FIDIS

Future of Identity in the Information Society

Title:	"D3.13: Study on Usability of IMS"
Author:	WP3
Editors:	John Sören Pettersson (KU, Sweden)
	Martin Meints (ICPP, Germany)
Reviewers:	Mark Gasson (University of Reading)
	Dennis Gessner (SIRRIX)
Identifier:	D3.13
Type:	[Deliverable]
Version:	1.0
Date:	Friday, 14 August 2009
Status:	[Final]
Class:	[Public]
File:	fidis D3 13 final 1.0.doc

Summary

Usability is an important factor influencing the acceptance of technology-related products and subsequently the decision process to use or buy them. For this reason usability has been of interest in market research for many years. However, in the context of user-controlled Identity Management Systems (IMS) hardly any comparative usability studies have been published.

This document gives an overview of established evaluation methods and criteria for usability and analyses which methods and criteria are suited for user controlled IMS. The selected methods and criteria are applied to twelve IMS in six functional classes. These classes include user controlled identifier management, policy management, form filling, context monitoring and history management. Nine of these IMS are further analysed in depth and the results of the tests are published in this report.

The FIDIS NoE receives research funding from the Community's Sixth Framework Program

Copyright Notice:

PLEASE NOTE: This document may change without notice – Updated versions of this document can be found at the FIDIS NoE website at www.fidis.net.

Members of the FIDIS consortium

1. Goethe University Frankfurt (JWG)	Germany
2. Joint Research Centre (JRC)	Spain
3. Vrije Universiteit Brussel	Belgium
4. Unabhängiges Landeszentrum für Datenschutz (ICPP)	Germany
5. Institut Europeen D'Administration Des Affaires (INSEAD)	France
6. University of Reading	United Kingdom
7. Katholieke Universiteit Leuven	Belgium
8. Tilburg University[1]	Netherlands
9. Karlstads University	Sweden
10. Technische Universität Berlin	Germany
11. Technische Universität Dresden	Germany
12. Albert-Ludwig-University Freiburg	Germany
13. Masarykova universita v Brne (MU)	Czech Republic
14. VaF Bratislava	Slovakia
15. London School of Economics and Political Science (LSE)	United Kingdom
16. Budapest University of Technology and Economics (ISTRI)	Hungary
17. IBM Research GmbH	Switzerland
18. Centre Technique de la Gendarmerie Nationale (CTGN)	France
19. Netherlands Forensic Institute (NFI)[2]	Netherlands
20. Virtual Identity and Privacy Research Center (VIP)[3]	Switzerland
21. Europäisches Microsoft Innovations Center GmbH (EMIC)	Germany
22. Institute of Communication and Computer Systems (ICCS)	Greece
23. AXSionics AG	Switzerland
24. SIRRIX AG Security Technologies	Germany

[1] Legal name: Stichting Katholieke Universiteit Brabant
[2] Legal name: Ministerie Van Justitie
[3] Legal name: Berner Fachhochschule

Versions

Version	*Date*	*Description (Editor)*
0.1	13.05.2008	• Initial release (Martin Meints), Contribution of TUD integrated (chapter 3)
0.2	02.06.2008	• Chapter structure revised and results of the preliminary scanning integrated (KU, chapters 4 and 9), minor editorial changes throughout the deliverable (ICPP and TUD)
0.3	10.07.08	• Integrative editing throughout the deliverable
0.4	09.09.2008	• Methodology integrated (Stefanie Pötzsch, chapter 3), minor editorial changes and additional subsection in Usability Metrics (Stefanie Pötzsch, chapter 4)
0.5		• John Sören added PET-USES paragraph and used methodology (not in detail about the usability studies). Deleted unfinished section i 5.4 (leaving only table).
0.6	09.10.08	• Structure finalised, comments resolved
0.7	22.01.2009	• John Sören cont. changing ch 6: • 6.2.3 inserted (new section numbering in 6.2) • 6.3 completely replaced by a version from 31 October – Martin needs to check this section.
0.8	06.03.09	• Contribution JWG on the economic relevance of usability integrated in the introduction
0.8.1	13.03.2009	• John Sören making correction, esp in Appendix (the IMS scan).
0.8.2	23.03.2009	• John Sören: Summary added
0.9		• Executive summary added, editorial changes throughout the deliverable
0.91		• John Sören went through Mark Gasson's review remarks.
0.92	14.08.2009	• John Sören made final editing from second reviewer.

Foreword

FIDIS partners from various disciplines have contributed as authors to this document. The following list names the main contributors for the chapters of this document:

Chapter	*Contributor(s)*
1 (Executive Summary)	All contributors
2 (Introduction)	Martin Meints (ICPP), Denis Royer (JWG)
3 (Methodology)	Stefanie Pötzsch (TuD)
4 (Usability Metrics)	John Sören Pettersson (KU), Stefanie Pötzsch (TuD)
5 (Selection of the IMS)	All contributors
6 (Usability Testing)	Maria Lindström, Nina Rönntorp, John Sören Pettersson (KU)
7 (Summary and Conclusions)	John Sören Pettersson (KU), all contributors
8 (Bibliography)	All contributors
9 (Annex 1: Indecies)	All contributors
10 (Annex 2: Pre-study)	Maria Lindström (KU)

Table of Contents

1 Executive Summary

Since the late 1990s models for understanding technology acceptance by the users were developed for computer applications. In these models perceived ease of use became a relevant factor. Perceived ease of use relates to the degree a user believes a given technology will be free from notable effort when being used. To this regard, usability of a technology is essential. This deliverable takes a look at the usability of user-controlled and user-centric identity management systems (IMS). This kind of IMS are called 'type 3' IMS in a typology formulated by Bauer, Meints, and Hansen (2005).

Firstly, methods for the investigation of usability of products and scales for the quantitative analysis of user feedback are introduced and based on the characteristics of IMS usable methods and scales are selected. In the following chapter relevant usability metrics and criteria are introduced, based on the literature and standardised methods. Criteria for the testing of 'type 3' IMS are selected. Most relevant criteria in this context are:

- Comprehensibility
- 'Error preventivness'
- Main task suitability
- Perceived controllability
- Likeability

For testing, twelve user-controlled IMS taken from the FIDIS IMS database[4] were selected, based on the variety of functions they show. In a pre-test these IMS were analysed and clustered according to their identity management functionality. The following are functions provided by the IMS:

- Identity management (including management of 'roles' and 'partial identities')
- Authentication management
- Form filling
- Policy management
- Context monitoring
- History management (i.e., management of records of data disclosures and policy agreements)

The pre-testing also showed that some of these selected IMS lack the maturity to be tested by users, for example because of difficulties to compile and install them on platforms available for the testing. The following IMS were investigated more thoroughly in the pre-test and some were selected for user-based testing. Each had at least one of the mentioned identity management functions:

- Privacy policy alerting: *Privacy Bird*

[4] See imsdb.fidis.net

- Anonymizers: *Tor* and *JAP/An.ON*
- Web browsers: *Opera*, *Safari*, *Firefox*, and *Internet Explorer*
- Dedicated identity managers: *CardSpace* and two *PRIME* prototypes

In order to highlight some specific features, two comparative user-tests have been performed, each involving two systems, as will be described immediately below. The user tests involved a relatively small number of test persons (10 to 12 persons) and included various tasks to perform within specific scenarios.

Web browsers (tested were Internet Explorer and Mozilla Firefox) are important, as they include some identity management functionality. In addition browsers are largely distributed and used by experienced as well as non-experienced users. In this context the usability of identity management functions is of general interest to IMS developers as well as privacy experts and the general public.

The browser test shows that identity management functions in web browsers can support the performance of the main-tasks the users had in mind. Security notifications remained largely unnoticed by the users in both browsers tested. Several test participants liked the idea of privacy-enhancing software and could imaging paying at least a small fee for it. A number of concrete observations and possible recommendations for further improvements are listed in section 6.2.6. Concerning the selected criteria the conclusions are:

- *Comprehensibility*: The test participants understood the concepts in Firefox better than in Internet Explorer.
- *Error prevention* was not really tested by the tasks, even if some problems with Firefox are notable. Security indications went unnoticed in both browsers. Additional inspection of the browsers can be done to assess what amount of context help, meaningful warnings, and feedback is given before privacy errors occur.
- *Main task suitability*: Even if test participants were unused to relying on auto fill functions, there was not much that really obstructed what could have been their main tasks during an ordinary browsing session (when privacy tasks are not the main tasks).
- *Perceived controllability*: Participants perceived Firefox as an easily manageable browser; however, Firefox's automatic completion of login data raised discomfort.
- *Likeability*: Most test participants liked the idea of privacy protecting software so much that they said they were willing to pay at least a small amount for such a product.

The second test concerned history functions that help users to remember what data disclosures they have made earlier. Microsoft's newly released IMS 'CardSpace' was compared with a prototype from a research project (*DataTrack* from the PRIME project). It turns out that the existing product is very 'card-centric'. The users are not only supposed to release data via a card metaphor, but the release history can only be tracked via each individual card. The research prototype offers more general search functions. Test participants performed better in the commercial product, but a majority liked the more general history function in the research product. Again, almost all participants liked the idea of identity management functions. Section 6.3.7 summarises test results according to the usability criteria developed for this study. Concerning the selected criteria the conclusions are:

- *Comprehensibility*: CardSpace's user interface was generally easy to deal with but some things were hard to understand; e.g. the history of transactions has to be searched for in all the so-called InfoCards that are used to group information. It was also hard to understand what had been sent, and if something had been sent. PRIME's history function DataTrack contains more functions but even if users failed initially on more tasks than when using CardSpace, they could manage the added 'complexity' of more search options (see also 'Likeability' below).
- *Error prevention*: The test tasks did not provoke any direct errors when testers used CardSpace but see the caveat in the first bullet. For PRIME, some test participants asked for a Help file.
- *Main task suitability*: This category must be reinterpreted when evaluating IMS history functions because for a history function the main task is the same as the relevant IMS-related tasks. Test users thought the PRIME history function DataTrack was suitable for its task. In CardSpace, the user needs to open each card to see the history of it. Also there is no search function inside the cards, so the user needs to go through every transaction if he or she does not remember when the information was sent or who it was sent to. Thus the CardSpace history function is harder to use and does not support identity management in the same way as the PRIME history function.
- *Perceived controllability*: CardSpace saves a lot of information about a data transaction. Thus, the user has in principle control over what data has been sent and who has received it. However, the test and post-test questions showed that people did not think they had good command over the history records. The PRIME history function also saves a lot of information about a data transaction. However, the test users felt more in control over the history function because of the search functionality within the DataTrack.
- *Likeability*: According to the post-test questions 3 of the 12 participants preferred CardSpace over PRIME while 8 of the 12 test participants preferred PRIME over CardSpace. Notable is that 11 liked the idea of identity management systems.

2 Introduction

In product development and design it has been realised that aspects of usability play an important role for many decades now. In the context of information and communication technology (ICT) related products usability increasingly became an area of research in the early 1980s, resulting in numerous scientific publications and even international standards. In addition, the hard- and software industry is an active stakeholder in the context of usability, as usability from an economic point of view is very important.

One of the success factors for multilayered and complex technologies, such as IMS, is the acceptance by the users of the technology itself. Looking at this, usability (not only with respect to IMS) can be considered an integral component. Davis (1989) formulated this notion in his Technology Acceptance Model (TAM), being composed of the two main factors: *perceived ease of use* and *perceived usefulness*. To this regard, the latter factor (perceived usefulness) relates to the degree to which a person believes that using a particular system will enhance his or her performance. Looking at IMS, a Single-Sign-On (SSO) system would enhance a user's performance by reducing the number of credentials being necessary or the process-time for log-ins, adding to a user's perception of the SSO system's usefulness.

Changing the perspective towards the first factor of the TAM (perceived ease of use), this factor relates to the degree a user believes a given technology will be free from effort when being used. To this regard, usability of a technology is essential. Looking at IMS, factors such as the comprehensibility, perceived controllability, or learnability are components to be considered when analysing IMS, as they influence the perception of their users. Among others, these components are discussed in this study (Section 4.2).

Extending the model of Davis, factors such as perceived controllability can be added as additional constructs (e.g. alternative belief factors) or for examining antecedents and moderators of perceived usefulness and perceived ease of use. By doing this, in-depth knowledge about the usage of IMS can be aggregated, analysed, and formulated into models and frameworks for enhancing their functionality. Ultimately this also leads to higher acceptance, especially with regard to complex technologies such as IMS, and opens opportunities for a greater diffusion of such technologies into a given market.

Focussing on the mechanisms that foster the diffusion of technologies, the Diffusion of Innovations (DoI) theory by Rogers (2003) can be taken as reference. The theory itself describes the process by which an innovation is communicated through certain channels over time among the members of a social system. In other words, the study of the diffusion of innovation is the study of how, why, and at what rate new ideas and technology spread through cultures[5].

Besides the adopters and the stages of adoption, several important characteristics for innovations can be observed. The first characteristic is the degree to which the innovation is perceived as being better than the practice it supersedes. This is referred to as the *relative advantage*. Next, the *compatibility* describes the extent to which adopting the innovation is compatible with what people do. Furthermore, the *complexity* is the degree to which an innovation is perceived as relatively difficult to understand and use. Moreover, the *trialability*

[5] For details on the Diffusion of Innovations theory → cf. FIDIS deliverable D11.3, pp.49.

focuses on the degree to which an innovation may be experimented with on a limited basis before making an adoption (or rejection) decision. Finally, the *observability* describes the level to which the results of an innovation are visible to others.

Looking at these key characteristics for innovations and applying them to IMS, the link to usability as an important factor for the adoption of IMS in general becomes visible – especially with regard to the following characteristics: IMS should be designed in a way that makes them *compatible* with the way people work. Specifically, IMS should be integrated into existing processes or technologies which can use them, so users do not need to integrate them on a separate basis. This notion also relates to *complexity* as IMS should be built in a way that (almost) everyone can use them with a limited amount of technical skills, also making an IMS *trialable* for the focussed user group. Also being important for fostering adoption, the usage of an IMS should be observable, so non-users can easily see the benefit of using IMS and finally experience the *relative advantage* for using IMS, such as SSO, higher security or privacy.

So, by closely observing the characteristics of innovations and their resulting effects, models for successful user adoption of an IMS can be developed. To this regard, usability seems to play an important role for the adoption and the user acceptance, in order to make identity management technologies a success. However, in the context of user controlled identity management tools and systems ('type 3' IMS, see Bauer, Meints, Hansen (2005) for the typology) test studies have not become widely known, and comparative usability studies have become available only recently.

This document gives an overview on established evaluation methods and criteria for usability, and analyses which methods and criteria are suited for user controlled IMS. In this document a selection of 12 IMS is taken from the IMS database[6] and tested with respect to usability. The selected methods and criteria are applied in six functional classes. These classes include user controlled identifier management, policy management, form filling, context monitoring and history management. The proceedings and the results of this testing are summarised in this document.

As background to the testing chapter, an overview on established methods for the evaluation of usability of (ICT) products is given in chapter 3. The selection of the methods used for the IMS testing is described. In chapter 4 criteria used in the usability testing of ICT and security related products are presented and the selection of criteria suitable in the context of 'type 3' IMS is described. In chapter 5 the results of analysing a number of systems with identity management functions such as widely used internet browsers, innovative identity management tools such as anonymisation services and history management tools, etc., are presented. Chapter 6 reports two usability tests focusing two different aspects: IMS background work during web browsing and IMS foregrounding when the user needs to know what he/she has done during earlier web sessions. The procedures in detail and the results of the pre-testing are documented in Annex 2 (chapter 10). Chapter 6 summarises the results of the testing, followed by the conclusions in chapter 7.

[6] The FIDIS database on identity management systems (IMS database) is publicly available at http://imsdb.fidis.net

3 Methodology

Usability testing aims to evaluate products and applications, e.g. software systems, and indicates suggestions for improvement. Results and conclusions of usability tests are based on qualitative and/or quantitative data and their statistical analysis. Usability testing combines technical development and approaches from social science in order to design products that are usable for the intended target group.

In the following, a classification of methods for usability tests is given. There are dependencies between some of the classes.

- **Considering Time: Formative Testing vs. Summative Testing**

A *formative test* aims at quality assurance. Therefore the primary goal is an identification of weaknesses during the development process of a system. Results from a formative usability test are used as input for improvements in further steps of the system's development (Tergan, 2000).

Summative tests are applied to assess and compare quality, impact and usefulness of systems which are already in use. It is tested whether the system fulfils defined expectations in practice (Tergan, 2000).

- **Considering Objectives: Explorative Tests vs. Assessment Tests vs. Comparative Tests**

Explorative tests are applied in early stages of the development process. The aim is to learn about thinking processes of users and the mental models that they have of a given system design. Hence, it can be checked to which extent both models match.

If a new system was designed, an *assessment test* can be done in order to check how well users can perform tasks and to find any deficits with regard to usability. Assessment tests focus more on user's performance than on thinking processes.

In *comparative tests* a system is compared with competing products in order to identify strengths and weaknesses with regard to usability. This kind of tests can be conducted in each stage of the development cycle.

- **Considering Scope: Full Tests vs. Tests of single modules/features**

In *full tests* the complete system is checked. For very comprehensive systems with a lot of features such tests can become especially complex and time consuming for the test user.

Tests of single modules focus only on a few features of system which may be comparable to other systems and should be re-engineered.

- **Considering Environment: Laboratory Tests vs. Field Tests**

The goal of *laboratory tests* is the measurement of the system's usability under defined conditions. Test users are confronted with typical tasks and observed using usability lab equipment. Independent variables can be easily monitored and interfering variables can be controlled. However, situations in

such test are unfamiliar for the participants and may influence their behaviour. Consequently, results from laboratory tests cannot be transferred easily to actual environments where the system is used.

Field tests provide the advantage of an authentic surrounding for the participants of the usability test, thus more natural behaviour is expected and results could be more realistic. However, it is impossible to know and control all interfering variables. Hence, it is more difficult to find causes of and explanations for usability problems which have occurred during field tests.

- **Considering Participants: Experts vs. End-users**

Expert-based usability tests, also called inspections, involve experts "examining" a system with reference to checklists or standards. These methods are efficient especially for testing in early stages of a system's development process, since expert-based tests offer a good trade-off between costs and the number of usability problems found.

Usability tests with *end-users* are usually more expensive in preparation. However these methods provide feedback from real users and help to identify actual usability problems of the target group of a system.

- **Considering Objectiveness: Objective Measures vs. Subjective Measures**

Objective measures such as time to complete a task, and number of errors can be collected indirectly during usability tests by observation and/or logging of users' interactions with the system.

Subjective measures collect the participants' opinion about a system, e.g. from questionnaires or interviews.

