Privacy Preferences 1.0 (P3P1.0) Specification. W3C Recommendation 16 April 2002. [online, cited 25.11.2002] www.w3c.org/TR/P3P
26. Bisdikian C, Christensen J et al (2001). Enabling location-based applications. WMC 01, Rome. ACM

Paper VI

Kaasinen, E. 2005. User acceptance of location-aware mobile guides based on seven field studies. Journal of Behaviour & Information Technology, Vol. 24, No. 1, January–February, pp. 37–49. Taylor & Francis.

Behaviour & Information Technology, Vol. 24, No. 1, January – February 2005, 37 – 49

Taylor & Francis
Taylor & Francis Group

User acceptance of location-aware mobile guides based on seven field studies

E. KAASINEN*

VTT Information Technology, PO Box 1206 (Sinitaival 6), FIN-33101 Tampere, Finland

Location-awareness provides mobile users with topical and personal contents that may increase the appeal of mobile guides in different application fields. Based on the results of seven field studies, this paper identifies factors of usability, utility and user trust that affect the user acceptance of location-aware mobile guides. The results indicate that on-the-move use sets high requirements for ease-of-use and, in particular, ease of taking the services into use. The need for location-aware mobile guides turned out to be occasional, but in many of those occasional usage situations the guides proved extremely useful. The trade-off between privacy protection and effortless use requires careful consideration, even though in these studies the users turned out to be trusting towards the service providers. The results indicate that location-awareness can provide the users with easy-to-use and situationally relevant mobile guides.

1. Introduction

Mobile devices provide good platforms for different guidance services. Potential application fields can be found in areas such as travel guidance, tourism services, shopping and event information. In this paper a mobile guide refers to a mobile device running an application or service providing guidance information or functions. A guidance service refers to the application or service apart from the device platform.

Varying contexts of use are typical for mobile services. Utilising information about the context of use provides a basis for new types of mobile services that are not just mere duplicates of fixed network services but services that are specifically intended for mobile use. These services can give the user quick access to the information or services that s/he needs in his/her current context of use. Services can even be activated automatically according to the identified context.

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task (Dey 2001). The location of the user is one of the few elements of the context that currently can be measured fairly accurately, depending on the positioning system in use. In this paper, location-aware services are defined as context-aware services that utilise the location of the user to adapt the service accordingly. In mobile guides, the information and service needs often depend on the location of the user, thus providing potential for location-awareness.

In the following, the results of seven field studies of different location-aware mobile guides are analysed and, based on the analysis, key factors that affect user acceptance of location-aware mobile guides are identified. According to the Technology Acceptance Model (TAM) by Davis (1989), user acceptance of technology is built on perceived ease of use and perceived utility. In this study, user trust is added as the third element of acceptance, indicating how confident the user feels about using the technology in real usage situations. Fogg and Tseng (1999) define trust as an indicator of a positive belief about the perceived reliability of, dependability of, and confidence in a person, object or process. Shneiderman (2000) refers to the definition of trust by Fogg and Tseng and further emphasises that trust is about expectations of the future. Shneiderman defines trust as the positive expectation a

*Email: eija.kaasinen@vtt.fi

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2005 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01449290512331319049

person has for another person or an organisation based on past performance and truthful guarantees. In this paper, user trust is defined according to Fogg and Tseng (1999). User trust includes elements such as perceived technical reliability, user trust in him/herself being able to control the technology and user trust that the technology does not include any unwanted features, e.g. related to privacy protection.

The research was carried out as a part of the Personal Navigation Research and Development Programme (NAVI) in Finland (Rainio 2000). The aim of the programme was to facilitate cooperation between different actors who were developing personal navigation products and services. Within a horizontal usability support project, our research group carried out usability case studies with the projects of the NAVI programme. The aim was to support the individual projects and, beyond this, to identify guidelines for acceptable personal navigation services.

2. Related research

Context-awareness can be implemented as an adaptation of the user interface or the contents of the service. Both approaches can be utilised in mobile guides. A context-aware user interface can select the appropriate modes for service interaction. A context-aware user interface can also be implemented, e.g. as context-aware text prediction (Ancona *et al.* 2001) or a location-aware remote control for the environment (Fano 2001). A major challenge for the context-aware user interface is that the context may be continuously changing. This raises the problem of integrating changes into the user interface in such a way that the user remains in control (Cheverst *et al.* 2000).

In previous research, context-aware guidance services have been studied in application areas such as tourist guidance (Cheverst *et al.* 2000), exhibition guidance (Bieber and Giersich 2001), event information (Fithian *et al.* 2003) and shopping (Fano 1998). These studies indicate that in well-focused application areas it is possible to predict the other elements of the context according to the location of the user. Mobile guides are often designed for focused user needs and contexts of use and are thus good candidates for utilising location-awareness.

During the first phase of our project in 2001, we interviewed different focus groups and evaluated some of the first commercially available personal navigation products and services. Based on these studies, we identified key user needs for location-aware services: topical and comprehensive contents, smooth user interaction, personal and user-generated contents, seamless service entities and privacy protection (Kaasinen 2003). The results of these studies formed a good basis for further studies in connection with the design of more advanced mobile guides.

The need for comprehensive studies on user needs and user acceptance of mobile services has been identified by other researchers as well. Sarker and Wells (2003) claim that mobile devices and service manufacturers continue to produce wireless handheld devices based on quite abstract conceptions of what the 'generalised' mobile user might value and desire. According to them, a clear understanding of the motivations and circumstances surrounding mobile device use and adoption from the perspective of the consumers themselves is still missing.

In their study of Internet and Wireless Application Protocol (WAP) enabled mobile phones, Teo and Pok (2003) identified personalisation and convenience of use as key success factors of mobile Internet services. They claim that convenience of use can largely be achieved by time- and context-sensitivity. Barnes and Huff (2003) have applied the Technology Acceptance Model to understand why the Japanese mobile Internet solution, iMode, has become so popular in Japan. According to their study, key success factors include relative advantage compared to other alternatives for Internet access, low costs, simple and intuitive interface, compatibility with the user's previous experiences with phone usage, trialability and observability (seeing how other people use the system) as well as the good image of the service provider (NTT DoCoMo) and user's trust in this well-known brand. In addition to usability, utility and trust, these success factors include factors affecting the social acceptance of the services. In the iMode case, individual success factors may be culturally specific but the classification can be applied more widely.

3. The empirical studies

This paper refers to a series of field studies that we carried out together with the projects of the NAVI programme. Our aim was to identify factors that affect the user acceptance of location-aware guidance services. The findings could then be utilised as guidelines for designing other location-aware mobile guidance systems. Because we were dealing with prototype services and relatively small user groups, the research focus was on the attitudes of individual users rather than on social acceptance. The field studies were selected so that they represented different types of services, from simple Short Message Service (SMS) based services to mobile Internet service entities. Maps and route guidance are important complementary information for location-aware guides and that is why two case studies were focused on these issues. The studies included three short-term field evaluations, two long-term field evaluations, one case study with field interviews and one survey study. Table 1 gives an overview of the case studies.

The device platforms in the field studies included mobile phones, media phones with larger colour screens (Nokia 7650) and personal digital assistants (Compaq iPaq). The media phones were operated by a joystick with five-way navigation, whereas the PDAs had touch screens with stylus. The positioning methods included GPS and cell-based positioning. The test users in the field evaluations and interviews were members of the general public; they represented both sexes, different age groups and different backgrounds. Table 2 gives an overview of the test users in the field evaluations.

The evaluated services and service platforms were different, but table 2 aims to classify the users in each field evaluation into three groups, depending on how experienced the users were with the device platform or the service type. Most users were inexperienced or had only some experience. The users received thorough guidance on how to use the devices so that in the evaluations they were not distracted by problems related to device usage but could concentrate on using the services to be tested. For the same reason, we selected for the SMS service evaluations (field studies 1 and 2) users who were at least somehow familiar with mobile phones and SMS usage; for the WAP evaluations (field study 3) we selected users who were at least somehow familiar with the Internet. In the relatively small user groups, we could not identify usability problems related to user gender or age.

Table 1. The field studies.

No.	Evaluated services	No. of users	User device	Usage environment	Method
1	Location-aware SMS services: service directory, event calendar and route guide	6	User's own mobile phone	City centre	Two-hour field evaluation
2	Weather and road conditions by SMS	10	User's own mobile phone	Highway	One-month field evaluation
3	Location-aware integrated service directory: several service directories, event directory, maps	7	Media phone (Nokia 7650)	Different cities	Three-week field evaluation
4	Mobile topographic maps	6	PDA (Compaq iPaq)	Nature park	Three-hour field evaluation
5	Mobile 3D maps	6 + 4	PDA (Compaq iPaq)	City centre	Laboratory evaluation and two-hour field evaluation
6	Location-aware tourist information	300 + 70	Mobile phones and PDAs	Ski resort	Web survey and survey at the ski resort
7	Scenarios of context-aware consumer services	28	Mobile phones and PDAs	Bowling hall, cafe, concert hall and similar urban public places	Interviews in anticipated contexts of use

Table 2. An overview of the test users in the field evaluations.

No.	Evaluated services	No. of users and ages	Inexperienced	Some experience	Experienced
1	Location-aware SMS services: service directory, event calendar and route guide	6 users ages 25–55	2 users had never used value- added SMS services	3 users had tried SMS services a few times	1 user had been using SMS services weekly for a few years
2	Weather and road conditions by SMS	10 users ages 31–52	2 users had never used value- added SMS services	5 users had been using SMS services occasionally	3 users were often using SMS services
3	Location-aware integrated service directory: several service directories, event directory, maps	7 users ages 25–54	All users were unfamiliar with mobile Internet on a phone	All users had at least some experience of Web usage	2 users were very experienced mobile phone users
4	Mobile topographic maps	6 users ages 24–60	4 users had not used a PDA or GPS before	1 user was using a PDA monthly and GPS occasionally	1 user had before been using PDA daily and had tried GPS once
5	Mobile 3D maps	4 users ages 27–54	1 user had not used a PDA or 3D models before	2 users were sometimes using a PDA and had been using 3D models but not often	1 user was using a PDA regularly, but rarely 3D models

Field studies 1 and 2 dealt with SMS-based services. In field study 1, we evaluated a service entity that consisted of a location-aware service guide, a location-aware event calendar and a route guide, each being a separate SMS service. The user could search for services or events nearby or close to a given address, and get route guidance as written turn-by-turn instructions. Field study 2 was carried out in cooperation with the Finnish Road Enterprise. The pilot service provided location-aware weather and road condition information upon user request (figure 1). The service was operational on Highway 1, between Helsinki and Turku (two major cities in Finland). Usage data was logged and the test users were interviewed before and after the trial (Rämä 2003). Field studies 1 and 2 demonstrated the possibilities of current SMS-based mobile services that can be provided to practically all currently available mobile phones. These services were easy to adopt, but remembering the key words and inputting them caused usability problems (Södergård *et al.* 2003).

Field study 3 dealt with a mobile Internet service entity that provided the users with location-aware access to different service directories and databases (figure 1). The service entity also included access to a map service and to a location service from where user location could be retrieved, irrespective of the positioning method in use on the user's device. In the field trial, we used cell-based positioning and WAP (Wireless Application Protocol) enabled mobile phones. This case study demonstrated the possibilities of current mobile Internet services and integration of different services into a comprehensive service entity. The users appreciated the versatility of the service entity but the complexity of the service architecture caused problems in starting and learning to use the service as well as in coping with error situations (Kaasinen *et al.* 2003b).

Field studies 4 and 5 dealt with route guidance. In field study 4, mobile, adaptive topographic maps (figure 1) were evaluated in a nature park in Finland with six users. The participants used topographic maps and GPS positioning on PDAs as part of a guidance system (Nivala *et al.* 2003). In field study 5, mobile route guidance was provided in the form of 3D city models. User evaluations were carried out first in laboratory conditions and later in the city of Helsinki with four test users (Kulju and Kaasinen 2002). These studies, as well as field study 3, which also included maps, showed that maps are essential elements in location-aware services. Even if the realistic 3D city models facilitated route guidance in field study 5 by providing easily recognisable landmarks, the users readily referred to traditional maps to get an overview of the environment. The results of field studies 3, 4 and 5 indicated that map information needs varied largely according to the user and the context of use. Personalised and context-aware contents and presentation of mobile maps could make the maps much more useful and functional for the users.

The aforementioned five field evaluations were complemented with survey studies (field study 6) and contextual