3.1 Methods for Expert-based Usability Testing

3.1.1 Cognitive Walkthrough

A *cognitive walkthrough* is a method for usability inspection where experts act on behalf of imagined users. Before the test starts, experts need to agree on the imagined users' profiles, their experience and the tasks which should be conducted. During performance of the tasks, experts estimate users' cognitive workload and put focus on the 'ease of learning' of the systems (Barnum 2002). "A cognitive walkthrough makes explicit the important design decisions that have been made either implicitly or explicitly in the process of designing an interface and the implications of those decisions for the problem-solving processes and learning by exploration." (Lewis et al. 1990)

This method allows investigation of usability issues early in the development process before the system is tested by real users. (Lewis et al. 1990) proved in their study that almost 50% of all usability problems which can be found during empirical tests with real end-users, can be identified using the less costly and less time-consuming method of cognitive walkthroughs. In order to be effective, experts who have a background in psychology and who are trained in cognitive walkthroughs are required.

3.1.2 Heuristic Evaluation

The method of *heuristic evaluation* is – as its name implies – based on heuristics. Experts inspect a system in order to find usability problems, which can be attended to in the next steps of development. Relying on former studies, (Nielsen 1993) concludes that only five experts are necessary to identify almost 75% of all usability problems. Hence, he suggests having three to five experts for a heuristic evaluation of a system is an optimal trade-off between the proportion of usability problems found and costs.

General usability principles that could be used for heuristic evaluation are for instance (Nielsen 1993, p. 20):

- Simple and natural dialogue,
- Speak user's language,
- Minimise the user's memory load,
- Consistency,
- Feedback,
- Clearly marked exits,
- Shortcuts,
- Good error messages,
- Prevent errors,
- Help and documentation.

Heuristic evaluation is a cheap method considering financial costs and time. This is especially an advantage when these resources are tight and qualitatively good results are needed. Since heuristic evaluation does not include end-users, they might have quite different usability problems in addition than these ones found by the experts. However, this shortcoming applies to each expert-based method.

3.2 Methods for User-based Usability Testing

User-based usability testing involves end-users from the intended target group. They may interact with the system in question considering given workflows and state their opinion by choosing between proposed options or users may be allowed to test the system on their own and formulate individual feedback statements. Figure 1 shows a plot, which gives an overview about methods for user-based usability testing. The vertical axis indicates the degree of users' liberty for feedback. The objectiveness of the collection of data is represented by the horizontal axis.

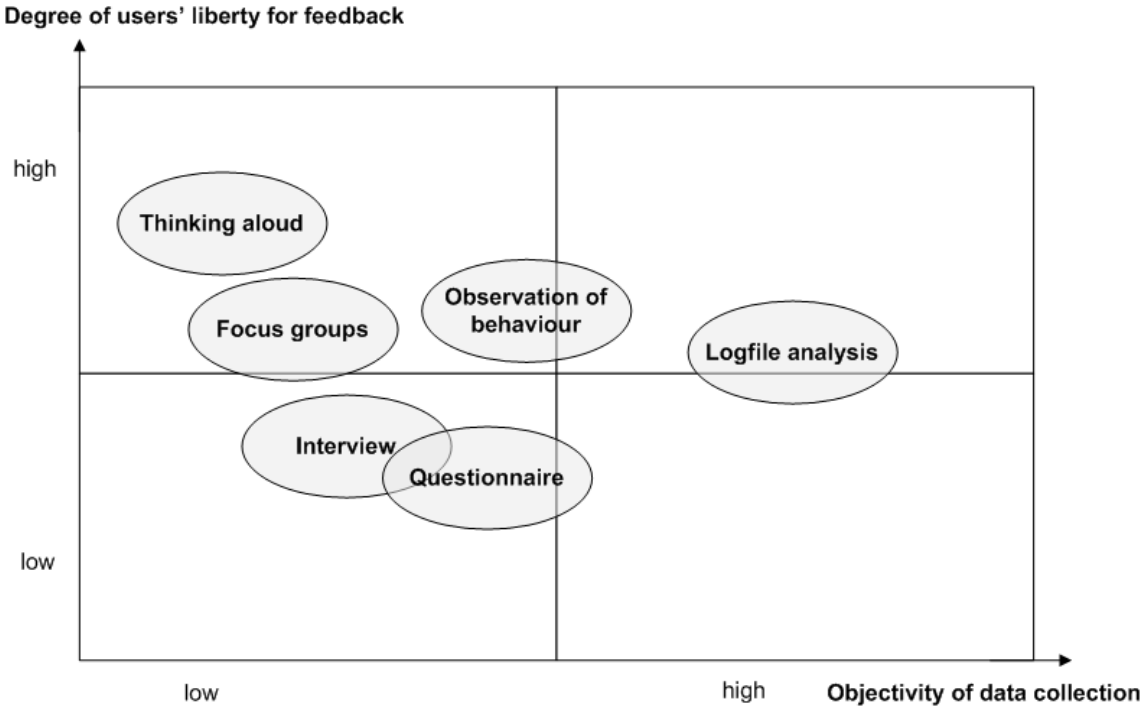


Figure 1: Methods for user-based usability testing (source: Görner, Ilg 1993)

A description of these methods, including their advantages and disadvantages, is presented in the following section.

3.2.1 Thinking Aloud

The thinking aloud method requires participants to verbalise their thoughts during interactions with the system. By speaking out their thoughts, users give insight in to how they view the system and what problems they have to work with it.

Since it is unusual to most participants to speak every thought out loud during interactions with technical devices, they need to be continuously prompted to talk, e.g. by questions such as 'What are you thinking now?' or 'Is that what you expected would happen?' from an experimenter. Due to the phenomena of 'social desirability', users may feel that they need to fulfil someone's expectations, which adulterates results from this method. However, thinking aloud is a good option to understand at which points of the system users have problems and misinterpret dialogues or other elements of the interface.

Further variations of thinking aloud are *constructive interaction*, *retrospective testing* and the *coaching method* (Nielsen 1993, pp. 198).

3.2.2 Observation

Observation of users can be done with support of technical means, e.g. by videotaping the users during the usability test or by a human observer who takes notes about users interactions with the system. The observer should not interfere in the user's workflow and not answer any requests for assistance. After the user test is finished, notes or videotapes are analysed.

Observing users can provide interesting feedback since they might find workflows and usability problems that developers would never have thought of and that therefore would not have been tested in experiments with detailed planned given tasks (Nielsen 1993).

A disadvantage can be the presence of the observer or video cameras as unnatural objects in the environment of users. They may feel controlled and therefore their behaviour is influenced by the test setting. Further, it is not easily possible to conclude reasons for users' behaviour from observations. If videos are recorded and need to be watched and analysed later, the timely effort should not be underestimated. Baber and Stanton (1996) assume "that one hour of recording could require around ten hours of analysis". The analysis has to be manual.

3.2.3 Logfile Analysis

Recording logfiles is an option of technically supported observation of users' interactions with the system that could be analysed in an automated way. A logfile contains information about system events and user input. Analyses may consider frequency of error messages, frequency of use of functions, number of steps that have been necessary in order to complete a task, etc. Eye tracking equipment also generates logfiles which can be easily summerised.

The advantages and disadvantages of observation methods in general, are also true for the analysis of logfiles (c.f. Nielsen 1993). From an ethical as well as from a privacy perspective, users need to be informed about the logging of their actions. Hence, even if no additional physical object (video camera, human observer) is installed, users are aware of the observation which may influence their behaviour.

3.2.4 Focus Groups

A focus group consists of a number of users – Nielsen (1993) suggests six to nine – who discuss concepts and identify usability problems. The discussion usually lasts about two hours and is run by a moderator who has a pre-planned script of what issues to bring up. However, the moderator should not inhibit the free flow of comments and ideas. Focus groups are intended to bring out users' spontaneous reactions and ideas and let the moderator observe some group dynamics and organisational issues. This method assesses what users think about a system. After the session, a report that sums up the discussion, the mood of the focus group and includes quotes from the participants is written by the moderator.

Demonstrations for a focus group of systems that do not exist yet can be made by oral presentation, user interface animations, and demos of interactive mock-ups.

Nielsen (1993) recommends to run more than one focus group discussion since the outcome of a single group may not be representative and involves the risk of overemphasising minor issues of the system. Further, this method collects users' thoughts about what they need instead of detecting real needs. In comparison to methods that require technical settings and/or detailed preparation of user tasks, focus groups are less costly.

3.2.5 Oral Interview

Oral Interviews are another option to gather users' impressions and opinions about a system as well as ask them about problems during a previous test with the system in question. Interviews can be used to collect quantitative as well as qualitative data. The variants of this method can be distinguished according to the following dimensions (cf. Bortz, Döring 1995, pp. 217-231).

- Extent of structuring
    - Fully structured: Order and exact wording of all questions is given and mandatory for the interviewer.
    - Semi-structured: The interviewer has a road map for the conversation with regard to content and way of asking; however no concrete questions are given.
    - Unstructured: The course of conversation depends on the skills of the interviewer. Merely the broad topic of the interview is given beforehand.
- Way of contact
    - Personally: Interviewer and interviewee meet in person for the conversation. This allows for control of the concomitant circumstances and the interviewer can record also non-verbal responses (e.g. gesture, behaviour).
    - Via telephone: Contacting the interviewee via telephone provides him/her with more perceived anonymity than a face-to-face conversation. However, it is not possible to work with pictures, graphics or extensive response options.
- Number of users involved
    - Single interview: Interviews with single persons are necessary for sensitive topics and in order to avoid responses that may otherwise be adulterated by group pressure.
    - Group interview/discussion: Group interviews are rather done in written form. An exception are group discussions in which the interviewer has the role of a moderator and give impulses to the discussion in order to direct its topic.

3.2.6 Questionnaire

Questionnaires are a special kind of interview since they are in written form. All participants get the same wording for the same question; however it cannot be ensured that everyone assigns the same meaning to a term. Further, with the help of written questionnaires it is possible to avoid an *interviewer bias*, i.e. participants are much less influenced by the persons who conduct the questionnaire than in oral interviews.

Besides these advantages, questionnaires also have shortcomings. Since there are no explanations from an interviewer, misunderstanding of terms and questions cannot be avoided. There is no possibility for tailored questions based on the interactive discussion in oral interviews. In addition, it is easier for participants to leave questions unanswered, thus data sets are incomplete. If standardised response options are used, there is no space for individual comments from the participants.

3.3 Scales for Quantitative Data Collection Methods

For collecting quantitative data from usability tests a number of scales exist. In the following we present the most popular scales. For more details we refer to (Bortz, Döring 1995). Once scales are developed for testing certain criteria for usability, these scales can be applied to different software systems in order to ensure comparability between these systems. If only a single system is to be tested, it is necessary to use normalised scales, i.e. the software system in question is compared to the average values of all systems that are considered during the

development of the scale. Hence, available normalised scales are already tailored to the aspects of usability that can be measured by using them.

3.3.1 Likert Scale

This type of scale was introduced by Rensis Likert and aims to cover attitudes, preferences and subjective reactions of test participants. Therefore clear statements are given and the user has to tick the degree to which he/she agrees or disagrees with the particular statements.

Once the statements are carefully worded, Likert scales are easy to set up and therefore often used. The number of items can be odd or even. The latter option is applied if a middle item (neither agree nor disagree) does not make sense for the statement or if participants should be pushed to take a decision. Figure 2 shows an example of a five point Likert scale.

I like this software.

( ) Strong agree

( ) Agree

( ) Neither agree nor disagree

( ) Disagree

( ) Strong disagree

Figure 2: Example of Five point Likert scale

3.3.2 Guttman Scale

Guttman scales are also used to measure users' attitudes towards facts and circumstances. Participants are confronted with several statements touching the same fact and should say whether they agree or disagree with the statement. The order of statements goes from the most basic step by step to the most extreme. An example is shown in Figure 3.

(1) I think it is important to take measures for protecting privacy.

(2) I think it is important to use software that supports protection of private data.

(3) I use software to protect my private data.

(4) I update my software for privacy protection regularly.

Figure 3: Example of Guttman scale

Participants who agree to statement (3) for instance, have also to agree to (1) and (2), if they give true responses. The critical point is to find the right wording for statements in a consistent order.

3.3.3 Thurstone Scale

The Thurstone scale measures users' opinion on a subject or circumstances by letting him/her agree or disagree on different statements related to the same subject or circumstances in question. Each of the statements is located beforehand by experts on an eleven point scale. After users have decided about each given statement, their personal opinion can be measured by the underlying expert scale. The advantage of this scale is that people are more used to agreeing or disagreeing on concrete statements than to assess their own opinion on a scale between zero and ten for instance. However, the critical point is to find appropriate wording for the statements and to validate the scale by the help of experts. Figure 4 shows an example of a Thurstone scale for usability testing.

The software had many convenient features.	*yes / no*
I always found what I wanted.	*yes / no.*
The software was hard to use.	*yes / no*
…	

Figure 4: Example of Thurstone scale

3.3.4 Semantic Differential Scale

Semantic differentials are a special type of rating scale. Users get a number of bipolar pairs of attributes and need to assess an object (e.g. software) by deciding for one or the other attribute on a seven point scale. AttrakDiff (c.f. AttrakDiff 2008) is an example of a semantic differential scale that is developed and applied for testing usability and joy of use. Figure 5 shows some of the semantic differentials used in AttrakDiff.

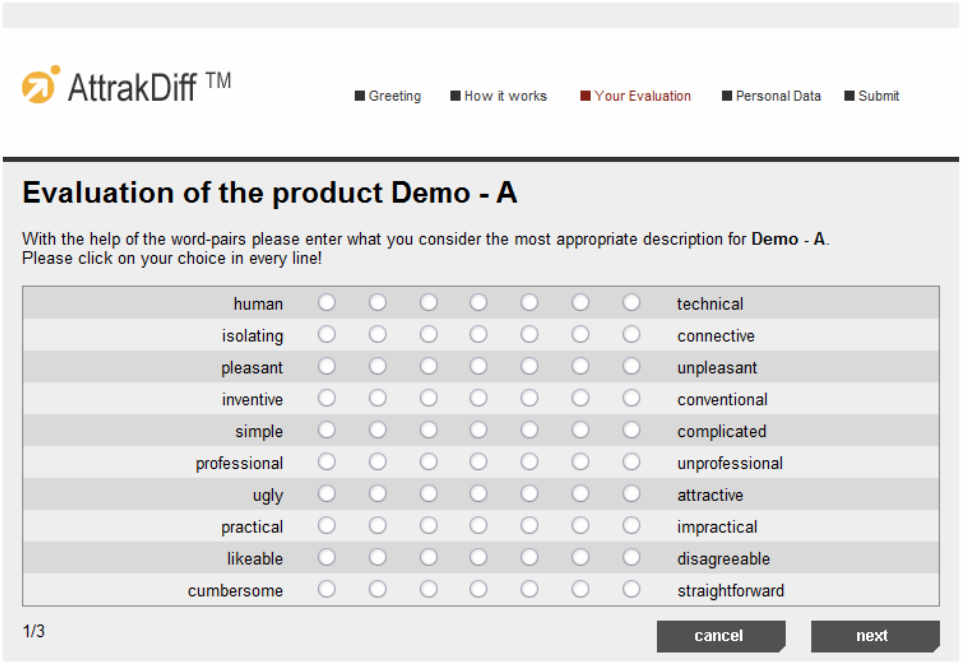


Figure 5: Example of semantic differential scale (source: AttrakDiff 2008)

3.3.5 PET-USES

PET-USES is a questionnaire presently being developed within the PrimeLife project. The name is an acronym for Privacy-Enhancing Technology Users' Self-Estimation Scale and covers such things as test participants perceived level of control and understanding of a system (i.e., a user-estimation of usability) as well as how they perceive the usefulness of the system's functions and, in general, if they like it.

3.4 Applied Methods for the Study

Table 1 summarises the methods, which are described in the previous sections. The column "Participant" is restricted to the "main informant"; there are naturally other people present (i.e. participating in some sense) in, for instance, a Think Aloud session or a Focus Group workshop. Also for the "expert" methods, there might be a test leader collecting protocols and conducting interviews with the experts; an analysis is also needed because the experts will in many cases not give the same comments and recommendations.

The word "users" is of course not meant as real users, when there is only a prototype tested. This word should in such cases be understood as "persons from the intended future user group"

Method	**Participants**	**Principle**	**Main Advantage**	**Main Disadvantage**
Cognitive Walkthrough	Experts	Performing tasks on behalf of imagined users	Efficient and cheap	Experts with background in psychology needed, no real users
Heuristic Evaluation	Experts	Inspection of system based on heuristics	Efficient and cheap	No real users involved
Thinking aloud	Users	Verbalisation of thoughts during interaction with system	Compare users thinking processes with system design model	Permanent verbalisation unnatural for users
Observation	Users	Observe interactions and reactions	Learn about users' workflows	Difficult to conclude reasons for behaviour
Log file analysis	Users	Logging of system events and user input	Automated analysis of quantitative data possible	May cause privacy issues, much raw data
Focus groups	Users	Moderated discussion of concepts	Spontaneous reaction, group dynamics	Not representative
Interview	Users	Asking users	Non-verbal	Interviewer bias

		about opinion and problems	reactions, Interviewer can react flexible on responses	
Questionnaire	Users	Asking users about opinions and problems in written form	Similar wording for all participants, automated analysis	Missunder-standing of wording, incomplete data sets

Table 1: Overview on established evaluation methods for usability

The present study of usability of IMSs consisted of two major parts:

1. Analysis of several systems to give there main characteristics as to their IMS functions and their user interaction;
2. Collecting/conducting usability studies of a few specific systems chosen on the basis of the overview given in 1.

3.4.1 Part 1 – Analysing IMS Systems

Part 1 used one HCI evaluator to read about the systems, install the systems, and finally do a heuristics evaluation along five dimensions: comprehensibility, error prevention, task suitability, controllability, and likeability.

These criteria for usability have been further elaborated during the whole study, and the resultant definitions are presented in the next chapter, section 4.2. It should thus be noted that in the detailed report in chapter 10, "Annex 2: Results of the preliminary analysis", somewhat different definitions for the five criteria were employed. "Task suitability" was later redefined (and limited) to "*Main* task suitability", i.e. it concerns only the IMS as a facilitator of other tasks rather than identity management *per se*. Likewise "Controllability" had to be more specific in its name – we have later chosen "Perceived Controllability" – so as not to confuse an evaluator to use it for an objective measurement of people's ability to control the system and thereby overlapping the first three criteria of comprehensibility, error prevention, and (main) task suitability.

In part 1, "likeability" could of course only be the expert evaluator's hypotheses about how actual users, or test users, of the investigated system would react emotionally to each system.

Part 1 is summarised in chapter 5, "Selection of IMS for the Usability Testing", with a table in section 5.4 presenting the main function of each IMS. Chapter 10 gives a detailed account of the findings of this analysis of IMS systems.