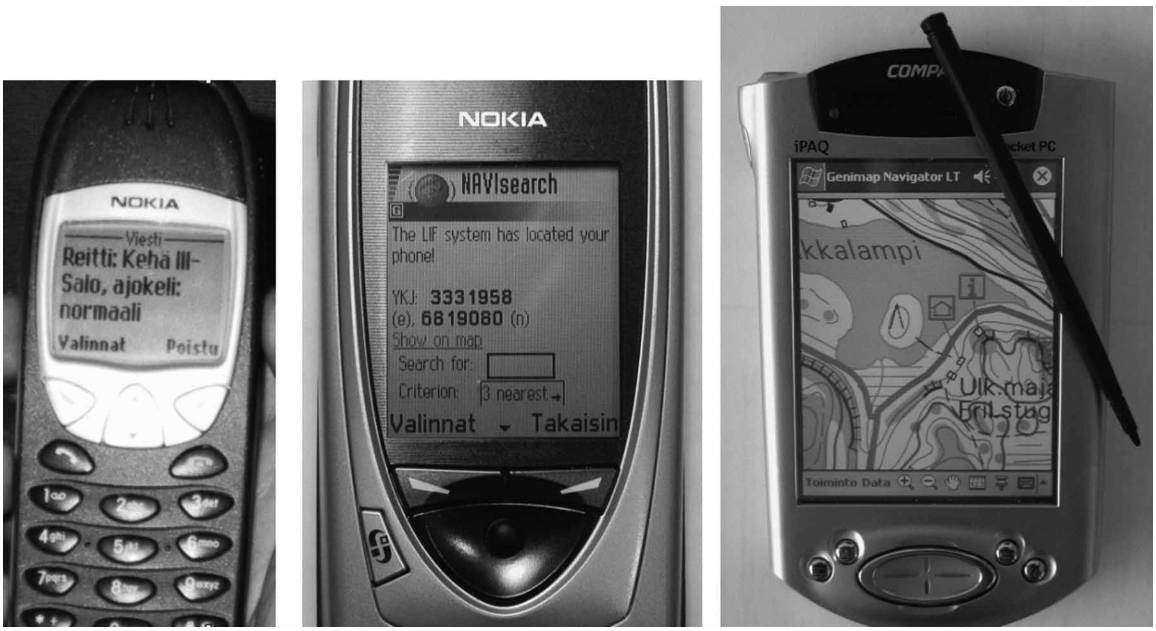


Figure 1. The device platforms included mobile phones, media phones and PDAs. From left: an SMS-based road condition guide on a mobile phone (Nokia 6210; field study 2), a location-aware service directory on a media phone (Nokia 7650; field study 3) and a topographic map on a PDA (Compaq iPaq; field study 4).

interviews (field study 7). Field study 6 identified user needs for location-aware and personalised tourist information in a traveller service targeted at Levi ski resort in northern Finland. The results of a web survey of potential Levi visitors (300 interviewees) and a survey study at the Levi resort (70 interviewees) were analysed in design walk-through meetings. The meetings brought together experts on technology, usability, legislation and ethical issues. Already the active response to the survey revealed that tourists were keen on location-aware and personal tourist information services (Kaasinen *et al*. 2003a). The results of the survey indicated needs for community-based services for regular visitors to Levi but also needs for more traditional guidance for newcomers. During the peak season, the main need for location-aware guidance was to make reservations, e.g. at rental services and restaurants. At other times of the year, the focus shifted to obtaining information on special offers and opening hours. User needs for context-awareness were also identified in connection with the phase of the holiday. The tourists would need the service in slightly different forms before, during and after their holidays, and context-awareness could provide them with situationally relevant information and services.

In field study 7, contextual interviews were carried out to identify how people describe different contexts in their lives, how location is related to these contexts, and what the information and communication needs in those contexts are (Kolari and Virtanen 2003). Twenty-eight people representing different and partially overlapping user groups were interviewed. The interviews took place in real contexts typical for the group, e.g., a bowling hall, a cafe and a music event. The results of this study showed that there are contexts of use that can be identified automatically based on time and the location of the user, especially contexts related to events, public places and workplace environments. It seems that location-aware guidance services have good potential in these fields.

Rather than presenting the detailed results of each field study, this paper will outline the bigger picture revealed by the studies. The following chapters will outline key factors affecting user acceptance of location-aware mobile guides, based on the combined results of the field studies and classified as findings on usability, utility and trust.

4. Findings on usability

4.1 *Facilitating momentary usage sessions on the move*

In all the field evaluations it became evident that location-aware guidance services were not just mobile in the sense that they could be carried with the user, but that they were often really used while on the move (figure 2). The user should be able to devote his/her attention to his/her main task of moving, and using the service should not require too much attention. The mobile phones used in our trials were designed for one-hand use, which made them easier to use while moving. Using a PDA proved more difficult because its use required both hands and the stylus easily got lost in mobile use (Nivala *et al*. 2003).

The usage sessions were typically not continuous but the users occasionally wanted to re-access the service. The mobile devices used in the case studies did not have a function allowing the input to be locked in the middle of a usage situation. This caused problems when the user wanted, for instance, to concentrate on other activities and put the device into his/her pocket in the middle of a usage session (Kaasinen *et al*. 2003b, Nivala *et al*. 2003). In line with our findings, Fithian *et al*. (2003) emphasise the importance of task resumability and point out that, in particular, the soft keyboard of the PDA is not suited to on-the-move use.

4.2 *Effortless user input by complementary input methods*

User input is one of the most demanding tasks in mobile services. Text input turned out to pose the most serious usability problems in the field evaluations (Kaasinen *et al*. 2003b, Södergård *et al*. 2003). On-the-move use made the already cumbersome input of text even worse. In field studies 1 and 3, the participants used text input for addresses and search terms. This kind of user input is typical for mobile guides. The street names were often long and tedious to write and sometimes the users had problems in spelling them. The same problems were encountered with search terms, although not as often. In both studies, the users were often unsure of what they could search for, and in which form they should input the search term. Five out of the seven test users in field study 3 said that instead of writing the search term they would have preferred to pick a service class from a ready-made list (Kaasinen *et al*. 2003b).

Chae and Kim (2004) found out in their study of mobile Internet usage on mobile phones that as screen sizes are reduced, the horizontal depth of navigation paths has a considerable negative effect on usability. The negative effect increases drastically with more complicated user tasks. In mobile guides, the horizontal depth of the navigation paths easily increases if the user is provided with detailed alternatives, e.g. Shopping/Clothes/Women/Casual. Although mobile Internet usability guidelines recommend minimised text input, other recent studies suggest that users may prefer keyword search to navigation with links that may introduce complicated navigation paths (Kaikkonen and Roto 2003). Our results indicate that instead of simply minimising text input, a more comprehensive view of effortless user input is needed. Ready-made selection lists are especially useful when the

Figure 2. On-the-move use was often quite demanding.

user is getting acquainted with the service, whereas experienced users may prefer text input (Kaasinen *et al.* 2003b, Södergård *et al.* 2003). Preferably, both alternatives should be available. Text input should be predicted and interpreted to correct misspellings and to make input less tedious. Location-awareness itself could also be utilised in text input, as suggested by Ancona *et al.* (2001). For instance, nearby streets or the most popular search terms at a certain location could be suggested to the user.

4.3 *Mobile maps and unambiguous route guidance are essential elements in mobile guides*

Recent studies of mobile services (Bornträger *et al.* 2003, Fithian *et al.* 2003, Kaasinen 2003) suggest that users appreciate service entities where maps complement location-aware services. There are already a number of systems supporting the display of maps on mobile devices but these systems are mainly focused on the presentation of traditional maps on small screens. Patterson *et al.* (2003) point out that maps on mobile devices have potential for many new features beyond mere presentation of maps on a small display. Brown and Perry (2002) claim that maps on mobile devices do not need to be just 'handy' representations of space but could allow a variety of re-representational capabilities, tuning the spatial representation to the user's needs at that moment. As examples, Brown and Perry name overlaying information onto objects in the environment and integrating user-generated information into places.

In field study 4, which was focused on mobile topographic maps, the users appreciated the possibility of using the map as a user interface. However, this two-fold role also caused usability problems (Nivala *et al.* 2003). For example, when the user pointed to a target on the map, it was not clear whether s/he was showing the target to a friend, whether s/he wanted to get more information about the target, or whether s/he just wanted to scroll the map in that direction. Utilising fluently interactive maps would have required more learning than was possible in our short field evaluations.

On the small screen, only a small proportion of the map could be shown at a time (figure 3). Parts of street names were truncated or they were not displayed at all. Also, the users had problems in associating the small map with the larger map entity (Kaasinen *et al.* 2003b, Nivala *et al.* 2003). This would have required additional identification information on the map, such as the name of the city district or indication of landmarks outside the map.

The users appreciated the different map levels available. A large scale was needed when following a route, whereas a small scale was needed to get an overview of the environment, the route or nearby points of interest (Kaasinen *et al.* 2003b, Nivala *et al.* 2003). The results of field studies 3 and 4 indicated the need for personalised and context-aware map presentations: the users had different preferences in the information that they would have liked to see on the map and these preferences were often related to their current context of use. For instance, in the nature park in field study 4, experienced outdoor people wanted to travel independently, and therefore needed detailed terrain information on the move (Nivala *et al.* 2003). People travelling with small children needed estimates of the difficulty of the guided routes available as well as detailed information about services available on the route.

In our field evaluations, the users got route guidance as textual directions (field study 1), maps (field studies 3 and 4) and routes in a 3D-city model (field study 5). On the route, the users needed information about the direction, the distance to the next turn, where to turn next and the total distance to the destination. The route guidance did not succeed in being unambiguous in any of these studies (Kulju and Kaasinen 2002, Kaasinen *et al.* 2003b, Nivala *et*

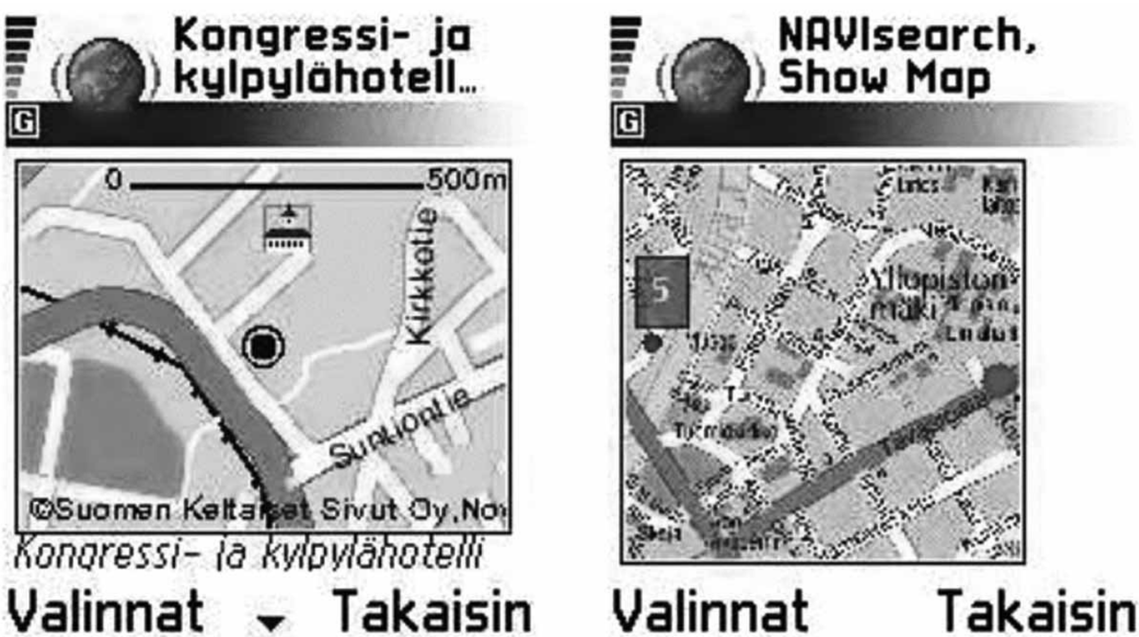


Figure 3. Typical usability problems on small maps: not all street names are shown and it is not easy to recognise the district.

al. 2003, Södergård *et al*. 2003). The textual guidance in the SMS route guide worked well but the users pointed out that they would appreciate simplicity of the route more than an optimally short route with a lot of turns (Södergård *et al*. 2003). Starting on the route was the most critical phase of the route guidance. In the SMS service, the starting direction was expressed as a point of the compass, e.g. 'Go west along Mikonkatu Street for 200 meters' but in the real environment quite a few users did not know which way was west, not even in familiar regions (Södergård *et al*. 2003).

With mobile maps and cell-based positioning, the user could see his/her location and the destination on the map. Due to inaccurate positioning, the actual location of the user was sometimes on the edge of the map, where it was not easy to recognise (Kaasinen *et al*. 2003b). The users suggested e.g. a circle surrounding their location on the map to indicate the estimated accuracy of the location. Feedback on accuracy was especially important because it turned out that the users had quite unrealistic expectations of the positioning accuracy of cell-based positioning (Kaasinen *et al*. 2003b). With GPS positioning the user saw his/her own location on the map quite accurately, as well as his/her progress along the route. However, not even GPS positioning accuracy was adequate enough for unambiguous start-up on the route (Kaasinen *et al*. 2003b, Nivala *et al*. 2003). This would have required the user to be located on the right side of the street on the map. In their studies of location-based games, Benford *et al*. (2003) suggest two strategies to cope with inaccurate GPS positioning: hiding uncertainty or revealing uncertainty. By hiding they mean correcting automatically unacceptable positions, e.g. from inside a building to one side of the building. Revealing uncertainty can be done e.g. by colour-coding maps to indicate good and bad areas for GPS positioning. Both strategies could also be utilised in mobile guides.

Orientation information is important in addition to the location. The GPS devices in our field evaluations provided orientation information only when the user was moving because the devices did not include a separate compass. Unfortunately, the orientation was needed most when standing still and trying to decide which way to start walking (Kaasinen *et al*. 2003b, Nivala *et al*. 2003, Södergård *et al*. 2003). Continuous orientation information would be especially useful to provide the user with explicit guidance to start on the route. Today many GPS devices already include a separate, integrated compass, making them more useful in route guidance.

The 3D city model (field study 5) provided easily recognisable landmarks that helped the user to recognise his/her location on the route. The users utilised e.g. the colours, materials or shapes of buildings as landmarks to locate themselves in the 3D model (Kulju and Kaasinen 2002). Landmarks could be utilised with traditional maps as well. The types of good landmarks varied by user, environment and task. In our studies, parks turned out to be good landmarks in cities (Kulju and Kaasinen 2002), whereas in the nature park, man-made constructions like bridges and campfires were favoured (Nivala *et al*. 2003). Even landmarks outside the actual map area could be used to help the user to locate the small map on the screen in the actual environment.