3.4.2 Part 2 – Usability Evaluations of Selected Systems

The methods used for the usability evaluations in Part 2 were mainly Think Aloud, Observation, and Interview/Questionnaire. However, for the browser study, the experimenter had to familiarise themselves with the systems and find relevant IMS tasks which meant doing cognitive walkthroughs and looking for possible usability traps.

The studies in Part 2 are reported in chapter 6, "Usability Testing".

4 Usability Metrics

To assess the usability of systems, appropriate metrics are needed as a basis for the evaluation. Since usability is a hypothetical property of a system that cannot be measured directly, it is necessary to break down usability into its components. These components can be further used to define a number of criteria, which can then be applied to evaluate usability of systems.

In the following, five models of usability and their underlying components are presented. The selection considers well known approaches (e.g. Nielsen 1993) as well as usability models that are described in more recently published research papers (e.g. Seffah 2006).

At the end of this section, we select the criteria to apply for the study on usability of IMS and argue why we have chosen not to considered other usability criteria.

4.1 Existing approaches

4.1.1 Usability model by Nielsen

Nielsen (1993), who is one of the leading experts in the field of usability, defines usability as a quality attribute of a software system that estimates how easy a user interface can be used. He describes a tree-like structure where usability is a sub-node of system acceptability (cf. figure 6). In Nielsen's view a software system is usable if it fulfils the following criteria:

- The system is **easy to learn**: How easy is it for users to accomplish basic tasks the first time they encounter the design?
- The system is **efficient to use**: Once users have learned the design, how quickly can they perform tasks?
- The system is **easy to remember**: When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- The system produces **few errors**: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- The system is **subjectively pleasing**: How pleasant is it to use the design?

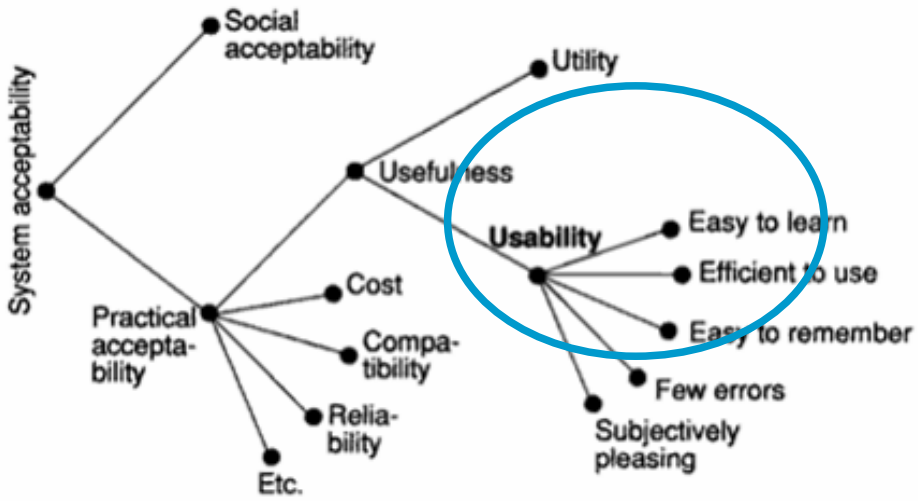


Figure 6: A model of the attributes of system acceptability (Nielsen 1993).

4.1.2 ISO guidelines

ISO standard 9241 establishes general guidelines for ergonomics of human-computer interaction. It contains a set of guidelines for user interface design originally described in part 10. As part of a major revision in 2006, they were moved to part 110.

They are:

1. Suitability for the task: The interface supports the user's task effectively and efficiently. Allowing easy navigation and emphasising the parts of the dialogue which are essential for the task, contribute towards this goal.

2. Self-descriptiveness: Each part of the interface has to be explained and the user needs to be able to obtain information about it. This can be achieved through easily identifiable controls and context-sensitive help.

3. Controllability: The user has to be able to control the direction and speed of the process leading to his goal.

4. Conformity with user expectations: The interface should try to fulfil the user's expectations. This can be obtained by considering the user's background and sticking to generally agreed upon conventions.

5. Error tolerance: It should be easy for the user to recover from errors. Committing an error must not lead to undefined states and incomprehensible error messages.

6. Suitability for individualisation: It should be possible to customise the interface to better suit the user's needs, expertise, preferences and current task.

7. Suitability for learning: Learning to use the program should be assisted and facilitated by design.

This set of heuristics is quite general and applicable to a wide range of systems in which human-computer interaction is a concern. To successfully implement it, it is important to consider the specifics of the project as many of the guidelines are relative to the user and the

task. While they were originally meant to aid system developers with the design of usable systems, these guidelines are also successfully used as a benchmark for usability. Furthermore, being part of an international, European and in many countries national standard, they have been strongly influential for both design and evaluation of usable systems.

4.1.3 Freiburg guidelines for security tools

Usability for security-related systems has been discovered to be a special case. In their influential paper, Whitten und Tygar (1999) observed the failure of traditional usability engineering in the field of security: Although the designers of the evaluated PGP encryption tool observed the rules of classic usability, the users committed severe security compromising errors. The authors identified five special properties of security:

1. Unmotivated user property: As a secondary goal, security is not the main task. The user is unlikely to invest a lot of time and effort.

2. Lack of feedback property: The user needs precise feedback to make well-informed decisions, but security configurations tend to be complex and are difficult to communicate to the user efficiently.

3. Abstraction property: Security concepts often use sets of abstract rules. That makes them hard to understand for laypeople.

4. Barn door property: "The futility of locking the barn door after the horse is gone" (Whitten und Tygar, 1999, p.3): Once committed, a security critical user error cannot be undone. Once unprotected, a secret is gone.

5. Weakest link property: The security of a system is only as strong as its weakest link. So the user must be guided through all relevant parts of his security configuration.

Building upon the ISO 9241-110 guidelines for dialogue design, Gerd tom Markotten (2004) proposed a new set of guidelines that take these special properties of security into account. They are shown in table 2 next to their ISO counterparts.

ISO 9241-110	***Freiburg guidelines***
Error tolerance	Error prevention
Suitability for learning	Ease of use for first-time users
Suitability for the task	Task-suitability for laypeople
Conformity with user's expectations	Conformity with user's expectations
Self-descriptiveness	Self-descriptiveness
Controllability	User guidance
Individualisation	
	Trust

Table 2: ISO-Guidelines and the Freiburg guidelines for security tools

The reasons for these changes are outlined in the following:

• **Error prevention**: This central guideline replaces ISO's error tolerance: Errors have to be prevented as far as possible and the user is warned when an action is potentially security critical. This has to be ensured to take the 'barn-door' property of security into account.

• **Ease of use for first-time users**: As security is only a secondary goal, it cannot be assumed that users are willing to spend time learning the interface. Therefore learnability is replaced with the demand to make the first use of the dialogue as intuitive and understandable as possible.

• **Task-suitability for laypeople**: In a similar vein, task-suitability should be explicitly focused on laypeople: No expertise in security issues can be expected from the user.

• **User guidance**: Controllability and Suitability for individualisation are two problematic guidelines when talking about security tools. If the user can freely move through the interface and deactivate warning messages at will, it is likely that a weakness in the security configuration is not brought to his attention. As the weakest-link-property states, this lowers the security of the whole system and is for this reason unacceptable for secure usability. Instead, the user should be guided step-by-step through important interactions.

• **Conformity with user's expectations** and **Self-descriptiveness**: These are both kept, as they are fully applicable to secure usability.

• **Trust:** This newly added guideline is necessary, because of the mostly invisible results of security tools. The user cannot possibly see if, for example, the antivirus software stays quiet because there are no viruses or because it is just not detecting them. To achieve a subjective feeling of security, it is for this reason important to achieve trust. This can be achieved through constant and precise feedback about the actual level of security.

These guidelines can be used in design but also as a set of heuristics to measure and evaluate usability of security tools. This flexibility as well as closeness of the guideliness to a well established standard should make them easy to embed into existing usability engineering procedures.

4.1.4 Enhanced usability model

The enhanced usability model developed by Abran et al. is based on an analysis of a number of usability models, e.g. Nielsen (1993), Dix (1993), and the ISO 9241 Guidelines. Additionally it is emphasised that further approaches towards usability for instance from other standardisation organisations refer to security as a usability factor as well (cf. Table 3).

ITSEC Information Technology Security Evaluation Criteria.
Standard IEC 300 presents software as security-critical.
ISO 13407 (1999) describes human-centred design as a multidisciplinary activity incorporating human factors and ergonomic and technical knowledge with the objective of raising efficiency and effectiveness, improving human working conditions, and opposing possible unfavourable effects of use on human health, security and performance.
ISO/IEC 9126 defines security, which is a sub-characteristic, as a set of software attributes which relates to its ability to prevent unauthorised access, whether accidental or deliberate, to programs and data.
Security is a characteristic of CHI, which is particularly important in an industrial context (FAA. 1998).

Table 3: Security as usability factor (Abran 2006)

Abran et al. specify five factors of usability in their enhanced model:

- **Effectiveness** as a percentage of tasks achieved per unit of time,
- **Efficiency** measurable by time spent on errors and number of failed commands,
- **Satisfaction** quantified by the number of times that users express their frustration, favourable and unfavourable comments,
- **Learnability** described by the time that people need to learn to use the system.
- **Security** includes access controllability, data corruption prevention and data encryption as capability of the software system to protect data so that unauthorised persons or systems cannot read or modify them and authorised persons or systems are not denied access.

4.1.5 QIUM model

The Quality Integrated Usability Measurements (QIUM) Model by Seffah (2006) describes an integrating model of usability. It provides a consistent framework for understanding and measuring software usability by specifying usability factors, criteria and metrics.

As a starting point the QIUM Model specifies ten abstract factors for usability. Those factors were selected based on the analysis of earlier usability models and standards, for instance the approaches presented above. The ten factors according to Seffah (2006) are:

- **Efficiency***:* the capability of the software product to enable users to expend appropriate amounts of resources in relation to the effectiveness achieved.

- **Effectiveness***:* the capability of the software product to enable users to achieve specified goals with accuracy and completeness.
- **Productivity***:* the level of effectiveness achieved in relation to the resource (i.e. time to complete tasks, user efforts, materials or financial cost of usage) consumed by the users and the system.
- **Satisfaction**: the subjective response of the user while using a software product (i.e. is the user satisfied).
- **Learnability**: the ease with which features required for achieving particular goals can be mastered.
- **Safety**: whether a software product limits the risk of harm to people or other resources such as hardware or stored information.
- **Trustfulness**: whether users trust a particular system in its use.
- **Accessibility**: the capability of a software product to be used by persons with some type of disability (e.g., visual, hearing, psychomotor).
- **Universality**: a software product accommodates a diversity of users (e.g., local culture is considered).
- **Usefulness**: a software product actually helps users to solve real problems.

The QIUM Model breaks down the ten abstract factors into 26 sub-factors, which are more concrete and measurable e.g. via metrics (cf. Seffah 2006). Those sub-factors are called criteria. In contrast to the models that are presented above, QIUM includes privacy as a particular criterion. By means of a factor-criteria-matrix, the relationship between the ten factors and the 26 usability criteria is indicated. After collecting data and evaluating specific criteria, the matrix is necessary to re-abstract from the particular results to a more overall usability assessment of the analysed system.

Criteria	Factors									
	Efficiency	Effectiveness	Satisfaction	Productivity	Learnability	Safety	Trustfulness	Accessibility	Universality	Usefulness
Time Behavior	+			+						
Resource Utilization	+			+						+
Attractiveness			+						+	
Likeability			+							
Flexibility		+	+					+	+	+
Minimal Action	+		+		+			+		
Minimal Memory Load	+		+		+			+	+	+
Operability	+		+				+	+		+
User Guidance			+		+			+	+	
Consistency		+			+	+		+	+	
Self-Descriptiveness					+		+	+	+	
Feedback	+	+							+	+
Accuracy		+				+				+
Completeness		+				+				
Fault- Tolerance						+	+			+
Resource Safety						+				
Readability								+	+	
Controllability							+	+	+	+
Navigability	+	+					+	+	+	
Simplicity					+			+	+	
Privacy							+		+	+
Security						+	+			+
Insurance						+	+			
Familiarity					+		+			
Loading Time	+			+					+	+
Appropriateness							+	+	+	+

Table 4: Relationship between factors and criteria (Seffah 2006)

4.2 Usability Criteria for the FIDIS study

The scope of this deliverable is to investigate the usability of several identity management systems by conducting a summative evaluation. For that evaluation, criteria are needed that reflect a privacy and security aware view on usability. It is neither manageable nor intended to apply a comprehensive list of usability criteria and according metrics, e.g. such as the list from QIUM, to all of the selected systems. However, during our evaluation, we focused on a few particular criteria that stood out as particularily relevant. These criteria were assessed *a posteriori* for each system to be able to compare them. These criteria were further refined during our study on usability of identity management systems and are presented in their most developed forms in the following paragraphs.

4.2.1 Comprehensibility

Identity management and other systems as well use special metaphors that encapsulate complex concepts and are intended to make them more understandable to the user. For

instance the idea of "partial identities" is used in several identity management systems in order to describe the concept of having several sets of personal attributes. *Comprehensibility* investigates whether users are able in general to understand such underlying concepts of the human-computer-interactions as these are presented by the system. If users are not able to understand the concepts, they will in consequence not be able to use the system in the right manner.

4.2.2 Error prevention

Identity management systems are responsible for supporting users in managing and disclosing personally identifiable information. If errors happen and such private data is unintentionally disclosed, this mistake cannot be made undone. Thus, for the assessment of the usability of identity management systems it is important to consider error prevention as a criterion instead of just error correction or tolerance. *Error prevention* captures whether the user interface of an identity management system provides context-sensitive help, meaningful warnings and feedback before privacy-critical errors occur.

4.2.3 Main task suitability

Identity management systems are important for preserving user's privacy. However they are mostly used to fulfil that as a secondary purpose while the user's main intention is another one, surfing on the Internet or buying something at a web shop for instance. Therefore identity management systems' interfaces should especially enable users in carrying out tasks effectively and efficiently. This means providing easy navigation and indisputable user dialogues for instance. *Main task suitability* describes the capability of the identity management system to enable users to fulfil tasks with accuracy and with appropriate amounts of resources.

4.2.4 Perceived controllability

Since user-controlled identity management systems are in the scope of our this study, a focus on controllability as a criterion for usability of the selected systems is needed. From a usability point of view it needs to be investigated not only whether the user really is in control of the system, its features and functionalities, but also if he/she had the impression that he/she controls the system. Thus, *perceived controllability* indicates whether a user feels, that he/she is in control of the identity management system and its functionalities.

Note: Tthe last sentence stresses the user experience rather than an objective measurement of people's ability to control the system – thus it does not overlap with the previous three categories. (This condition was not used in the analysis in section 5.4 and annex chapter 10.)

4.2.5 Likeability

Usability is always subjective and a user-dependent attribute. This is true for identity management systems as well as for any other application. *Likeability* captures users' overall impression of the identity management systems and investigates whether users like the systems in general and, e.g., would spend some money on it or recommend the IMS to friends.

4.3 Criteria not used in the study

4.3.1 Learnability

We have not used the notion of "learnability". A privacy-enhancing IMS must be rather intuitive to use, because IMSs and IMS features deal with possibly sensitive data and if the user makes mistakes once, data disclosure cannot be made undone.

Admittedly, one may think that users do not use a new IMS for the most sensitive information at first, thereby any necessary learning may take place before full understanding of the system is required (cf. our comparison of two web browsers in chapter 6 where test participants had to execute the same tasks on both browsers; learning was quick). However, it is not evident that all users are always able to estimate what are less critical personal information (currently and in the future), and therefore we do not include the "Learnability" among our criteria as comprehensibility (and thus correct use) is important from the start.

5 Selection of IMS for the Usability Testing

Based on the FIDIS IMS database a wide range of IMSs were analysed to find out particular functions that various developers have found essential to be included into their software. Not all systems include all functions. Therefore, to study the usability of IMSs a selection of systems for deeper analyses was made by inspecting systems both for functionality and usability.

5.1 IMSs Analysed

The following twelve systems have been analysed with respect to the functionality they provide as well as the usability of these functions.

- Privacy Bird
- Tor
- JAP
- iJournal
- PiMI
- Opera
- Safari
- Firefox
- Internet Explorer
- CardSpace
- PRIME History function
- PRIME IPV3

The PRIME project has produced several prototypes so this process relied on an evaluation of an html mock-up of the history function, while the final delivered prototype, IPV3, was assessed as concerns other functions.

5.2 Criteria and Methodology

The analysis of the different IMSs was performed in the following manner. We appointed an assistant familiar with usability testing but less familiar with IMS systems to use the different systems to be sure that functions claimed for a system were also reachable for a non-expert.

In short, the workflow for each IMS has been:

1. Finding facts about the IMS
2. Installing the IMS on a local computer
3. Understanding how the product works (instructions, help files, trial and error)
4. Classifying available functions
5. Usability probing based on simple inspection of the IMS systems
6. Compare to usability criteria of section 4.2

5.2.1 Finding facts about the IMS

The facts (about individual IMSs) which this study is based on are mostly taken from the different IMS websites or from help files. Naturally, these sources may give a too positive view of their own software.

5.2.2 Installing the IMS on a local computer

During this study some of the IMS were problematic to install. This is both a problem for the usability of the product and a problem for our evaluation as we have a time restriction.

5.2.3 Understanding how the IMS works

A lot of time was spent using the product, especially learning by trial and error how it worked, how it was supposed to be used, etc. This was seen as a necessary step in order to be able to classify the functions in a proper manner in the next step (5.2.4).

5.2.4 Classifying available functions

In the study information about available *functions* of each IMS is collected. In the PRIME-project (Privacy and Identity Management for Europe) some functions or "characteristics" of what is called "user-controlled identity management system" are mentioned. These characteristics have been used as a basis when categorising available functions within the evaluated IMSs. They are described in the following subsections except for two, namely "user interface" and "user control", because these are not functions but *about* functions (and belong in the present context to the section about usability criteria).

The texts below (5.2.4.1-5.2.4.6) are taken from the PRIME website (PRIME General Public Tutorial v1, 2008).

5.2.4.1 Role and (partial) identity management

"The primary functions of a user-centred identity management system are creating, managing, manipulating and deleting of digital partial identities and the corresponding roles of a user. The manipulation of a partial identity includes adding, deleting or changing some of its attributes, such as the postal address, the e-mail address, or the hobbies of the user. These attributes can be filled in by the user or – in some cases – automatically configured by the system. For these partial identities and the corresponding roles unique pseudonyms are used as identifiers. These pseudonyms make it easier to handle and to reuse partial identities.

It is important that the personal data disclosed by the user cannot be linked with attributes this user uses with other partial identities. In case of such linkability, the identity management system should inform the user and point out which data are linkable, to enable the user to avoid linkage between his partial identities."