5. Findings on utility

5.1 *User attitudes indicate good potential for location-aware mobile guides*

In all the evaluations, the users were quite keen on location-aware guides. Service and event guides were found useful and the users commented positively on the detailed information provided, such as ticket price, contact information and distance (Kaasinen *et al*. 2003b, Södergård *et al*. 2003). The users of the weather and road condition service (field study 2) had several ideas to enhance the service concerning the traffic situation, route guidance, gas stations, maintenance actions, forecast for the next day and so forth (Rämä 2003). In field study 6, both newcomers and regular visitors pointed out different kinds of information that they considered very useful in the skiing resort guide (Kaasinen *et al*. 2003a).

In the contextual interviews (field study 7) event information was appreciated by all users. They were keen to access topical information such as news and changes to event programs. The need for such information increased with the geographical spread and/or timescale of the events. The younger interviewees liked the idea of peer messaging that let them participate in content creation instead of being just passive information consumers (Kolari and Virtanen 2003). In the survey at the ski resort (field study 6), the users broadly liked the idea of peer messaging through a public bulletin board available on ski resort monitors, accessible by personal mobile phones (Kaasinen *et al*. 2003a).

5.2 *Usage needs are typically spontaneous and occasional*

User feedback and the results of the field trials indicated that the need to use location-aware guides was quite occasional. For instance, in field study 2 (road conditions) the number of usage sessions per user varied from 2 to 50 during the one-month trial (Rämä 2003), and in field study 3 (integrated service directory) from 5 to 20 during the three-week trial (Kaasinen *et al*. 2003b). Occasional usage is understandable: route guidance is needed when travelling in unfamiliar regions; event information is related to a particular occasion that may last a few days or weeks at most. Some services are available only in certain locations. Occasional usage requires the services to be easily available on standard phone and PDA platforms rather than providing the users with dedicated mobile guidance devices.

Occasional usage also requires ease of taking the service into use. The first-time use of a service involves finding out about the existence of the service, assessing whether and how the user could utilise the service, installing it and, finally, using it. All this should be made effortless for the user in the spontaneous situation in which s/he feels motivated to start using the service.

The usefulness of the services should not be assessed by the usage figures only. In our evaluations, although the usage was occasional, the users assessed the guidance services as being very useful in many of those occasional usage situations (Kaasinen *et al.* 2003b, Rämä 2003, Södergård *et al.* 2003).

5.3 *Location predicts the context of use in mobile guides*

Context is much more than location and time, but in field study 7 we wanted to study the extent to which these attributes can be used to predict the usage context in different environments. In the interviews, we were able to recognise some common contexts that most users described in a similar way and that could largely be related to certain locations (Kolari and Virtanen 2003). These common contexts were ones related to landmarks, city areas, sport fields or event venues. These contexts of use could often be recognised by the user's location and, due to the nature of the contexts, they provide possibilities for context-aware guidance services.

The results of the other field studies also indicate that location is the key element of the context of use in mobile guides (Kaasinen *et al.* 2003a,b, Nivala *et al.* 2003, Södergård *et al.* 2003). Location as such or together with time can in many cases indicate the context of use.

5.4 *Users need comprehensive service entities*

In their foundational studies of groupware adoption, Grudin and Palen (1995) pointed out that adoption can succeed without mandated use. Groupware features, some of which provided benefits to individuals, attracted a critical mass of users. After reaching a critical mass of users, social pressure by peers and others started extending the service usage. In services targeted at consumers, mandated use is out of question. Sarker and Wells (2003) point out the importance of critical mass in service adoption, both in terms of the number of users in the user's social network and sufficient number of services. In our earlier studies (Kaasinen 2003), we also identified the importance of comprehensive service entities that support the user throughout his/her task and include a critical mass of points of interest as well as enough information about each individual point of interest.

In field study 3, the users appreciated the comprehensiveness of the service contents: the wide selection of points of interest, broad information about each individual point of interest, and the completeness of the service (Kaasinen *et al.* 2003b). The primary need of the users was to get access to an adequate amount of information and services. When sufficient content was available, the users wanted to get easy access to personally relevant information. Systems that adapt according to the user's previous behaviour or utilise group profiles will not require much user effort and may thus be successful personalisation solutions. In field study 6 at the skiing resort, we found that the same user may need quite a similar guidance service at other skiing resorts as well (Kaasinen *et al.* 2003a). The user would benefit if s/he could export his/her user profile from one service to another. As proposed by Cheverst *et al.* (2000), the user would benefit if s/he could export his/her user profile from one service to another instead of defining the attributes separately, e.g. for different city guides.

6. Findings on user trust

6.1 *Solving trade-offs between effortless use and privacy protection*

Privacy protection in location-aware services is related to the right to locate a person, use the location, store the location and forward the location. The legislation in most countries requires the user's permission before s/he can be located. This is the basis for privacy protection, but social regulation can also create rules and norms for different situations in which location-aware services are used (Ackerman *et al.* 2001). In practice, trade-offs between privacy protection and effortless use need to be resolved. The more seamless and easy the use of a location-aware application gets, the fewer the cues required to remind users that their location is being monitored (Barkhuus and Dey 2003, Kaasinen 2003). The user should be reminded that s/he is being located but this should not require excessive user effort, e.g. in the form of acceptance requests. Barkhuus and Dey (2003) point out the trade-off between usefulness and lack of control. They claim that the more useful the user considers the system, the more s/he may be willing to give away control. Giving away user control should be voluntary and definitely not 'the price of usefulness'.

In the contextual interviews of field study 7, getting information about other people's location was considered intrusive (Kolari and Virtanen 2003), similar to previous studies (Cheverst *et al.* 2000, Fithian *et al.* 2003). The interviewees pointed out that in many cases it is better to convey location information indirectly, in the form of a user-defined description of the current context (e.g. downtown, at work, shopping or partying) rather than exact co-ordinates (Kolari and Virtanen 2003). In field studies 3 and 6, we found that transferring user permissions between location and service providers may get quite complicated (Kaasinen *et al.* 2003a,b). GPS and proximity-based positioning are easier in this sense since, with these systems, the location data is possessed by the user (GPS) or generated by the service provider.

In field study 3, the users recognised situations where they would like to convey their location to service providers, e.g. in a car breakdown situation (Kaasinen *et al.* 2003b). In field study 6, we identified situations where continuous positioning would be beneficial for both the user and the service provider (Kaasinen *et al.* 2003a). If the user rents a snowmobile, s/he would feel safer if s/he knew that in case of emergency, the rental service could locate his/her snowmobile. Similarly, rental service providers would appreciate a service that enables them to locate their fleet at any time. These findings indicate that the possibility of enabling the user to manage his/her location data him/herself may be useful in different kinds of location-aware services.