5.2.4.2 Authentication management

"Additionally, identity management systems support authentication and access control. In contrast to identification, authentication is the verification of a given identity of a user accessing a computer (or indeed of an application or a computer accessing a second application or computer). The verification is operated by showing a certificate and attributes such as a key, a password or by having some privileges. As soon as a user has authenticated himself to a system, he gets access to the service. These mechanisms are intended to prevent unjustified use of a service. A special case of authentication is the single sign-on, which

enables a user to authenticate himself just once and thereby gain access to the resources of multiple service providers".

"The standard functions of an identity management system concerning authentication and access control are:

- Support for access management (or enabling single-sign-on for each session between the client and the service provider) in order to get access to services of the service provider.
- Support for digital signatures. Digital signatures are digital data, which can be attached to an electronic file or message, and confirm its authenticity and integrity as sent by the user's computer, therefore guaranteeing the correctness of data in messages, documents, e-mails or contracts.
- Support for password management in order to enable the user to manage his login data including his login name, the corresponding password, and the URL of the service provider.

Of course we must remember that the user has to keep his password secret for authentication to work effectively."

5.2.4.3 Form filling

"If the user is browsing the World Wide Web and using web services, he often has to fill in data on a digital form. In this context, identity management systems offer the function "form filling". The form filling supports the user by automatically filling in entry fields with pre-configured data, or else by making suggestions for data to fill in. We can distinguish two categories of form filling. In the first case, the data is filled in only for a particular form of a particular web page. In the second case, the type of each entry field of a form is identified by categories: frequently-requested data such as name, address and credit card number can be filled in automatically, or can be suggested to the user as possible data to use. In principle, data which have been filled in former sessions are stored on the computer of the user and the user can manage these data. Some service providers offer the possibility to store personal data such as the delivery address of the user. If the user accepts this offer and hands over the storage of these data to the service provider, then he no longer exercises complete control of his personal data."

5.2.4.4 Policy management

"A further additional functionality of an identity management system is that of policy management. In this context a policy is a guideline concerning disclosure, use and manipulation of data, especially user-specific personal data. These policies are necessary, both for the user and for service providers like bookland.com. A special case of a policy is a privacy policy, whereby the service provider determines which data he will store for which purposes and to what extent he will make use of these data after the transaction has finished. The possibility of looking at these policies is realised by the function of policy management within an identity management system.

Besides this, the policy management function gives the user control of the use of his personal data, because in each transaction case he can agree or disagree with the use of the data as the service provider, e.g., bookland.com, suggests it. Therefore, on one side the user determines

which personal data he transfers to which service provider for which purpose. On the other side, the service provider also determines which data he needs from a customer and how he will use them e.g., how long he will store these data. Only if the policies of the customer and of the service provider match each other are the data transferred. At present policy management can take advantage of standard terms defined by the platform P3P (Platform for Privacy Preferences); only in this case can there be any automatic recognition of whether the user's and the service provider's requirements match. In Europe, we can observe the trend that more and more Internet service providers are using P3P" […].

5.2.4.5 Context-monitoring

"To help the user to handle his digital partial identities according to his current situation and activity, the identity management system should be able to identify the context of the user automatically. The context-monitoring component of an identity management system identifies contexts within the application or the service the user interacts with. In the case of a context switch, the identity management can act in two ways: either it makes some suggestions for further activities to ensure the privacy and security of the user, or else it takes action autonomously to this end.

To enable this, the system must offer the possibility of specifying contexts and of determine how the user can navigate through them.

One example could be when the user is shopping online: the system could, for example, suggest that after having bought an item the user switch identities so that he can browse anonymously, and then if he decides to buy further items the system could raise the possibility of returning to the previous identity to get the user a discount on his purchases."

5.2.4.6 History management

The phrase 'history management' refers to the management of records of past events:

"The user typically transfers some personal data during the interaction with a service provider. The history management of an identity management system gives him the possibility to go back and review all the data he has transferred to a particular service provider. The history management component therefore logs all transactions with a service provider, and keeps a record of all personal data transferred. The history management is able to present the data transferred, and the context of its transferral, to the user. However the history management does not log the content data the user discloses, for example during his interaction with a chat partner.

The history management helps the user to keep track of which personal data he has given to which communication partner in which context."

5.2.5 Usability inspection of the IMS

In the process of analysing different IMS systems, possible usability problems were also noted. Not every potential problem has been noted in this report as this was just a preparatory step for selecting a few systems for more extensive usability testing.

5.3 Usability criteria

The focus in the evaluation is on the *usability* of the IMS. To be able to compare the systems and estimate their usability we need some criteria. The usability criteria developed in course of this evaluation are found in section 4.2: “Usability Criteria for the FIDIS study”.

5.4 Results of the Inspection

Details of the analysis are found in the appendix (chapter 10), while the following table gives the available functions within the IMSs inspected.

	Role and (partial) identity management	Authentication management	Form filling	Policy management	Context monitoring	History management
Privacy Bird	-	-	-	Yes, its main functionality	System suggests activities when user is in wrong context (e.g. mismatch of policies).	
Tor	Only anonymisation	-	-	-	-	-
JAP	Only anonymisation of http- and DNS-traffic	-	-	-	-	-
iJournal	Only history management	-	-	-	-	iJournal tracks data transferred by the user via a local http-proxy and provides analysis and reporting functionality
PiMI	-	-	-	-	-	-
Opera	Allows *multiple* login identities, but no identifier to each login (e.g. pseudonym). The user cannot handle personal data, only login. Uses certificates to verify websites.	Password manager (e.g. remembers usernames and passwords)	Form filling function called "Wand" (data configurable within Opera)	-	Identifies both context within Opera (help files) and on the service side (user been here before)	-
Safari	Identities can be configured. (but does not assist use of pseudonyms and "Linkability")	Manages login and passwords (also certificates and credentials)	"Form AutoFill" completes online forms from address book (configurable) or previous	-	Identifies both context within Safari (help files) and on the service side (user been here	-

	Role and (partial) identity management	Authentication management	Form filling	Policy management	Context monitoring	History management
			forms.		before)	
Firefox	- (No configurable data within Firefox)	Facilitates password management (also main password on local computer).	On text field level this works automatically. There are add-ons for more advanced form filling functions.	-	Identifies both context within Firefox (help files) and on the service side (user been here before)	-
Internet Explorer	- (No configurable data within IE. No multiple login for single website.)	Facilitates password management (login and certificates)	Can remember login data from earlier fillings, but *cannot configure* these data within IE. Form filling works on text field level.	Policy "summary" can be *asked* for, but no matching, only read policy.	Yes, both within help files and on service side.	
Card Space	Allows multiple digital partial identities by creation of multiple cards; User chooses from her set of cards (partial identities) which data to reveal to the respective receiver.	No separate logon within the application. Possibility to set passwords for particular cards.	N/A because the CardSpace concept presumes servers sends questions to the user's CardsSpace client.	User decides how many and which cards to transmit to partner website; Besides necessary data which is requested by the service provider, optional information can be negotiated for.	If the user is asked for personal information on a website, CardSpace automatically suggests the most-used card for that website if it has been visited before.	History function with information about receiver website, time, date and the type of data sent (but not the data itself)
PRIME (two systems[7])	Allows *multiple* login identities with an identifier to each login (e.g. protection level). Also uses certificates, linkability etc.	Facilitates authentication and password management (login and certificates)	Yes, but in this prototype the link to *configure* form filling data was not implemented.		Yes, PRIME IPV3 is able to identify context.	History function with extensive search function

Table 5: Functions available in the analysed IMSs.

[7] IPV3 April 2008 and history function mockup 2006

6 Usability Testing

6.1 Introduction

The analysis of IMSs reported in chapter 5 indicated that web browsers have some IMS functionality and as this kind of software is used by most Internet users, a usability test was designed to gauge the usability of identity management functions of web browsers. Furthermore, as the more advanced 'type 3' IMS functionality has barely been implemented in everyday software, we include an account of a comparison of one new system and one research project system in prototype form – the comparison concerned the history functions of these advanced IMS systems. Such functions will grow in importance the longer a user has used the system and therefore system developers need elaborate experimental tests rather than field reports they cannot have in order to understand the user requirements for such functions.

Section 6.2 reports a comparison between Internet Explorer and Mozilla Firefox while section 6.3 compares the history management functions in CardSpace and a prototype of the PRIME history function called "DataTrack".

6.2 Comparing two Web Browsers

Two web browsers were tested in the beginning of May 2008 – Internet Explorer and Mozilla Firefox – with respect to their identity management functions. Altogether 10 people participated in the test (called P1-P10). The aim of the test was to see how ordinary users interpret these IMS functions and to compare the systems usability.

6.2.1 Approach

The usability test of Internet Explorer and Firefox were made on a portable computer in the participants' home environment. Each test session was performed in the same way - split in 5 different parts.

1. Introduction to test
2. Agreement
3. Questionnaire
4. Usability test
5. Interview

6.2.2 Introduction to test

Prior to the test each participant were introduced to the usability test by listening to a text read by the test leader. The introduction text informed about the following:

> "Karlstad University is a partner in different international projects working on ways to **increase people's anonymity** when using services on Internet. Internet users leave digital traces when **shopping**, **chatting,** or **searching** information, etc. If these traces are put together the user is profiled to an extent that he/she perhaps not at all like.
>
> Some digital traces are left with the user's **full awareness**, e.g. address data or credit card number when shopping on Internet. But sometimes the user is **not aware** of leaving traces, and not aware that these traces can be connected to them (e.g. search words, cookies, IP-addresses etc)

> The goal of this study is to look at some existing ways to protect users' privacy on the Internet to see if it is possible to increase users' privacy even more. Your participation in this study is therefore important and helpful!"

In this way, all participants got the same information prior to the test. Also an agreement was signed by each participant prior to the test.

6.2.3 Pre-test questionnaire

The pre-test questionnaire contained 9 questions that the participants answered by themselves prior to the usability test. The aim of the questionnaire was to get information about the participants' previous knowledge and Internet habits. Another purpose was to be able to group participants according to their age, sex etc. The questions and multiple choice alternatives are summarised below (in the actual questionnaire each alternative was preceded by a check box):

1. Your sex: Male; Female
2. Your age: 18-25; 26-35; 36-45; 46-55; 56-65; 66-
3. Your highest level of education: Compulsory (9yr), High school (3 yr), University (at least 3yr), (other) _____
4. How often do you use Internet? Every day, Few times a week, Few times a month, Few times a year, (other) _____
5. Which web browser have you used (previous experience of)? Internet Explorer, FireFox, Opera, Safari, (other) _____
6. How concerned are you about privacy protection when using e.g. Internet?
   - O Not concerned, "I feel safe"
   - O Somewhat concerned. I think about it, but do not protect myself
   - O Really concerned
   - O I have not thought about privacy issues prior to today
   - O _____
7. How often do you read Privacy policies on websites at Internet?
   - O Always
   - O Occasionally (now and then)
   - O Never/Seldom
   - O I don't know what a "Privacy policy" is
   - O _____
8. Have you ever (earlier/prior this test) made any settings to protect information about yourself?
   - O Yes
   - O I have tried but not succeeded
   - O No
   - O O I don't know

9. Try to give examples of factors that influence your trust towards websites you visit.

6.2.4 Usability test

The purpose of the usability test was to see how well "ordinary" users can interact with the IMS, to see if there were any difficulties in the usage etc.

The usability test included different parts:

1. Scenario
2. Fictive user identity
3. Tasks
4. Ongoing interview

Scenario

To prepare the participants for the tasks a scenario was presented:

> "You are going to test (use) two different tools for managing digital identities on the Internet - Internet Explorer and Mozilla Firefox. These tools help you to e.g. fill in web forms, remember your login data (passwords) etc.
>
> Imagine that you often use the Internet e.g. to chat, blog, email and shopping. You have several "accounts" on the Internet e.g.
>
> - Two Email-accounts
> - Work
> - Private
> - Bookshop accounts
> - www.adlibris.se
> - www.bokus.se
> - www.amazon.com
> - Blog-accounts
> - "Training"
> - "My garden"
>
> Thus you have at least **seven different accounts** (different digital identities) all demanding login data."

Fictive user identity

To facilitate the usability test the participants were given a fictive user identity to use during the test. The made-up user identity included name, address, phone number and login data (user name and password) to three different web-accounts (Bokus, Hotmail and Funbeat).

Tasks

After the scenario and fictive user identity had been presented the participants were given seven tasks to perform on the computer, each task written down on a paper. The users were verbally informed about what to do and were also allowed to ask the test leader if there was something not understood. The tasks were:

Task 1

Your first task is to make settings in your browser (Firefox/Explorer) to make it **remember your login data** (to facilitate login). How do you make these settings?

Task 2

Visit your account at www.bokus.se. Login and check that your phone number is correct. Log out when finished.

Task 3

Visit your training diary (www.funbeat.se) and read the latest news. Log out.

Task 4

Log in at hotmail (www.hotmail.com) and look if you have any new emails. Log out.

Task 5

Go back to your account at Bokus (www.bokus.se) and check that your email address is correct. Log out.

Task 6

Go back to your account at Funbeat (www.funbeat.se) again. Log out.

Task 7

You have visited different websites and thus left several digital traces. You want to end your Internet session and do not want any other user (using the same computer after you) to use your login data or see which websites you have visited. **How do you delete** your digital traces? [This to be understood as local deletion.]

These seven tasks were performed for one browser at the time, thus all seven tasks were repeated when testing the second IMS. Half group of participants started with Internet Explorer then continued with Firefox, while the other group of participants started with Firefox and continued with Internet Explorer.

Task-related questions during the test

When the participants performed the tasks stated above some questions were asked in connection with some specific task. These questions were:

Question 1 (after *task 1* was performed)

*How was it to make the browser **remember** your login data? (easy/hard/other)*

Question 2 (after *task 5* was performed)

*How was it to login when you **returned** to Bokus? (easy/hard/other)*

Question 3 (after *task 6* was performed)

*How was it to login when you **returned** to Funbeat? (easy/hard/other)*

Question 4 (after *task 7* was performed)

*How was it to **delete** your login data from the browser? (easy/hard/other)*

Question 5 (after *task 7* was performed)

Did you notice anything during the test indicating that the web browser was "secure". Explain!

6.2.5 Post-test interview

After both browsers had been tested the participants were interviewed. The aim of the interviews was to get more information (details) about the users' experiences of the IMS functions and the concepts connected to this kind of functionality.

Altogether 17 questions were asked out loud and the answers were written down by the test leader (and later translated into English to communicate the data within the research group).

1. What is your spontaneous comment after using these two web browsers?
2. Was there something that was hard to do?
3. Was there something that was easy to do?
4. Were there any words or phrases that you did not understand (recognized)?
5. Was there something that you missed – that would have made you more safe (secure)?
6. Which of the two browsers made you feel safer (secure)? Why?
7. Which of the two browsers inspire most trust? Why?
8. Would you buy a product that can protect your privacy when using Internet?
9. How much are you willing to pay for such a product? (e.g. compare price to a similar product)
10. How easy was it to reuse your login data in Internet Explorer?

(Feedback to task 6 & 7 in usability test)

11. How easy was it to reuse your login data in Firefox?

(Feedback to task 6 & 7 in usability test)

12. What do you think about the available functions? Did you miss any functions? Was some function superfluous?
13. Can you mention something that would have made it easier to note when you are anonymous/not anonymous when using Internet?
14. What determines if you trust a website or not?
15. Can you give any examples of things/words that you think inspire trust?
16. Was there anything during the test make you notice if you were protected or not, e.g. when logging in at websites?
17. Which of the two browsers did you like best?

6.2.6 Summary of test result

This section presents a summary by the test leader of the result from the usability test of Internet Explorer and Firefox. The following section contains a table compiled by PRIME

HCI researchers indicating evaluation according to the usability criteria developed for this FIDIS deliverable (cf. section 4.2).

1. **Problems occurred** when participants **made settings** in **Internet Explorer.** E.g. 7 out of 10 participants could not make "remember"-settings in Internet Explorer (even though six of these participants had Internet Explorer as their default browser). Noticeable is also that 9 out of 10 participants where able to make "remember"-settings in Firefox correctly, and five out of these 9 had never used Firefox before. This could be an indication that Firefox is easier to use than Internet Explorer.
2. **Firefox** was generally experienced as the **easiest browser** to use, but participants also thought there was a **risk** to use it because Firefox revealed all login data **automatically** (pre-filled) and therefore data can be misused by others. Some participants preferred simplicity (ease of use) before privacy.
3. **AutoComplete** (in Internet Explorer) was hard to understand for most users. Some did not understand that this was the correct place to make settings, while others were annoyed by the choice of word (name of the function).
4. Easier to make setting the second time. It is quiet obvious that participants easier found where **to make settings the second time**. Only participants who had made these settings earlier knew where to make settings fast. Some of the participants who never had made these settings before had really big trouble in finding where to make settings. It seems they also forgot (or did not understand) that the deletion of the settings is made in the same UI. One can believe that most participants would learn how and where to make these settings very fast (after two or three uses). Some renaming and movement of functions to other menus would probably facilitate usability even more.
5. Users **not aware** that login data is **automatically displayed** when typing first letter in Internet Explorer. E.g. P3, P4 **looked at the keyboard** when typing in login data, which made them not noticing that data was automatically filled in when typing. But in the second task they did notice it – probably because they looked at the screen to check if they had typed correct login data (very long user name used in that task).
6. Some users had problems understanding *where* to make settings in the browser. This can depend on that they **do not see the browser as a program.** One user (P3) clicked on "Start" in windows XP and looked for where to make these settings. She did not understand to make the settings *within* the browser. When asking her she said she didn't know that the menu had such options. She had never looked into the menus before (in her browser).
7. **Frequency of reading Privacy Policies** – All (10) participants answered that they "**Never/Seldom**" read privacy policies.
8. The indication of **security** was really **hard to see**. Only one participant (P7) saw the indication of security in both browsers. P8 said that she was expecting a padlock but she did not see any. In Firefox she saw both the yellow address field and the padlock, but it should be noticed that this time she was looking for it since she knew the question would come.
9. Among the three participants who were **not interested in buying** a product for protecting privacy were the two **older** participants (P5, P6).

10. Those who answered the question of prize on product for privacy, answered fast!
11. Every participant knew exactly what **amount** to spend on a **tool protecting** them on Internet. Participants answered this question within one second. But the amount varied between the participants (5-100 Euro; two of the seven expressed it as a sum 5-10 Euro to pay by month).
12. Interesting reflection from test was that **none** of the participants seemed to **care** (negatively) **about using a fictive user identity**. E.g. one woman said: "it feels strange that my personal login data is saved on someone else's computer…" even though she never entered her own data.

6.2.7 Browser comparison according to the usability criteria

In the table below the observations, when relevant, are related to the criteria elaborated in the preliminary analysis of different IMS systems. Three grades are used: *Poor*, *OK*, and *Good*, while a dash, '-', means that a criterion does not apply to the indicator.