User location is very personal information and it should be protected even when the user him/herself does not require it. The user should be provided with easy mechanisms for giving permission to use the data for a predefined purpose. The user should get clear feedback on whether the positioning is on or off (Kaasinen *et al.* 2003a,b). Location histories should not be stored purposelessly and without user consent (Kaasinen *et al.* 2003a). When location is conveyed to others, it is worth considering whether they will need the exact location co-ordinates or more descriptive but less intrusive description defined by the person to be located (Kolari and Virtanen 2003). It should also be considered whether it is necessary to connect the location to the user identity.

6.2 Push services should stay under user control

Location-aware push services provide the users with information automatically, activated by the location of the user. In many countries, legislation prohibits the sending of commercial messages without user permission. This saves the users from receiving spam messages, but also generates problems in informing the users about proximity-based services.

Users may accept push services because of the effortless use (Cheverst *et al.* 2002). Teo and Pok (2003) have made a survey study with 1000 users of mobile Internet in Singapore. Their study revealed that 60% of the respondents would not reject mobile advertisements provided that the advertisements are personalised and that accessing the advertisements will subsidise their phone bill. 65% of the respondents to the web survey in field study 6 indicated that they would like to receive special offers via their mobile phone while at the Levi skiing resort (Kaasinen *et al.* 2003a). In addition to commercial messages, nightly aurora borealis notifications and storm warnings were mentioned as examples of useful push services. The attitudes of the users were quite positive towards push services in the other field studies as well (Kaasinen *et al.* 2003b, Kolari and Virtanen 2003, Södergård *et al.* 2003).

The users expected that push services would provide them with personally and contextually relevant information (Kaasinen *et al.* 2003a, Kolari and Virtanen 2003, Södergård *et al.* 2003). The test users had no experience of mobile spam due to the strict privacy legislation in Finland. The situation may be different in countries with more permissive legislation, where the users already have seen the negative consequences of mobile spam, e.g. in Japan.

In the Levi ski resort case, the information will be provided by several local service providers – restaurants, shops, ski rental services etc. The service providers were discussing about common rules for providing push services, e.g. how many messages per day can each actor send and how to utilise common user profiles (Kaasinen *et al.* 2003a). Without these rules, no service provider would benefit: the users could easily be inundated with commercial messages. In a restricted area such as the ski resort, the service providers already had experience of similar cooperation. With other types of guides, the cooperation might need external activation.

User attitudes create good potential for push services, but the services should be implemented with care so that the users will feel, and really be, in control. Otherwise the attitudes of the users may soon change completely. The user should get clear feedback on the push services: which services s/he is registered to and how probable messages are. The user should be able to easily define which topics s/he is sufficiently interested in to be willing to receive push messages. The user should be able to fine-tune or cancel the push feature easily, ideally as s/he receives a push message.

6.3 Occasional usage sets high requirements for reliability

In field study 3, different services and databases were integrated into a comprehensive service entity. This complex service architecture caused problems in error situations because often it was almost impossible to analyse and indicate to the user the reason for the error (Kaasinen *et al.* 2003b). At times, the service entity did not work in an occasional situation where it would have been very useful to the user, e.g. when searching for nearby hotels after a car breakdown in the middle of the night. These kinds of unsuccessful usage experiences discouraged the users from using the service again. Users seem to set even higher service reliability requirements for occasionally used services than they do for more frequently used services.

One of the ideas in the skiing resort guide was to use a push service to inform the users about storm warnings (Kaasinen *et al.* 2003a). More traditional warning systems should be maintained in parallel with the introduction of such systems. It should not be assumed that all people would have the mobile services in use or their mobile devices with them.

People may become quite dependent on location-aware guidance services. The users should be provided with information about the reliability of the service so that they can assess whether they can rely on the service in different usage situations. It is also important to get feedback on the provider of the service and the timeliness of the information provided (Kaasinen *et al.* 2003, Södergård *et al.* 2003).

7. Experiences of the evaluation methods

Mobile guides aim at connecting information with physical places. They impact on and are used in specific physical environments, bringing space and place together in activity (Brown and Perry 2002). Because of the strong connection with physical environments, only very limited simulations of usage situations can be made in laboratory evaluations. That is why in most of our case studies we replaced laboratory evaluations with short-term field evaluations. The set-up of the evaluations resembled laboratory evaluations but the evaluations took place in real usage environments. One evaluator conducted the test and walked with the user. Another evaluator walked a bit behind as an observer. As the test user was moving during the test, we were using audio recording with a portable minidisk system. The observer took photos of key usage situations. All evaluations took place outdoors, making organisation of the evaluations quite challenging. We had to be prepared for different weather conditions and to remind the test users to have warm clothes and good shoes because the evaluation included quite long periods of standing still and operating the mobile guide.

Most mobile guides are targeted at visitors rather than local people and that is why we had to recruit test users who had to travel to the test site. We decided not to try to recruit people who were already visiting the holiday resort or other test sites because people on holiday are often very busy with their schedule and not willing to use their time for anything extra. Travelling to the test site required extra time from the users. Luckily, we also had test users who appreciated the free train tickets to the test site as a pleasant bonus. As it was quite laborious to get the test users in place, we made several backup plans for the evaluation sessions to make sure that we could get the evaluation done. For bad weather, we had alternative routes or extra coffee/chocolate breaks, technical staff were on duty in case of technical problems and we also had with us spare devices, spare batteries and spare evaluation tools. The only thing that we could not prepare for was test users not showing up. This happened a couple of times and there were also a few last-minute cancellations. These incidents resulted in lost working hours but as we had prepared with a couple of test users more than needed, the cancellations did not postpone the total evaluation schedule.

In field studies 2 and 3, we were using long-term field evaluations. As described earlier, the usage needs for mobile guides turned out to be quite occasional. In the 3–4 week field trials with 7–10 users, we captured quite a few real usage situations (Kaasinen *et al* 2003b, Rämä 2003). Setting up systems that were reliable enough to be used in field trials required quite a lot of implementation effort. We were short of time but in other similar evaluations it should be considered whether the evaluation could last longer so as to get the most out of the implemented test set-up.

In long-term field evaluations, log files were useful in collecting and analysing usage information. In field study 3, we had online access to the log files during the trial. In this way, we could monitor the usage and, for instance, contact users who had not been using the service for a while and check to see if they had any problems with the service (Kaasinen *et al.* 2003b). Unfortunately, our test set-ups did not enable logging of malfunctions and error situations. It would have been useful to get exact data on these situations for comparison with user impressions.

All of our field studies were independent activities. It would have been useful to plan the evaluation activities more as an entity. In that way, we would have been better able to compare and combine the results.

8. Conclusions

In environments such as public places, city areas, holiday resorts, sport fields and event venues, the context of use can often be predicted according to user location. Location-aware mobile guides have good potential in these fields. The users may need location-aware mobile guides only occasionally, but in these occasional situations the services may turn out to be extremely useful. Occasional use indicates that designers must give careful thought to how the users will get information on the existence of the locally available guidance services and their characteristics, and how they can easily install, take into use and then use the services.

Momentary usage sessions on the move set high requirements for ease of use. In particular, text input should be made as effortless as possible by ready-made selection lists, text prediction and location-awareness itself in user interfaces. Alternative input methods are needed according to the user and the usage situation. Maps are essential both as complementary information and as user interface elements that associate information and functions with certain locations. Making maps personalised and context-aware can further increase their usability. Current technologies could not provide the users with unambiguous route guidance. This will require more accurate positioning methods and continuous orientation sensing.

The users turned out to be quite trusting towards service providers in issues related to privacy protection and reliability of the services. User acceptance of push services was high. These user attitudes form a good basis for introducing location-aware mobile guides. The trust of the users should be respected. Push services should be designed carefully so that the user feels, and really is, in control. The trade-offs between privacy protection and effortless use need to be resolved for each individual service. Legal and ethical assessments should be included as integral parts of the design to make sure that the user's privacy will not be threatened when collecting, using and storing location information.

To achieve a critical mass of users, location-aware guidance services should be offered on standard PDA and mobile phone platforms and they should be available irrespective of the network and positioning system in use. To achieve a critical mass of content, service entities should support the user throughout his/her activity in locating him/herself, looking for nearby services and getting route guidance. Realising comprehensive service entities requires the co-operation of content and service providers such as directory services, map providers, telecom operators and location providers.

Acknowledgements

Our research work was carried out with the financial support of the Personal Navigation (NAVI) Programme, hosted by the Ministry of Transport and Communications of Finland. The author wishes to thank the members of the NAVI management board for their valued co-operation: Benefon, Digia, Fortum, Genimap, Hewlett-Packard, ICL Invia, Nokia Mobile Phones, Novo Group, Radiolinja, Satama Interactive, Sonera, Suunto, TietoEnator, the Ministry of Transport and Communications, the Ministry of Agriculture and Forestry, the Ministry of Education and Culture, the Ministry of Defence, the Ministry of the Interior, the Ministry of Social Affairs and Health, the Finnish National Fund for Research and Development (SITRA), the National Technology Agency (TEKES) and the Finnish National Road Administration as well as Finnish Road Enterprise and the SME enterprises of the NAVI network. The author would also like to thank our excellent project team and especially Ari Ahonen, Veikko Ikonen, Juha Kolari, Minna Kulju, Annu-Maaria Nivala, Pirkko Rämä, Rolf Södergård, Raisa Suihkonen, Teija Vainio and Tytti Virtanen who had key roles in carrying out the field studies.

References

Ackerman, M., Darrel, T. and Weitzner, D.J., 2001, Privacy in context. *Human–Computer Interaction,* **16,** 167–176.

Ancona, M., Locati, S. and Romagnoli, A., 2001, Context and location aware textual data input. In *Proceedings of the 2001 ACM symposium on Applied computing*, Las Vegas, Nevada, USA, 11–March 2001, ACM Press. pp. 425–428.

Barkhuus, L. and Dey, A., 2003, Is context-aware computing taking control away from the user? *UbiComp 2003 Conference*, Seattle, Washington, 12–15 October 2003, Springer, pp. 150–156.

Barnes, S.J. and Huff, S.L., 2003, Rising sun: iMode and the wireless Internet. *Communications of the ACM*, **46,** 78–84.

Benford, S., anastasi, R., Flintham, M., Drozd, A., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M. and Row-Farr, J., 2003, Coping with uncertainty in a location-based game. *Pervasive Computing,* July–September 2003, 34–41.

Bieber, G. and Giersich, M., 2001, Personal mobile navigation systems – design considerations and experiences. *Computers and Graphics*, **25,** 563–570.

Bornträger, C., Cheverst, K., Davies, N., Dix, A., Friday, A. and Seitz, J., 2003, Experiments with multi-modal interfaces in a context-aware city guide. In *Mobile HCI 2003 Conference*, Udine, Italy, 8–11 September 2003, Springer, pp. 116–130.

Brown, B. and Perry, M., 2002, Of maps and guidebooks. Designing geographical technologies. In *DIS2002 Conference*, London, UK, 25–28 June 2002, ACM Press, pp. 246–254.

Chae, M. and Kim, J., 2004, Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones. *Behaviour and Information Technology*, **23,** 165–185.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C., 2000, Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI 2000 Conference*, The Hague, The Netherlands, 1–6 April 2000, ACM Press, pp. 17–24.

Cheverst, K., Mitchell, K. and Davies, N., 2002, Exploring context-aware information push. *Personal and Ubiquitous Computing,*, **6,** 276–281.

Davis, F.D., 1989, Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, **13,** 319–339.

Dey, A.K., 2001, Understanding and using context. *Personal and Ubiquitous Computing*, **5,** 20–24.

Fano, A. E., 1998, Shopper's eye: using location-based filtering for a shopping agent in the physical world. In *Proceedings of the second international conference on Autonomous agents*, Minneapolis, Minnesota, United States, 10–13 May 1998, ACM Press, pp. 416–421.

Fano, A. E., 2001, What are the location's 'File' and 'Edit' menus? *Personal and Ubiquitous Computing*, **5,** 12–15.

Fithian, R., Iachell, G., Moghazy, J., Pousman, Z. and Statsko, J., 2003, The design and evaluation of a mobile location-aware handheld event planner. In *Mobile HCI 2003 Conference*, Udine, Italy, 8–11 September 2003, Springer, pp. 145–160.

Fogg, B.J. and Tseng, H., 1999, The elements of computer credibility. In *CHI 99 Conference*, Pittsburgh, Pennsylvania, USA, 15–20 May 1999, ACM Press, pp. 80–87.

Grudin, J. and Palen, L., 1995, Why groupware succeeds: discretion or mandate? *ECSCW95 Conference*, Stockholm, Sweden, 11–15 September 1995. http://www.ics.uci.edu/~grudin/Papers/ECSCW95/ECSCW.html

Kaasinen, E., 2003, User needs for location-aware mobile services. *Personal and Ubiquitous Computing,* **7,** 70 – 79

Kaasinen, E., Ikonen, V., Sotamaa, O. and Vainio, T., 2003a, Location-aware and personalised tourist information. In *Products and services for Personal Navigation – Usability Design. Part III. Case Studies and Guidelines*, E. Kaasinen (Ed.), pp. 70 – 75 (Publications of the NAVI programme), available online at: http://www.navi-ohjelma.fi

Kaasinen, E., Ikonen, V. and Suihkonen, R., 2003b, NAVISearch – an integrated location-aware service directory. In *Products and Services for Personal Navigation – Usability Design. Part III. Case Studies and Guidelines*, E. Kaasinen (Ed.), pp. 131 – 154 (Publications of the NAVI programme), available online at: http://www.navi-ohjelma.fi

Kaikkonen, A. and Roto, V., 2003, Navigating in mobile XHTML application. *CHI 2003 conference*, Fort Lauderdale, Florida, USA, 5 – 10 April 2003, ACM Press. pp. 329 – 336.