	Comprehens-ibility		Error prevention		Main task suitability		Perceived controllability		Likeability	
	IE	FF	IE	FF	IE	FF	IE	FF	IE	FF
Indicator										
1 setting	Poor	Good	?	(Poor)	-	-	-	-	-	-
2 intuitive	-	-	-	(Poor)	OK	OK	OK	Good	-	-
3AutoCom	Poor	-	OK	-	-	-	Poor	-	Poor	-
4 learnab	'Learnability' is a candidate for being a usability criterion; it is only weakly related to 'Comprehens.'									
5 auto fill	Poor	Poor	(Good)	(Good)	(Good)	(Good)	-	-	-	-
6 browser	(Poor)	(Poor)	-	-	-	-	-	-	-	-
8 Sec ind.	-	-	Poor	Poor	-	-	-	-	-	-
9-11 Buy	-	-	-	-	-	-	-	-	Good	Good

Table 6: Observations of testers (comparison of Internet Explorer with Firefox)

By selecting only participants not familiar with a particular browser, one might interpret the observation in 4, Learnability, as an indicator for Comprehensability and Error Prevention. However, one might also like to remove the cases where the unfamiliar browser was the second one, as there might be a learning effect (at least for some FF-IE orders). The data set on task completion is too small to make any definitive assertions on learning effects, and observation 2 already is an indicator of the "intuitivity" of the browsers, which has been interpreted in the table above.

Thus, we remove observation 4 and also observation 7 on privacy policy, and of course the reflection in observation 12, from the considerations of the five usability criteria. We then reach the following conclusions:

Comprehensibility: the test participants understood the concepts in Firefox better than in Internet Explorer.

Error prevention was not really tested by the tasks, even if some problems with Firefox are notable. Security indications went unnoticed in both browsers. Additional inspection of the browsers can be done to assess what amount of context help, meaningful warnings, and feedback is given before privacy errors occur.

Main task suitability: even if test participants were unused to rely on auto fill functions, there was not much that really obstructed what could have been their main tasks during ordinary browsing (when privacy tasks are not the main tasks).

Perceived controllability: as privacy tasks went better when participants used Firefox than Internet Explorer, it is not surprising to find they perceived Firefox as an easily manageable browser; however, the automatic completion of login data raised discomfort.

Likeability: most test participants liked the idea of privacy protecting software so much that they said they were willing to pay at least a small amount for such a product.

6.3 Comparing history functions in CardSpace and PRIME

The history functions of IMS systems provide their users with a lot of abilities to follow up on data use and increase their own awareness of how much data about them that might be in circulation. However, as data transaction history grows, the use of such functionality will be similar to using databases, a task which many people do not find easy at all. The history functions in CardSpace and PRIME were compared in a usability test in 2007 in order to see users' preferences and abilities. Twelve Internet users participated as test users (named as TP 1-12 below).

6.3.1 Test tasks and post-test questionnaire

The test participants were given some tasks to perform in the history function of each program: PRIME contained 100 faked records, CardSpace only 4 because there were not many CardSpace-enabled sites for the experimenter to visit in order to generate a history log. Before these test tasks each participant made one (1) data transmission with each system. However, these transactions were not recorded by the systems. After the test tasks the participants were given a form with questions to answer.

Six tasks were the same for both PRIME and CardSpace:

1. Find where in the program you can see history information about where you have sent information. Keep the window open to solve the rest of the tasks.
2. Which receiver was the last one you sent information to?
3. When did you last send something to Xxx?
4. Have you ever sent information to a company called Xxx?
5. Which template (for PRIME) / InfoCard (for CardSpace) did you use when you ordered something at Xxx?
6. Which web sites have received information from you when you used your template "Xxx" (for PRIME) / InfoCard "Xxx" (for CardSpace)?

These tasks were used to make it possible to compare PRIME and CardSpace. The number of tasks completed for each program was recorded as well as whether there were any specific recurring problems, etc.

In addition to the six first tasks, the test participants were given four extra tasks to perform in PRIME. The results of these tasks were later used to examine the extra functionality that DataTrack contains. The tasks were:

7. At which web sites have you registered a user name?
   Purpose: To solve this task, the test person needs to understand how to search for a specific date type. This task examines if the test person understand how to do that.
8. What information has the receiver Xxx got about you?
   Purpose: This task also examines if the user understands the search function.
9. Before a cinema visit the 3rd of February 2006, you visited the cinema web site to book some tickets. You gave the phone number of your mobile phone, but what other information did you send to the cinema web site?
   Purpose: The tickets were not booked the same date as the cinema visit, they were booked some days before. This task examines if the test person understands that he or she shall search for the cinema web site, and not only look at the specific date.

10. How many times has the receiver Xxx got information about your e-mail address xxx@xxx.xx?
Purpose: The receiver has got information about another e-mail address than the one asked for. This task examines how careful the test person performs the task.

When the tasks above had been performed, the test persons filled in a questionnaire about the programs. These questions were used to examine what the test persons were thinking about the programs, what they look for in such programs, and which program they prefer. The questions in the form were:

1. When you sent information when ordering, you used two different programs instead of filling in a form. What did you think about that, compared to fill in forms as usual?
   Purpose: Before the tasks were performed the test persons tried to send data with the programs (these data transmissions did not add to the faked history records used in the experiment). This question examines what the test persons thought about that.

2. Which one of the history functions do you prefer, and why?
3. Which advantages and disadvantages do you see with the CardSpace history function?
4. Which advantages and disadvantages do you see with the PRIME history function?
5. If you want to find out what you have sent to a specific web site, how do you want to do that?
6. Is there something you think was missing in the programs you tested?
7. Was there something you did not understand in the programs you tested? Describe concise.
8. Do you want to use such a program? Why/why not?
9. Other thoughts, proposals, or comments?

6.3.2 Test persons and main results

TP	*Gender*	*Age*	*Occupation*	*Computer experience*	*Internet use* *	*Ordering online?*
1	M	18-25	Web dev.	Used to	Daily	often
2	W	26-35	Designer	Used to	Daily	sometimes
3	W	26-35	Unemployed	Used to	Daily	sometimes
4	W	26-35	Project ass.	Used to	Daily	sometimes
5	W	18-25	Student	Used to	Daily	sometimes
6	W	18-25	Web ass.	Used to	Daily	sometimes
7	W	46-55	Bank empl.	Used to	4-7days/w	It has happened
8	W	18-25	Working	Used to	4-7days/w	sometimes
9	W	18-25	Student	Rather used	Daily	sometimes
10	M	26-35	Unemployed	Used to	Daily	often
11	M	26-35	Company owner	Used to	Daily	sometimes
12	M	26-35	Working	Used to	Daily	sometimes

Table 7: Test persons involved in the testing

* Several options were given; only two used: *Several times/day* and *4-7 days/week*.

6.3.3 Comparison PRIME - CardSpace

TP		*Score PRIME (1-6)*	*Failed tasks PRIME 1-6, [extra tasks 7-10]*		*Score CardSpace*	*Failed tasks CardSpace*		*Prefer*
1		6 (100%)	[9, 10]		6 (100%)	-		CardSpace
2		6 (100%)	[9, 10]		6 (100%)	-		PRIME
3		5 (83%)	5, [10]		5 (83%)	4		CardSpace
4		6 (100%)	[9, 10]		6 (100%)	-		PRIME
5		5 (83%)	6, [9, 10]		6 (100%)	-		PRIME
6		5 (83%)	5, [9]		6 (100%)	-		PRIME
7		5 (83%)	5		5 (83%)	1		PRIME
8		5 (83%)	5, [10]		6 (100%)	-		PRIME
9		5 (83%)	3, [10]		6 (100%)	-		PRIME
10		3 (50%)	3, 5, 6, [9, 10]		6 (100%)	-		CardSpace
11		4 (67%)	3, 6, [10]		6 (100%)	-		PRIME
12		4 (67%)	3, 5, [9, 10]		5 (83%)	4		None

Table 8: Results of the testing – PRIME vs. CardSpace

When only counting the tasks that were the same for PRIME and CardSpace and therefore comparable, the test persons scored on average 4.9 of 6 points in the PRIME tasks and 5.8 of 6 points in the CardSpace tasks. Notably, the possibility to give wrong answers was smaller in CardSpace because DataTrack contained more records and more search opportunities.

When it comes to PRIME, the men performing the test got a worse result than the women, but when it comes to CardSpace the result is the same. The fact that 88% of the women prefer PRIME, when only 1 out of 4 men prefer PRIME might be a reason.

6.3.4 Post-test questionnaire

All test persons except one were positive about using a program like PRIME or CardSpace when they send information online. They thought it was convenient, and that it would be even better when they got more used to the programs. One test person that did not want to use such a program pointed out that he does not want to install extra programs on his computer that "steal memory" from the rest of the programs.

8 test persons preferred the history function of PRIME, 3 test persons preferred CardSpace and one of the test persons did not want to choose since he did not understand any of the programs. Several test persons commented that the search function in PRIME is better, but the user interface of CardSpace is nicer.

The test persons answered a question about how they want to see the history of their transactions. Three of the test persons suggested a solution that allows to see the history already in the web browser. The other test persons suggest improvements in either PRIME or CardSpace, e.g. better explanations about how to find history and to show all history in the same place.

6.3.5 Observations

Several test persons looked in just one InfoCard to see the history. Most of them understood, after a while, that they should search in the other InfoCards when reading more test tasks. Some test persons commented during the test that it is hard to know if something has been sent in CardSpace, and what data have been sent.

When the test participants had clicked through CardSpace for a while, they became familiar with it and found operating it easy.

In PRIME, many test participants got stuck when they searched for where to find template information. They clicked at a lot of transactions and hoped to find it, and some of them commented that it is not possible to search for template information.

It was hard for many of the test persons to understand that it is possible to click on a record in the record list of DataTrack to see more information. This was in part because the mouse cursor in the prototype used in the test did not become a hand when the user pointed the cursor over a record. Some of the test participants found out that it is possible to click on a record when they by mistake clicked on a record.

Apart from these two remarks, it was easy to search the DataTrack. The users can search with "search sentences", e.g. "What information about me has <receiver> received?", or with the free text search. Several test participants used the "free text search" as the last resort when they did not know how else to find information, and most often this strategy worked.

6.3.6 A summary of usability problems found in the test

Among the problems and disadvantages when it comes to CardSpace, the following are the ones that are most important or voiced by most participants in the test:

- Several test persons did not understand that the history of transaction should be searched for in all the InfoCards. They first just looked in one InfoCard and thought it was all of it.
- A collection of the history (for all InfoCards) is missing.
- Search functionality is missing.
- It is hard to understand what has been sent, and check if something has been sent.

For the last statement it should be noted that it is hard for a user to know what information on a card was included in different disclosures as it is possible to suppress or add information when using a card for sending data.

Among the problems and disadvantages when it comes to PRIME, the following are the ones that are most important or voiced by most participants in the test:

- It was hard to find information about what template has been used in a transaction.
- It was hard for many of the test persons to understand that it is possible to click on a record in the record list of DataTrack. (In this prototype, the mouse cursor did not become a hand when the user points the cursor over a record.)
- It was not possible to see all records at the same time. When one first opens DataTrack, all records should be shown from the beginning.
- The language was complicated and was hard for some of the test participants to understand.

6.3.7 The history function by our usability criteria

In this section some reflections are made about the history functions of CardSpace and PRIME according to the usability criteria selected for this study (defined in 4.2). The comments below are based on the result summarised above (more extensive treatment in Rönntorp, 2008).

6.3.7.1 Comprehensibility

CardSpace's user interface was easy to deal with but some things were hard to understand, e.g. history of transaction should be searched for in all the InfoCards. Test participants initially looked in only one InfoCard. It was also hard to understand what had been sent, and if something had been sent.

According to the usability study PRIME's history function DataTrack was easy to use and most users liked it. Thus, the users also seemed to think the history function was quiet comprehensible.

6.3.7.2 Error prevention

No information about this for CardSpace is found in the test report, but see the note in the previous section about the difficulties to understand what had been sent.

For PRIME, some test participants asked for a Help file.

6.3.7.3 Main task suitability

This category is not quite suited for evaluation of a history function of IMS because for a history function the main task is the same as the relevant IMS-related tasks. In any event, the PRIME project also sketched functions for contacting data controllers and data protection boards. For such tasks, the IMS use would only be a means not an end. According to the study is seems as if the test users thought the PRIME history function was suitable for its task.

In CardSpace, the user needs to open each card to see the history of it, and there is no search function inside the cards, so the user needs to go through every transaction if he or she does not remember when the information was sent or who it was sent to. Thus the CardSpace history function is harder to use and therefore do not support tasks in the same way as the PRIME history function does.

6.3.7.4 Perceived controllability

CardSpace saves a lot of information about a data transaction. Thus, the user has in principle control over what data has been sent and who has received it. However, the test and post-test questions showed that people did not think they had good command over the history records.

Also the PRIME history function saves a lot of information about a data transaction. The test users felt more in control over the history function because of the search functionality within the DataTrack.

6.3.7.5 Likeability

According to the post-test questions 3 of the 12 participants preferred CardSpace over PRIME while 8 of the 12 test participants preferred PRIME over CardSpace. Notable is that 11 liked the idea of identity management systems.

6.3.8 How to make a useful IMS history function

To summarise suggestions on how to make a useful history function, this is what is important according to the test participants:

- A nice and simple user interface which is easy to understand and use.
- The history of all transactions saved in one place, and an easy way to see all records.
- Search functionality, especially free text search but also the search sentences that are used in PRIME are useful.
- The grouping of information in InfoCards in CardSpace is popular among the test persons. This grouping is easier to understand than the template concept in PRIME.
- A language that is easy to understand.
- Much information about transactions. This is an advantage in the PRIME test prototype; CardSpace needs to show more information.
- Help functionality.

If these criteria are fulfilled, the history function should be useful for both experienced users and users with average computer experience. Computer and Internet experience did not seem to play a decisive role in how well the twelve test participants understood the programs. Neither did age or occupation – what the test leader could see matter was how careful the test persons were when they perform the tasks.

6.3.9 What the test participants asked for in a history function

To summarise the question about what the test participants asked for in a history function, this is what came up in the post-test questionnaire:

- Browser integration. Three test persons suggested a solution where it is possible to see the history of their transactions already in the web browser.
- One test person suggested an Identity Manager that works as e.g. MSN messenger – it is possible to use it from all computers; you just sign in to your account.
- Several test persons suggested improvements in the current user interfaces and functions, e.g. better help functionality.

7 Summary and Conclusions

In the introductory chapter, it was mentioned that by closely observing the characteristics of innovations and their resulting effects, models for the successful willingness of users to adopt an IMS can be developed. To this regard, usability seems to play an important role for the adoption and the user acceptance in order to make identity management technologies a success. Therefore this report set out to discuss and develop a set of usability criteria for IMS as well as give an overview of usability methods and measurements (chapter 3).

Usability criteria developed in this FIDIS study were (section 4.2):

- Comprehensibility
- 'Error prevention'
- Main task suitability
- Perceived controllability
- Likeability

Each of these five criteria is given a comprehensive discussion in section 4.2, as is also a criterion mentioned in section 4.3, 'Criteria not used in the study', namely 'Learnability'. A system that is easy to learn is of course good, but if an IMS is not intuitive enough to provide immediate instructions on the correct use, there is a risk that the user will perform actions he or she will regret later. 'Comprehensibility' and 'Error preventive' thus seem more important than a gradual understanding of an IMS.

Analysing and testing IMS systems: Chapter 5 gave an account of how a wide range of IMSs were analysed based on the FIDIS IMS database. This analysis of systems revealed what IM functions developers have found useful to include in their software. The scan also revealed that not all systems included all functions – rather, IM functions have been included in a rather piecemeal fashion.

The selection of IMS systems for this study included Privacy Bird, Tor, JAP, iJournal, PiMI, Opera, Safari, Firefox, Internet Explorer, CardSpace, PRIME History function, and PRIME IPV3. The analysis of these different systems was performed by a HCI assistant going through documentation for each system, installing the system, familiarising with system, classifying available functions, usability probing based on so-called 'inspection' of the system, and finally comparing with the usability criteria selected and developed for this FIDIS report (section 4.2).

A table in section 5.4 demonstrates the function types identified and the extent to which each system implements this function (the Appendix to this report gives a more detailed commentary to each system and the results from the inspection evaluation). The functionalities provided by these IMS were:

- Role and (partial) identity management
- Authentication management
- Form filling
- Policy management
- Context monitoring

- History management (i.e., management of records of data disclosures and policy agreements)

Furthermore, this study has reported two smaller usability tests. Section 6.2 compared two web browsers (Internet Explorer and Mozilla Firefox) because the scanning of IMS systems just mentioned showed that browsers – i.e. the software used by almost all Internet users and definitively by all non-savvy users – have some identity management functionality and the usability of such functions in such applications would thus be of general interest to IMS developers as well as privacy experts and the general public.

The test showed that web browsers may support main-task performance even if an automatic completion of login data raised discomfort. Security notifications went unnoticed in both browsers tested. Several test participants liked the idea of privacy-enhancing software and could imagine paying at least a small fee for it. (Section 6.2.7 discusses the browser comparison according to the usability criteria developed for this study.)

The other test (section 6.3) compared history functions in a newly released IMS (Microsoft's CardSpace) with a prototype from a research project (DataTrack from PRIME). It turned out that the existing product was very 'card-centric': not only was the user supposed to release data via a card metaphor, but release history could only be tracked via each individual card. The research prototype offered more general search functions. Test participants performed better in the commercial product but the majority liked the more general history function in the research product. Again, almost all participants liked the idea of identity management systems. (Section 6.3.7 summarises test results according to the usability criteria developed for this study, noting in particular that criterion "Main Task Suitability" can have several interpretations when evaluating history functions.)

Conclusions:

This study has elaborated on a limited set of broad criteria for IMS usability evaluation, defining five, namely: Comprehensibility, Error prevention, Main task suitability, Perceived controllability, and Likeability. Objective 'controllability' is well covered by the first three, while the *perceived* controllability needs a particular focus, as also likeability does. 'Learnability' was excluded with the motivation that the user cannot risk learning by making mistakes; however, it was noted that this exclusion is debatable (section 4.3).

Besides giving guidelines for future usability evaluations of IMS systems, this report has demonstrated the limited scope concerning IMS of the applications existing on the market. It has furthermore found six major IMS functions present in existing products and research prototypes, i.e., role and identity management, authentication management, form filling, policy management, context monitoring, and history management. It has also analysed the last IMS function mentioned as this one has so far has received little attention in different applications, namely a history function for personal identity management. The demand for such a function among Internet users as well as the demands on such a function in future IMS systems were explicated through a usability test comparing Microsoft's CardSpace and the PRIME project's DataTrack.

8 Bibliography

8.1 Scientific References

Abran, A. et al., 'Usability Meanings and Interpretations in ISO Standards', in: *Software Quality Journal*, 4/11, pp.325-338, 2003.