Kolari, J. and Virtanen, T., 2003, In the zone: views through a context-aware mobile portal. *Mobile HCI 2003 Workshop HCI in Mobile Guides*.Udine, Italy, 8 September 2003. http://www.mguides.info/

Kulju, M. and Kaasinen, E., 2002, Route guidance using a 3D city model on a mobile device. *Mobile HCI 2002 Workshop on Mobile Tourism Support Systems*. Pisa, Italy, 17 September 2002. http://www.mguides.info/

Nivala, A.-M., Sarjakoski, T., Jakobsson, A. and Kaasinen, E., 2003, Usability of topographic maps on mobile devices. *21st International Cartographic Conference*. Durban, South Africa, 11 – 16 August 2003. http://www.icaci.org/

Patterson, C.A., Muntz, R.R. and Pancake, C.M., 2003, Challenges in location-aware computing. *IEEE Pervasive Computing*, April – June 2003, 80 – 89.

Rainio, A. (Ed.), 2000, Personal Navigation NAVI Programme 2000 – 2002. VTT Research Notes 2038, available online at: http://www.navi-ohjelma.fi

Rämä, P., 2003, Mobile information on weather and road conditions – a user study. In *Products and Services for Personal Navigation – Usability Design. Part III. Case Studies and Guidelines*, E. Kaasinen (Ed.), pp. 68 – 70 (Publications of the NAVI programme), available online at: http://www.navi-ohjelma.fi

Sarker, S. and Wells, J.D., 2003, Understanding mobile handheld device use and adoption. *Communications of the ACM,* **46,** 35 – 40.

Shneiderman, B., 2000, Designing trust into online experiences. *Communications of the ACM,*, **43,** 57 – 59.

Södergård, R., Kulju, M. and Ikonen, V., 2003, Usability test of three location-based SMS-services. In *Products and Services for Personal Navigation – Usability Design. Part iii. Case Studies and Guidelines*, E. Kaasinen (Ed.), pp. 52 – 67 (Publications of the NAVI programme), available online at: http://www.navi-ohjelma.fi

Teo, T.S.H. and Pok, S.H., 2003, Adoption of the Internet and WAP-enabled phones in Singapore. *Behaviour and Information Technology*, **22,** 281 – 289.

Published by

VTT

Series title, number and report code of publication

VTT Publications 566
VTT–PUBS–566

Author(s)
Kaasinen, Eija

Title

User acceptance of mobile services – value, ease of use, trust and ease of adoption

Abstract

This dissertation studies user acceptance of mobile services based on a series of case studies of mobile Internet and location-aware information services targeted at consumers and accessed mainly by mobile phones. The studies were carried out in connection with technology research projects, where the objectives were to develop technical infrastructures, to pilot the technology in demonstrator services, and to collect user feedback in field studies to assist in focusing future commercial deployment of the technology. A Technology Acceptance Model for Mobile Services is proposed on the basis of the results of the case studies. According to the model, user acceptance is built on three factors: perceived value of the service, perceived ease of use, and trust. A fourth user acceptance factor: perceived ease of adoption is required to get the users from intention-to-use to actual usage. Based on the Technology Acceptance Model for Mobile Services, design implications for each user acceptance factor are proposed.

Instead of implementing collections of useful features, the design of mobile services should be focused on key values provided to the user. The value of mobile services can be built on utility, communication or fun. Successful service content is comprehensive, topical and familiar, and it includes personal and user-generated content. The users appreciate seamless service entities rather than separate services. Ease of use requires a clear overview of the service entity, fluent navigation on a small display, and smooth user interaction with the service. The users should get personally and situationally relevant services and information without needing to expend effort on personalisation. The services should be designed to be adaptive to the wide variety of devices and networks. As the services increasingly support individual users in their daily tasks and the services are increasingly dealing with personal data, user trust in the services is becoming more and more important. The user should be able to assess whether (s)he can rely on the service in the intended contexts of use. The user needs to feel and really be in control, and the privacy of the user must be protected.

Occasional usage and momentary usage sessions on the move are typical of mobile services. In addition, services are increasingly available only locally or in certain contexts of use. This indicates the need for disposable services: services that are easy to find, take into use, use and get rid of when no longer needed. The user needs realistic information about the actual values of the services, so that (s)he can realise how to utilise the service in his/her everyday life and to innovate new usage possibilities.

Designing mobile services that will be accepted by users calls for changes also to the design methods. Value-centred design aims to identify the key targeted value(s) for end users, service providers and other stakeholders, and then focuses on these values in the design. The proposed Technology Acceptance Model for Mobile Services provides a tool to communicate key user acceptance factors and their implications to the design.

Keywords

mobile services, user acceptance, Technology Acceptance Model, human-centred design, mobile Internet, location-based services, value, ease of use, usability, trust, ease of adoption

Activity unit

VTT Information Technology, Sinitaival 6, P.O.Box 1206, FI–33101 TAMPERE, Finland

ISBN			Project number
951–38–6640–8 (soft back ed.) 951–38–6641–6 (URL:http://www.vtt.fi/inf/pdf/)			
Date	Language	Pages	Price
May 2005	English	151 p. + app. 64 p.	E
Name of project		Commissioned by	
Series title and ISSN		Sold by	
VTT Publications 1235–0621 (soft backed.) 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)		VTT Information Service P.O.Box 2000, FI–02044 VTT, Finland Phone internat. +358 20 722 4404 Fax +358 20 722 4374	

This dissertation discusses user acceptance of mobile services based on a series of field trials and other user evaluation activities. The services studied include mobile Internet and location-aware information services targeted at consumers and accessed by mobile phones. The case studies were carried out in connection with technology research projects at VTT in the years 1998–2002. A Technology Acceptance Model for Mobile Services is proposed on the basis of the synthesized results of the studies. According to the model, user acceptance is built on three factors: perceived value of the service, perceived ease of use, and trust. A fourth user acceptance factor, perceived ease of adoption, enables users to move from intention to use to actual usage. Based on the proposed model, design implications are suggested for future services.

Tätä julkaisua myy	Denna publikation säljs av	This publication is available from
VTT TIETOPALVELU PL 2000 02044 VTT Puh. 020 722 4404 Faksi 020 722 4374	VTT INFORMATIONSTJÄNST PB 2000 02044 VTT Tel. 020 722 4404 Fax 020 722 4374	VTT INFORMATION SERVICE P.O.Box 2000 FI–02044 VTT, Finland Phone internat. +358 20 722 4404 Fax +358 20 722 4374

ISBN 951–38–6640–8 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–6641–6 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)

VTT PUBLICATIONS 566 User acceptance of mobile services – value, ease of use, trust and ease of adoption Eija Kaasinen