Bauer, M., Meints, M., Hansen, M. (eds.), *FIDIS Deliverable D3.1 Structured Overview on Prototypes and Concepts of Identity Management Systems,* Frankfurt a.M., 2005. Download: http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.1.overview_on_IMS.final.pdf.

Barber, C., Stanton, N., 'Observation as a technique for usability evaluation', in: Jordan, W.; Thomas, B.; Weerdmeester, B.; McClelland, I. (Eds.): *Usability evaluation in industry*, pp. 85-94London: Taylor & Francis, 1996.

Barnum, C. M., *Usability Testing and Research*, Longman Publishers, New York, NY, 2002.

Bortz, J., Döring, N., *Forschungsmethoden und Evaluation*, Springer, Berlin 1995.

Chappell, D., *Introducing Windows CardSpace*, Redmond 2006. Download: http://msdn2.microsoft.com/en-us/library/aa480189.aspx

Cranor, L. F., 'User Interfaces for Privacy Agents – discusses Privacy Bird design and user studies', in *ACM Transactions on Computer-Human Interaction,* 2006. Download: http://lorrie.cranor.org/pubs/privacy-bird-20050714.pdf

Davis, F. D., 'Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology', *MIS Quarterly* Vol. 13, No. 3, pp. 319-339, 1989.

Dingledine, R. Mathewson, N. Syverson, P, 'Tor: The Second-Generation Onion Router'. In *Usenix Security*, 2004. Download: https://www.torproject.org/svn/trunk/doc/design-paper/tor-design.pdf

Dix, A. et al.. *Human-Computer Interaction*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

FAA, *Standard Terminal Automation Replacement System, Human Factors Team Report of the Computer–Human Interface Re-Evaluation*, 23 February 1998.

Fischer-Hübner, S., Nilsson, M., Lindskog, H., 'Self-Determination in the Mobile Internet - PiMI Prototype Results'. In: *Proceedings of IFIP TC11 17th International Conference on Information Security (SEC 2002)*, Cairo/Egypt, 7-9 May 2002, Kluwer Academic Publishers 2002.

Gerd tom Markotten, D., *Benutzbare Sicherheit in informationstechnischen Systemen,* Albert-Ludwigs-Universität Freiburg, Dissertation, 2004.

Görner, C., Ilg, R.. ‚Evaluation der Mensch-Rechner-Schnittstelle.' In: Ziegler, J.; Ilg, R.: *Benutzergerechte Software-Gestaltung*, pp. 189-206, Oldenbourg Verlag, München 1993

ISO 13407:1999. Human-centred design processes for interactive systems.

ISO 9241-110:2006. Ergonomics of human-system interaction – Part 110: Dialogue principles.

ISO/IEC 9126:2001. Software Product Evaluation—Quality Characteristics and Guidelines for the User, Geneva, International Organization for Standardization.

Lewis, C., Polson, P. G., Wharton, C., Rieman, J., 'Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces', in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People CHI '90*. ACM, pp. 235-242, New York, NY, 1990.

Nielsen, J., *Usability Engineering*. Academic Press, Boston, MA. 1993.

Pettersson, J. S., *PRIME HCI Guidelines D06.1.f*, Brussels 2008. Download: https://www.prime-project.eu/prime_products/reports/arch/pub_del_D06.1.f_ec_wp06.1_v1_final.pdf

Rönntorp, N., *Personal identity management systems - A comparison with a special focus on the history function in CardSpace and the PRIME prototype*, BA thesis in Information Systems, Karlstad University 2008.

Rogers, E. M., *The Diffusion of Innovations*, 5th Edition, Free Press, New York, 2003.

Seffah, A. et al., 'Usability measurements and metrics. A consolidated model.' In: *Software Quality Journal*, 2/14, pp.159-178, 2008.

Tergan, S.-O., 'Grundlagen der Evaluation: Ein Überblick', in: Schenkel, P. , Tergan, S.-O., Lottmann A. (eds.): *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme. Evaluationsmethoden auf dem Prüfstand*. pp. 22-51Nürnberg: BW-Verlag, 2000.

Whitten, A., Tygar, J. D., 'Why Johnny can't encrypt: A usability evaluation of PGP 5.0', in: *8th USENIX Security Symposium*, 1999.

8.2 Web Links

Apple website (2008) "Safari". http://www.apple.com/se/safari/ (2008-04-01)

AttrakDiff (2008). http://www.attrakdiff.de/en/Home/ (last access 18 Sept. 2008)

FIDIS website (2008) "IMS db/**MozPET**". http://www.fidis.net/interactive/ims-db/?tx_fidisdb_pi1%5BshowUid%5D=106&cHash=2d796c7f5a

FIDIS website (2008) "IMS db/Opera". http://www.fidis.net/interactive/ims-db/?tx_fidisdb_pi1%5BshowUid%5D=19&cHash=1de5fb21d2 (2008-03-28)

IT transfer office website (2008) "About **iJournal**". http://www.ito.tu-darmstadt.de/projects/prima/ijournal/index_en_html)

JAP Website (2008). http://anon.inf.tu-dresden.de/index.html (2008-03-17)

JonDonym Test site (2008). https://www.jondos.de/en/ (2008-03-17)

Microsoft (2008b) "CardSpace". http://www.microsoft.com/net/cardspace.aspx (2008-04-10)

Microsoft (2008) "Internet Explorer 7.0". http://www.microsoft.com/windows/products/winfamily/ie/default.mspx (2008-04-04)

Mozilla (2008) "Firefox". http://www.mozilla.com/en-US/firefox/ (2008-04-03)

MozPET website, (2008) "**iJournal**". http://mozpets.sourceforge.net/)

Nielsen, Jakob (2008). Useit.com: Jakob Nielsen's Website. Online available http://www.useit.com.

Nielsen, J. (2008) "Ten heuristics". http://www.useit.com/papers/heuristic/heuristic_list.html (2008-04-21)

Opera Website (2008) "Products". http://www.opera.com/products/desktop/ (2008-03-28)

PRIME website (2008) "Start". https://www.prime-project.eu/ (Read 2008-04-20)

PRIME website (2008) "Tutorial/IDM Systems/ Properties of IDMs". http://blues.inf.tu-dresden.de/prime/EUT_Tutorial_V0/PRIME_nm.htm (2008-03-04)

Privacy Bird website (2008). http://www.privacybird.org/ (2008-03-05)

Privoxy website (2008). http://www.privoxy.org/ (2008-03-11)

Tor website (2008). http://www.torproject.org/ (2008-03-07)

Torbutton (2008). http://torbutton.torproject.org/dev/ (2008-03-11)

Vidalia website (2008). http://www.vidalia-project.net/ (2008-03-11)

9 Annex 1: Indices

9.1 Index of Figures

9.2 Index of Tables

10 Annex 2: Results of the preliminary analysis

Below is a presentation of all the IMS analysed in this study. The analysis was conducted by Maria Lindström in Sweden, and the preliminary assessments are in most cases hers.

This appendix is structured in the following way. First a *short description* (section 9.x.1) of the IMS is presented and then a description of *available functions* (section 9.x.2) and a *summary of hypothetical usability problems* (section 9.x.3) detected during inspection of the product. Finally some reflections are made about the IMS according to the *usability criteria* (section 9.x.4).

10.1 Privacy Bird

The Privacy Bird® - a free tool available for download (added to the user's web browser). The Privacy Bird was originally developed at AT&T Corp, and is now maintained by "CMU Usable Privacy and Security Laboratory" at Carnegie Mellon University, USA (Privacy Bird, 2008).

10.1.1 Short description

The Privacy Bird is a tool that automatically *searches for privacy policies* at websites. The Privacy Bird reads the privacy policies and notifies the user whether the site's policies match the personal privacy preferences or not. Note that the Privacy Bird does not hinder data transaction – it only notifies the user about a mismatch. The user is allowed to configure its own personal privacy preferences.

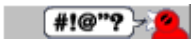
According to the Privacy Bird website (Privacy Bird, 2008) the tool allows the user to ask for *warnings* at Web sites that may:

- Sell your name and address to other companies without your permission
- Use your health or medical information for marketing
- Put you on mailing lists that you can't get off of etc.

The Privacy Bird uses **icons** to display (visualise) the *level of match* between users' preferences and the websites privacy policy. Below are some examples of these icons (Privacy Bird, 2008).

The singing **green bird** appears when Privacy Bird determines that a web site's privacy policy **matches** the users' preferences. When a red exclamation point will appear next to the notes in the bird's song bubble it tells the user that the website contains **embedded content** that do not have any privacy policies associated with them, or that the privacy policies do not match the user's preferences.

The uncertain **yellow bird** appears when Privacy Bird is **unable to fetch** or read a privacy policy from the web site you are visiting.



The angry **red bird** appears when Privacy Bird determines that a web site's privacy policy **conflicts** with your preferences.

10.1.2 Available functions

The Privacy Bird facilitates **Policy management -** Privacy Bird is a tool for *matching* privacy policies. More details in Appendix 1.

10.1.3 Summary of usability problems

When walking through the Privacy Bird® trying to use it potential usability problems were noticed, a summary of these problems below. Note that some of these problems can depend on that the browser used in this test (Internet Explorer 7.0) has new features (e.g. tabbed browsing) which are probably not applicable on Privacy Bird.

Browser hides the icon

The first problem – and a serious one – is that the **browser hides the bird icon**. To be able to see which colour the bird has the user need to click on an extra (empty) Internet Explorer icon in the lower status bar (marked in picture below).

Tabs activates one bird icon per tab

A second problem is that the browser (Internet Explorer) uses tabs when having several websites active at the same time. When *using tabs the Privacy Bird activates one bird icon per tab*. This makes the lower status bar crowded of hidden bird icons. And when surfing on Internet the user is not able to know which bird goes with which website.

Not resizable windows

Another problem is that the pop-up windows were not resizable. This is both annoying and hindering for users.

Icons representing important information

It is obvious that the icon represents a bird, but it is not obvious *what* the bird is trying to say to the user. The bird has different colour representing different types of match/mismatch in policies/preferences. But the bird icon also has a "song bubble" with different meaning. According to the Privacy Bird website the bird icon can have different moods (e.g. sometimes the bird is "angry", "uncertain", "sleeping" etc) representing e.g. different levels of match/mismatch. But is this obvious to all users? Only the eyebrows of the bird changes when representing different moods. Cranor, the principal developer of Privacy Bird, explains that selected symbols

> *"…convey the messages "your preferences are matched" and "your preferences are not matched" rather than "your privacy is protected" and "your privacy is not protected."(Cranor, 2006)*

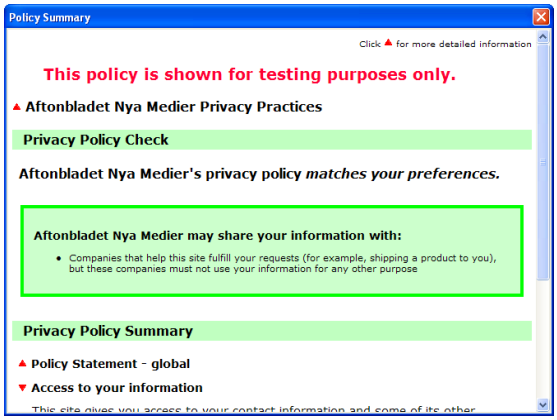
Pop-up windows

A lot of pop-up windows appear when using the Privacy Bird. After a while it is hard to know which window is "old" or which is "new".

Policy summary window

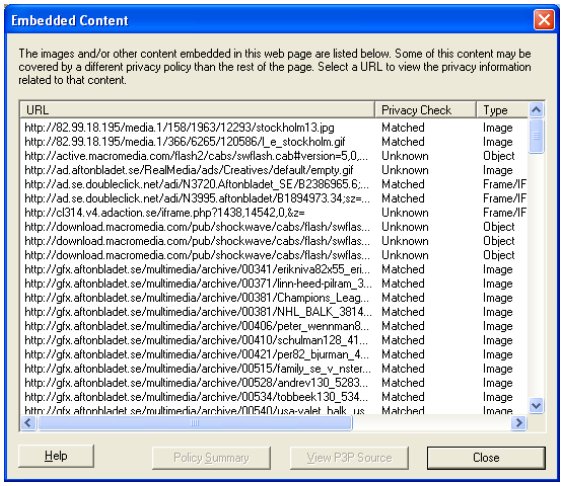
The UI for **"Policy summary"** was very content rich making important information "lost in the crowd" (see picture below). Other problems connected to "Policy summary":

- the text structure may not be optimal
- the UI had many different colours, which might be disturbing
- Used warning signs (red triangles) for expanding hidden text (e.g. instead of links)
- Window is not resizable etc.



Embedded Content

The information in this UI (picture below) can be hard for "ordinary users" to understand. What to do with this information? Another problem is that this UI is **scrollable sideways.** This is not good from a usability perspective. This UI is not resizable.



"View P3P Source" (activated by button in UI above) can be hard to understand for some users; its purpose should be explained (that the natural language translation can be checked).

Bird – yellow all the time

Another problem is that **the bird stayed yellow at most websites** which were confusing – most sites are not P3P-enabled. JSP&SFH want to have red for non-P3P sites.

The option "**No-prompt Sites…**" was hard to understand. The **text could be shorter and more concrete**. What exactly does it mean?

The option "**Clear cash**" did **not give any feedback** of what happened when clicking on it. Another problem with "clear cash" is that it is a technical term perhaps not understood by all users.

Finally - the problem with **long** "summaries". A summary should be short and only the most important information should be presented. One summary was so long **I needed 5 (five) screenshots to capture the text**. In PRIME a multi-layer structure is used for policies.

10.1.4 Usability criteria

In this section some reflections are made about the Privacy Bird according to the usability criteria selected for this study.

10.1.4.1 Comprehensibility

The Privacy Bird offers the user to check if a privacy policy match or mismatch. As an illustration of this check a bird icon is used (a bird in different colour and different mood) to reflect how the matching went. Since the bird icon can appear in many different shapes (different colour and different moods) it can be hard for users to understand what the Bird is telling. Thus the comprehensibility can be violated.

Also, some users can have problems in understanding the concepts of "privacy policies". And since these policies are very extensive it is possible that users not even read them. The Privacy Bird offers a "summary" of the policies that could be improved by explaining to users what this policy mean and what effect it has to the user.

The overall use of the Privacy bird is okay. Because the privacy bird **only matches policies** (not hindering) many other privacy protective features (which are harder to comprehend/use) are not evaluated here.

10.1.4.2 Error prevention

In the Privacy Bird the user gets both a signal (sound) and symbol (colour) reflecting the status of the matching. But the system does not prevent the user from doing mistakes. Since the Privacy Bird **only warns** the user (not hinder) errors may happen. This is serious since it is affecting a person's privacy. The developer (Cranor, 2004) comments this choice:

> "Some P3P critics have argued that for a privacy agent to be effective it must bring about a direct increase in privacy protection for the user through its support of the collection limitation principle. Thus, these critics view software that anonymizes a transaction as an effective PET, but claim that a tool that simply informs users about privacy practices cannot be effective".
>
> "We find this view of PETs to be overly narrow, as increased transparency about privacy practices can enable users to make informed decisions about when to provide their data. As a secondary effect, the transparency provided by the P3P protocol in combination with P3P user agents may motivate web sites to improve the privacy protections they offer, or it may highlight areas where further privacy regulation might be needed"

10.1.4.3 Task suitability

The Privacy Bird uses a small icon in the browser and is therefore available all the time. It also offers the user to make its own shortcuts to some features – which is good. But the use of technical terms, too long and bad structured privacy policy notifications do not encourage the task. Pop-up windows hinder users in their main task. But the icon and sound makes it easy to use (*if* the users know what the icon and sound means) and is not hindering usage.

10.1.4.4 Controllability

In the Privacy Bird the users is allowed to set their own privacy preferences which in a way give the user control. But the user is only in control over the preferences the user has set – not in control over that these preferences *are followed*. The Privacy bird only checks privacy policies at service providers. But the user does not know if the service provider really follows these policies. The user gets a link to the website to be able to check the policy. But the user has no control in deleting its data from a service provider. So no – the Privacy bird do not give the user full control.

Privacy Bird allows setting preferences according to data types, but not for individual web sites. In principle, if Privacy Bird is combined with data release functions, one would like to set preferences also as concerns specific data and not only data types.

10.1.4.5 Likeability

Presumably, some parts of the Privacy Bird are likable by most people but a lot of privacy protecting features was missing, e.g. it has no history function. On the other hand, the high degree of "ease of use" can be connected to the low number of features.

10.2 Tor

The Tor Project is a non-profit project based in the United States. The organization consists of many volunteers and a few employees.

10.2.1 Short description

This is a short presentation of the software "Tor". The information in the description is from the Tor website (Tor, 2008).

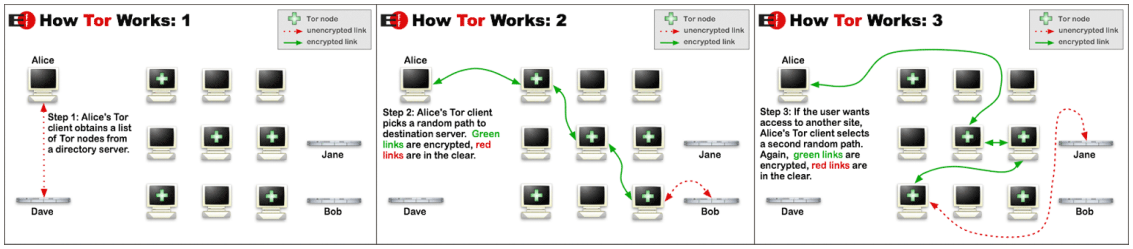
Tor – The second-generation Onion Router "is a software project that helps you defend against traffic analysis, a form of network surveillance that threatens personal freedom and privacy, confidential business activities and relationships, and state security". It focuses only on **protecting the transport of data**.

According to the Tor website hundreds of thousands of people around the world use Tor. Some examples of users are "journalists and bloggers human rights workers, law enforcement officers, soldiers, corporations, citizens of repressive regimes, and just ordinary citizens".

Examples of what "ordinary internet users" can gain if using Tor is:

- Protect their privacy from unscrupulous marketers and identity thieves
- Protect their communications from irresponsible corporations
- Protect their children online (IP-address)
- Research sensitive topics

How does it work? According to their website "Tor protects you by bouncing your communications around a distributed network of relays run by volunteers all around the world", illustrated in the figures below:



"Tor is a network of virtual tunnels that allows people and groups to improve their privacy and security on the Internet […]. Tor helps to reduce the risks of both simple and sophisticated traffic analysis by distributing your transactions over several places on the Internet, so **no single point can link** you to your destination".

10.2.2 Available functions

Tor **only hides IP-addresses**, making it possible to use internet partly anonymous. The user cannot hide their data (the information sent is not hidden). Thus all information can be "listen to" if not encrypted by the user.

Tor **does not** facilitate **any** of the requirements (see 2.1) for an IMS, neither basic functionality (identity management, authentication management) nor additional functionality.

10.2.3 Summary of usability problems

After using the Tor software some reflections are summarized about its usability.

Puts burden on users

When collecting facts at Tor website (2008) some statements got special attention since they potentially will worry the users. At the website all responsibility for usage is also put on the users. E.g. in the lower hierarchy on the website the user is told that "Tor does not protect you if you do not use it correctly". The user is also urged to read their "list of warnings and make sure to follow the instructions for your platform carefully". This is not user-friendly, but the warning is indeed needed: cf. the "Embassy Password Leak" The website also informs the user that "even if you configure and use Tor correctly, there are still potential attacks that could compromise Tor's ability to protect you".

Incorrect information given to users

Some statements at the website can be challenged, e.g. the statement "protect children online" which is incorrect information since the underlying concept is about hiding IP-address.

Too technical terms:

"Using Privoxy is **necessary** because browsers leak your DNS request when they use a SOCKS proxy directly, which is **bad for your anonymity**. Privoxy also removes certain dangerous headers from your web requests, and blocks obnoxious ad sites like Doubleclick."

Unprofessional language:

"This depends entirely on how you installed it. If you installed a package, then hopefully your package has a way to uninstall itself. If you installed by source, I'm afraid there is no easy uninstall method. But on the bright side, by default it only installs into /usr/local/ and it should be pretty easy to notice things there".

Problems installing the product

When installing the product a lot of potential usability problems were found. First of all the installation guide is probably **not written for ordinary users** – e.g. had too technical words and deficient information if the product was correctly installed.

Conflicting messages

After installing Tor some time were spent on understanding what the UIs were communicating. E.g. some messages (text and icons) were **conflicting** (figure below).

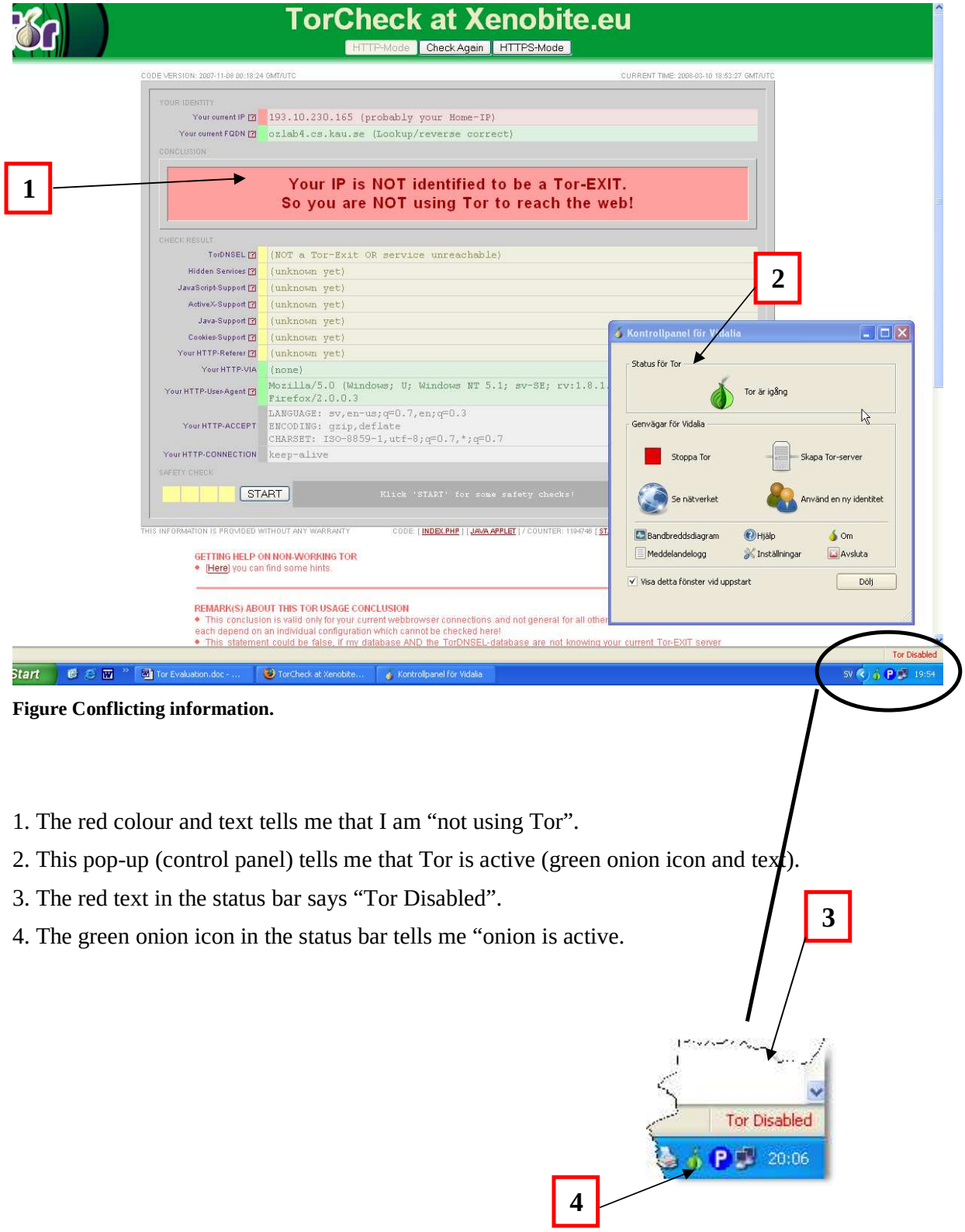


Figure Conflicting information.

1. The red colour and text tells me that I am “not using Tor”.
2. This pop-up (control panel) tells me that Tor is active (green onion icon and text).
3. The red text in the status bar says “Tor Disabled”.
4. The green onion icon in the status bar tells me “onion is active.

The use of short cut

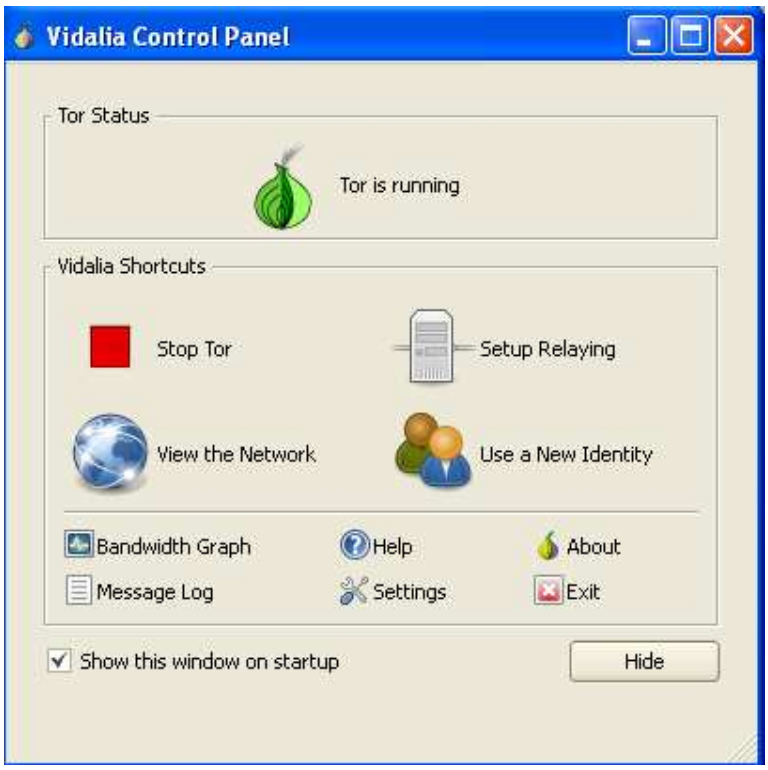
In the UI there is a shortcut called Torbutton (point 3 in picture above), for enable/disable the Tor. But the "button" is not a button – it is a text in the status bar.

The use of colours

Another drawback with Tor is the **use of many different colours** (e.g. not seen by colour-blind and not in a black and white printout of this document…). In Tor colours are used on icons, in UI objects, on background and in text. Colour-coding is good – sometimes – e.g. to strengthen something important. But if too many colours are used these important things disappear.

No visual difference of clickable or non-clickable objects

There is no visual difference of clickable or non-clickable objects in the user interface (UI). E.g. in the picture below the green onion in the upper part of UI is *not clickable*. But the red square (for stopping Tor) is clickable. Nothing in the UI shows any difference between these two objects. It should have - either by underlining the text (indicating it is a link) or by making the object look like a button (indicating it is clickable).



Other reflections:

- Tor seems to be the **number one choice for criminals** (according to statements at the Tor website and an extern "blog" about Tor)
- Tor was **slowing down the web browser**

10.2.4 Usability criteria

According to online document (Dingledine, R. Mathewson, N. Syverson, 2004) Tor defines usability not only as a convenience, but also as a **security requirement.** "A hard-to-use system has fewer users — and because anonymity systems hide users among users, a system with fewer users provides less anonymity".

In this section some reflections are made about the Tor according to our predefined usability criteria.

10.2.4.1 Comprehensibility

Tor does not make underlying concepts comprehensible. The product is hard to understand. And the only way to see if the IP-address is hidden the user need to visit a website that checks this for them. Even though it is good that extern websites check this, the Tor product by itself is not giving the user any trust, at all.

10.2.4.2 Error prevention

Tor is not giving the users any feedback – the user does not know what happens and it is hard to understand what the product is helping with. Tor is not experienced as error preventive. Many messages to the user is either conflicting (e.g. "Tor is running" and "Tor Disabled" at the same time) or nonsense messages (e.g. in the "Message Log").

10.2.4.3 Task suitability

If (or when) the user has understood how Tor works and what the purpose of using it is, then it can be useful. But Tor does not have the "basic" features of an IMS (2.3.1-2.3.6).

10.2.4.4 Controllability

No, as far as this inspection shown, the user does not feel in control of this product. It is hard to be sure of what happens and hard to make sure it works at all (if the users doesn't visit the "check IP-address" website).

10.2.4.5 Likeability

The Tor product is hard to like since it does not inspire trust – most likely, it is not making the user feel safe. Moreover, because it is only hiding the user's IP-address Tor can be experienced as only wanted by criminals. But the idea is good – if all other parts of an IMS was included (e.g. different levels of anonymity, policy checking, preference setting, history function etc.)

10.3 JAP

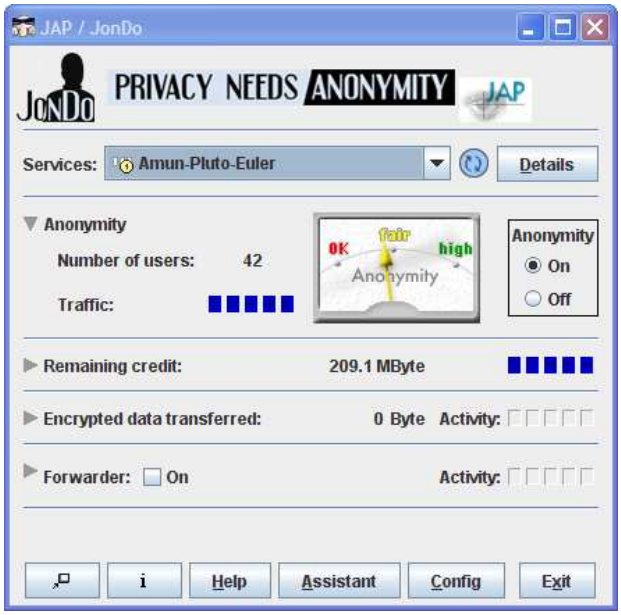
The JAP Project is run by University of Regensburg, Department of Business Informatics, and Technical University Dresden, Department of Informatik, according to JAP website (JAP Website, 2008). Also ICPP in Kiel, is involved in developing JAP.

10.3.1 Short description

According to JAP website (JAP, 2008) JAP makes it possible to surf the internet anonymously and unobservable. Without a*nonymization*, every computer on Internet communicates using a traceable address. That means:

- the website visited,
- the internet service provider (ISP),
- and any eavesdropper on the internet connection

can determine which websites the user of a specific computer visits. Even the information which the user calls up can be intercepted and seen if encryption is not used. JAP uses a single *static address which is shared by many JAP* users. That way neither the visited website, nor an eavesdropper can determine which user visited which website.



"Instead of connecting directly to a web server, users take a detour, connecting with encryption through several intermediaries, so-called *Mixes.* JAP uses a predetermined sequence for the mixes. Such a sequence of linked mixes is called a *Mix Cascade*. Users can choose between different mix cascades.

Since many users use these intermediaries at the same time, the internet connection of any one single user is hidden among the connections of all the other users. No one, not anyone from outside, not any of the other users, not even the provider of the intermediary service can determine which connection belongs to which user. A relationship between a connection and

its user could only be determined if *all* intermediaries worked together to sabotage the anonymization" (JAP Website, 2008).

10.3.2 Available functions

The JAP system only facilitates "anonymization" by "connecting with encryption through several intermediaries". Thus this is not an IMS offering identity management, policy management etc.

10.3.3 Summary of usability problems

After using the JAP system a summary of reflections of its usability and functions will be presented.

First of all, the JAP **website looked professional**. This is needed to give the user (at least a potential user) a good first impression and inform the users about the product, e.g. the Privacy Bird and Tor systems had websites looking less professional. One problem connected with the website was that it was **hard to understand what the product was helping** the user with - and what it was not helping with. This problem belongs to the difficult solution rather than to the website itself.

After reading the JAP system was installed on a local computer. The **installation guide was helpful** and quiet informative. But after awhile the inspection evaluator **was unsure if the product had been configured right.** There are a lot of settings the user have do set and since no "green" confirmation was displayed insecurity took place. Another problem was that the "turn-JAP-on-action" was hindered by a usability problem. **The user has to spend too much effort** and the **presentation of "how to do" it is deficient**.

No help to overcome problems

When checking level of anonymity detailed information is given to the user about possible problems (activated by using mouse over). But they **do not tell how to overcome the** problem.

Slow web browsing

JAP really slows the web browsing down. It seems like a faster connection (version) is available but costs money.

Some labels are missing in the UI which would have helped to understand it better. E.g. in the lower part of UI (close to the two server icons) a label would make it easier to understand.

Other reflections about JAP

- Hard to understand that the JAP UI was extendable
- Too much settings to make (burden on users)
- "Certificate" with information hard to interpret

Conflicting words

One of the most used buttons in JAP is called "Details" but the activated window is called "Settings". Another example is the first label in JAP UI – called "Services" – is not presenting services, but **servers** (available to the user).

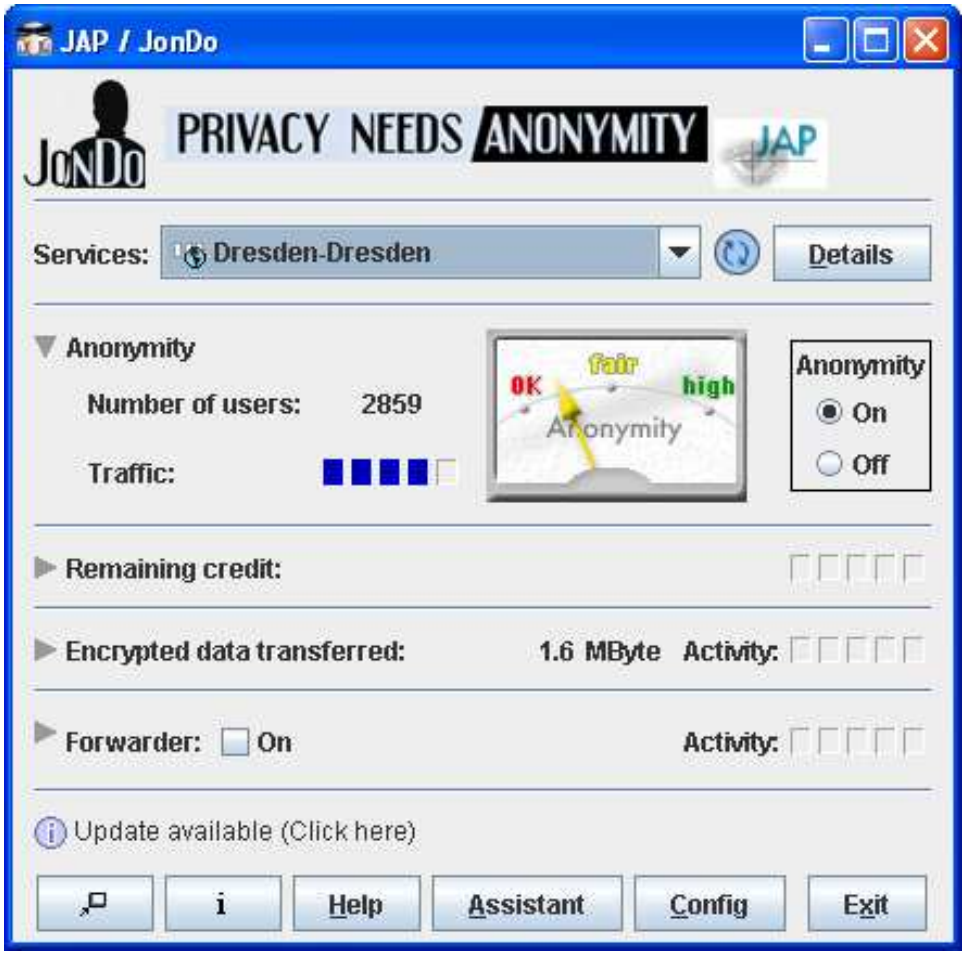
Confusing text colour

When using the "Assistant" the user should select options by using radio buttons. The user is has different options to choose from, but the strange thing is that **one option** is coloured (font

colour) in green while the other options are in black font. At first is seems like the user should choose the green option (green colour indicated "OK" or "good alternative") but it was not correct. Why is the text green? And If the purpose was to give the user a hint of "the proper" option - then this option should have been pre-selected (with a radio button) instead of coloured in green.

Colour-blind are hindered

Colour-blind people are possibly hindered when using JAP because of the frequent use of colours representing different level of anonymity.



Picture "JAP UI"

My spontaneous (at first glance) comments about the JAP UI:

- Some **icons** are not understandable (e.g. in the lower UI)
- **OK** is coloured in red (**red is not** indicating OK)
- "**Traffic**" is not understood
- **"Number of users"** is good to see
- **"Remaining credit"** is not understood

- "**Encrypted data transferred**" is not intuitively understood
- "**Forwarder**" is not understood

User control

The user of JAP has options to select and deselect options of different parts of JAP-system. This gives the user freedom to make settings of their choice, and makes the user in control of the system (at least the parts which are able to configure). An example of settings the user can make: (the text is taken from "Help files" in JAP)

> "You may block single services by unchecking the checkbox to the left of the service. JonDo does not connect to blocked services, and they are not listed in the choice menus. This is useful if you do not trust single services, or their service quality (speed/stability) seems to low".

The user is also able to make personal settings to the user interface etc. But **no** settings are available to protect personal data, no privacy policy management etc.

10.3.4 Usability criteria

In this section some reflections are made about JAP according to its usability.

10.3.4.1 Comprehensibility

When using JAP it was hard to understand what it was helping the user with, and what it was not helping with. A lot of settings were available for the user to configure, but JAP failed to make the user understand the underlying concept of these settings.

10.3.4.2 Error prevention

JAP gives a lot of feedback information about eventual problems, but no information is given about how to overcome the problem or prevent it. The help-files included in JAP are informative and offers contextual help, but the text is too extensive and does not by itself prevent errors.

10.3.4.3 Task suitability

The main functionality in JAP is to offer "anonymization" (not identity management, policy management etc). Thus JAP can be used for this purpose, but some usability problems (see above) will probably hinder the user in carrying out its primary task.

10.3.4.4 Controllability

The user probably will not feel fully in control of JAP. The user is allowed to make (a lot of) settings, to control how JAP behaves (maybe too much settings). It was really hard to make settings to be fully "anonym" (all fields coloured in green), which should have been the easiest setting to do if the system would have been "in the hand of the user" (controllable). To be able to know if JAP was correctly configured the user also needed to "test" this at a "test site".

10.3.4.5 Likeability

JAP is probably more likable than e.g. Tor (offering almost the same feature). JAP inspires more trust but was too hard to configure.

10.4 iJournal

No available download of a compiled version of the IMS.

iJournal now is a small part of a larger system – called MozPET.

10.4.1 Short description

The goal of the MozPETs project is to produce a web browser and mail client with enhanced privacy and security features.

The iJournal allows the user to keep track of personal information he has given to web sites. If the user enters sensitive information (e.g. his name), the **iJournal** intercepts the transmission and tries to profile the web site by analyzing its P3P policy, SSL certificate and WHOIS entry MozPET website, 2008)

According to MozPET website (2008) a plan is to "add as many privacy enhancement technologies (PETs) as possible".

10.4.2 Available functions

The MozPET facilitates the following functions:

- "We modified some settings to have a more privacy-friendly default value. For example, MozPETs does **not accept third-party cookies** at all and treats all other cookies as session cookies
- MozJAP is a wrapper to integrate a proxy for **anonymous web access** with Mozilla. It supports JAP and Tor.
- MozPAw checks the current page for potential harmful content like web bugs. The result is displayed as an emoticon to alarm the user if a website makes intensive use of tracking technologies.
- Allows the user to **block third party content** (e.g. webbugs) embedded within web pages to reduce cross-side clickstream tracking by advertisers
- Tracknosis is a tool that reveals cross-side tracking activities used by advertisers
- **History function - The iJournal allows the user to keep track of personal information he has given to web sites. If the user enters sensitive information (e.g. his name), the iJournal intercepts the transmission and tries to profile the web site by analyzing its P3P policy, SSL certificate and WHOIS entry. The data is stored for later analysis.**
- Add PGP Support to Mozilla Mail. EnigMail is developed by the EnigMail project. Please check their pages for more info on EnigMail (and use the original EnigMail if you want to use it in production
- Send **e-mail anonymously** through the remailer"
  (MozPET website, 2008)

iJournal has these functions (IT Transfer Office, 2008):

- "Record who has your private information with the iJournal (screenshot).
- Gather as much information about those people who track you as possible (screenshot). Then you can legally require them to remove your information.
- Easily configured as a standard Mozilla component. (screenshot).
- Pull information through whois, ssl, and p3p.

- Permute your data slightly so you can tell who was using it.“

10.4.3 Summary of usability problems

MozPET are still in an early developer version. Compiled sources are not available yet. Source code needs to be compiled and linked, using certain tools, certain operation system’s platforms and requiring fundamental knowledge in setting up make-files.

The steps required are not documented in detail so that the evaluators were not able to set up a running version of MozPETs for Windows systems.

10.5 PiMI – Not evaluated

No available download of this IMS. According to one of the developers, Prof. Simone Fischer- Hübner (Karlstad University), the PiMI software is not maintained any longer.

10.5.1 Short description

PiMI system is for **mobile phones**.

Fischer-Hübner, Nilsson and Lindskog (2002) present results of the PiMI prototype project with participants from Ericsson AB and Karlstad University. "In the PiMI project a browser built-in and a proxy-based P3P (P3P) user agent for controlling the dissemination of CPI in mobile Internet environments by the means of Minimal Profile Conveyance was developed. The approach of Minimal Profile Conveyance requires that the user defines a minimal CPI profile, containing only information that the user considers completely harmless, or where there is an understanding that this information may be necessary for some reason. In the extreme case, the profile could be empty. This minimal profile can be used:

1. for accessing non-P3P enabled web sites or web sites that do not meet the user's P3P privacy preferences
2. for serving third party requests to the WAP Gateway for cached profiles (such as for WAP push content generation)
3. for communication in the "safe-zone" (as defined in P3P) before a P3P agreement

The end user also has to define a full CPI profile to be used when there is a successful P3P agreement, i.e. the site is P3P compatible and the site's P3P policy file matches the end user's privacy preferences.

Even though P3P can enhance transparency and control over data disclosure for users, it has also been criticised as it does not ensure compliance of privacy policies with privacy laws, it does not guarantee a minimum and non-negotiable level of privacy protection for individuals and in its current form it does not fulfil the legal requirements for obtaining technically user consent.

Privacy-enhanced mobile identity management systems can go a step further and should provide a means for privacy control (consent, objection, disclosure, correction, deletion and addition) and for privacy-compliant data processing of CPI and other personal data belonging to a user's mobile identity".

10.5.2 Available functions

See the previous paragraph.

10.6 Opera

The web browser Opera (Version 9.26) was installed and inspected the 28th of March 2008.

10.6.1 Short description

According to the Opera website (Opera, 2008) they mean that "Opera Web browser is faster and more secure than other browsers available on the market"

FIDIS website (FIDIS, 2008) defines Opera as: "Opera contains simple form-filler, so personal data apart from usual interaction-data can be stored "

10.6.2 Available functions

Opera offers many different features - according to their website - e.g. "Site Preferences" if you want to block pop-ups, "Cookie handling" and "Tabbed browsing" etc.

Interesting for this evaluation is the following features:

- **Fraud Protection**
  Opera's advanced fraud protection protects you against web sites that try to steal your personal information.
- **Password Manager**
  "The password manager remembers your usernames and passwords so you will not have to."

These features are within IMS functionality "**authentication management"** and "**context monitoring".**

10.6.2.1 Fraud Protection

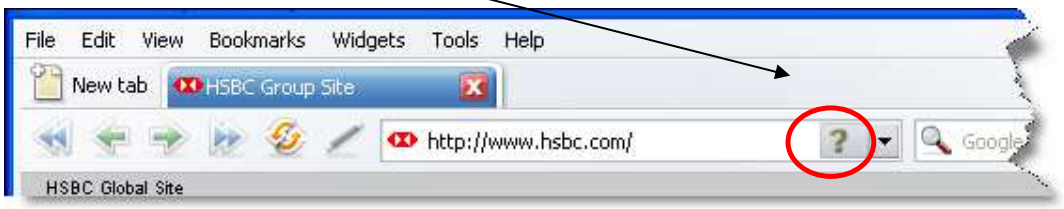
According to "Help" files within Opera "Fraud Protection **warns you about suspicious** Web pages by checking the page you visit against a database of known "phishing" Web sites". The Fraud Protection gives the user a "security status" of every visited website, and the result of the security status is presented as an icon on the right side of the address field (in picture below).

There are different kinds of status, as can be seen in picture below.

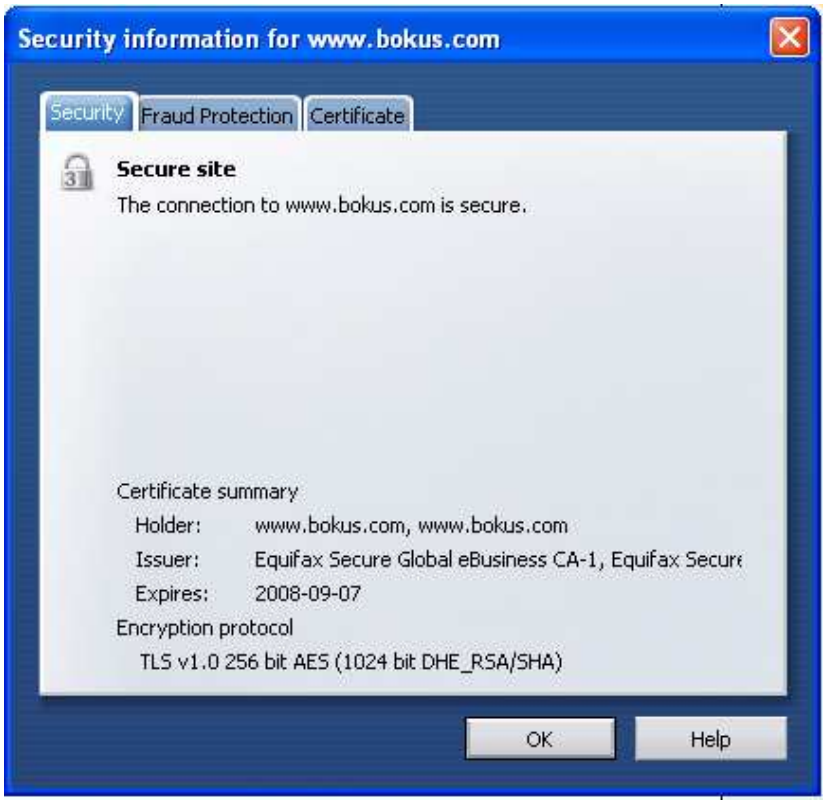
Address Field	Fraud Protection Dialogue
	VERIFIED
	NOT VERIFIED
Fraud Site	WARNING
	VERIFIED

An exclamation mark next to the address field means that the fraud check has **succeeded** without any problem – the website is verified.

A question mark next to the address field (http) in Opera means "the Opera **hasn't checked** the site" or the website is not verified e.g. due to the lack of information.



If clicking on the icons (security status) the security information for the page to be displayed, including information about the Web site's "Security", "Fraud protection" and "Certificate" (tabs in picture below). Below is a screen shot showing the "security" check of a website.



Picture: "Security"

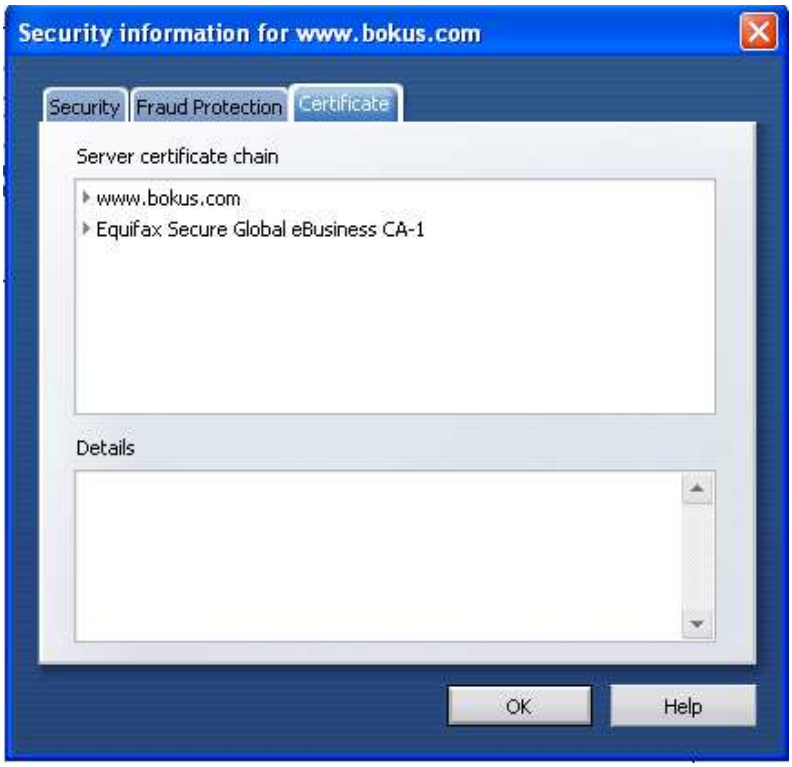
The "Fraud Protection" (picture below) states that this website is "verified by a trusted third party and is **not listed** on Operas blacklist of suspected fraud sites". The user is able to report this site "as fraudulent" if they want to.

Picture: "Fraud Protection"

If a Web site **is found** on the blacklist, the user will be presented with a warning page, and must decide whether to visit the fraudulent Web site, or not.

Picture: "Fraud Warning"

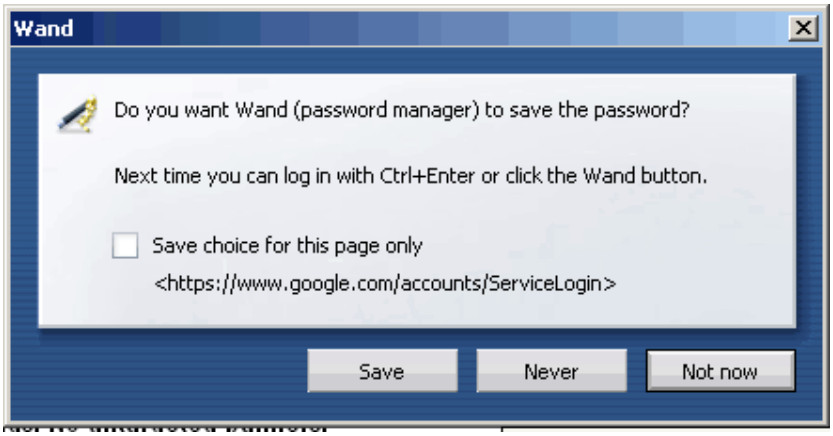
"Certificates" are used to verify that a Web site is secure to use. If there is something questionable about a certificate, a warning dialog will be displayed. Certificates like credit cards have an **expiry date**, and must be renewed on a regular basis by the people maintaining the site. A certificate is issued by an authority for a single site to use, and sites cannot borrow certificates from each other as this invalidates the whole concept of certificates.



Picture: "Certificate"

10.6.2.2 Password Manager

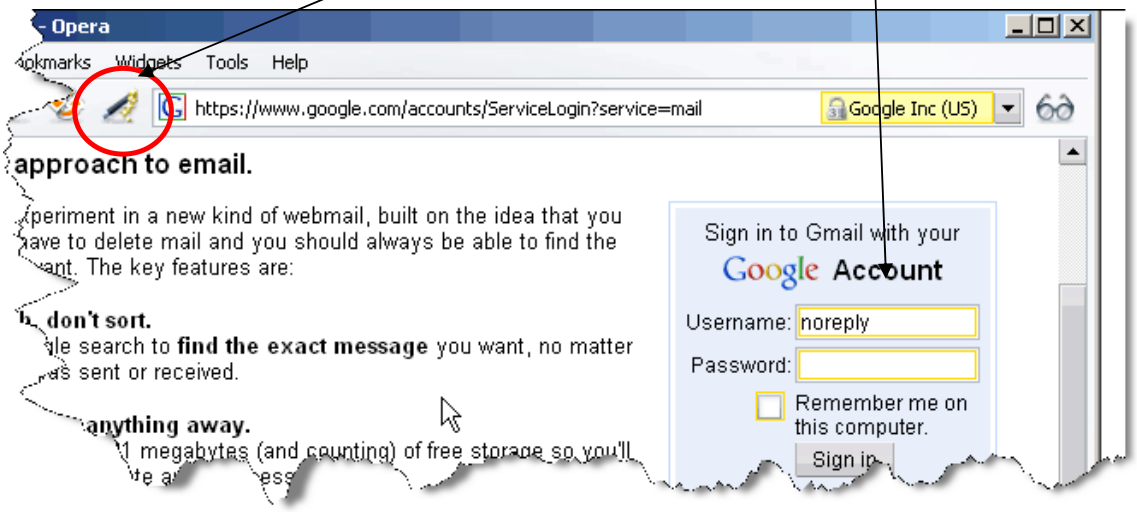
Opera's password manager is called "Wand". The Wand is automatically activated when entering a website asking for your password, or can be manually activated by clicking a "wand icon" in the address field.



Picture: "Wand – Save password…"

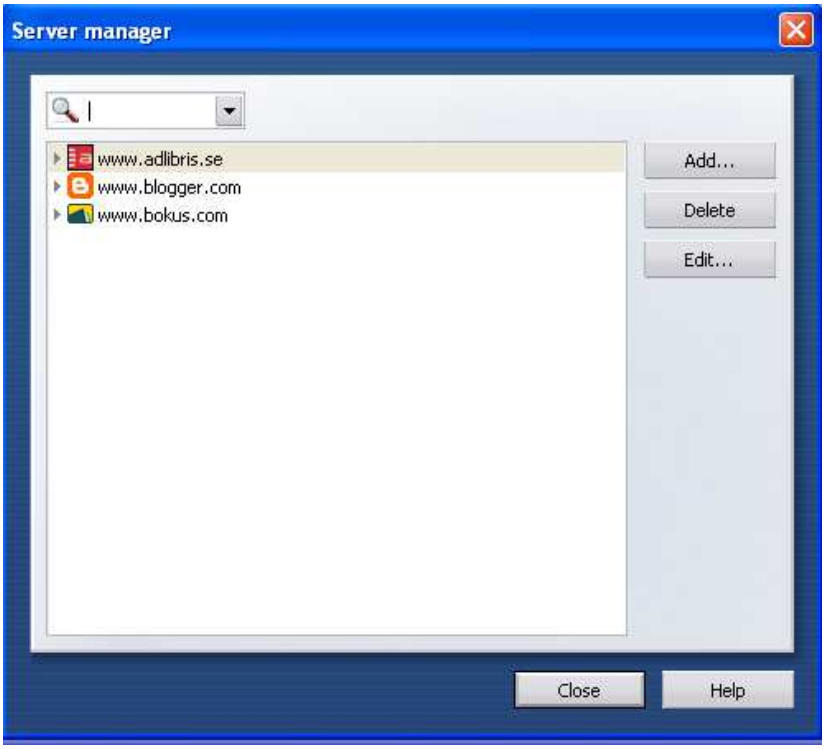
The user can decide if Opera should remember the password. The UI above have a **checkbox for saving password** "for this page only", which allows the user to save multiple passwords per domain. If the user leaves this checkbox unchecked it means that the password is stored for the entire website.

The next time the user enters this website Opera will highlight the form fields (yellow as in picture below). The user only has to click on the "wand" icon, and the form will be automatically filled in (or press ctrl+enter).



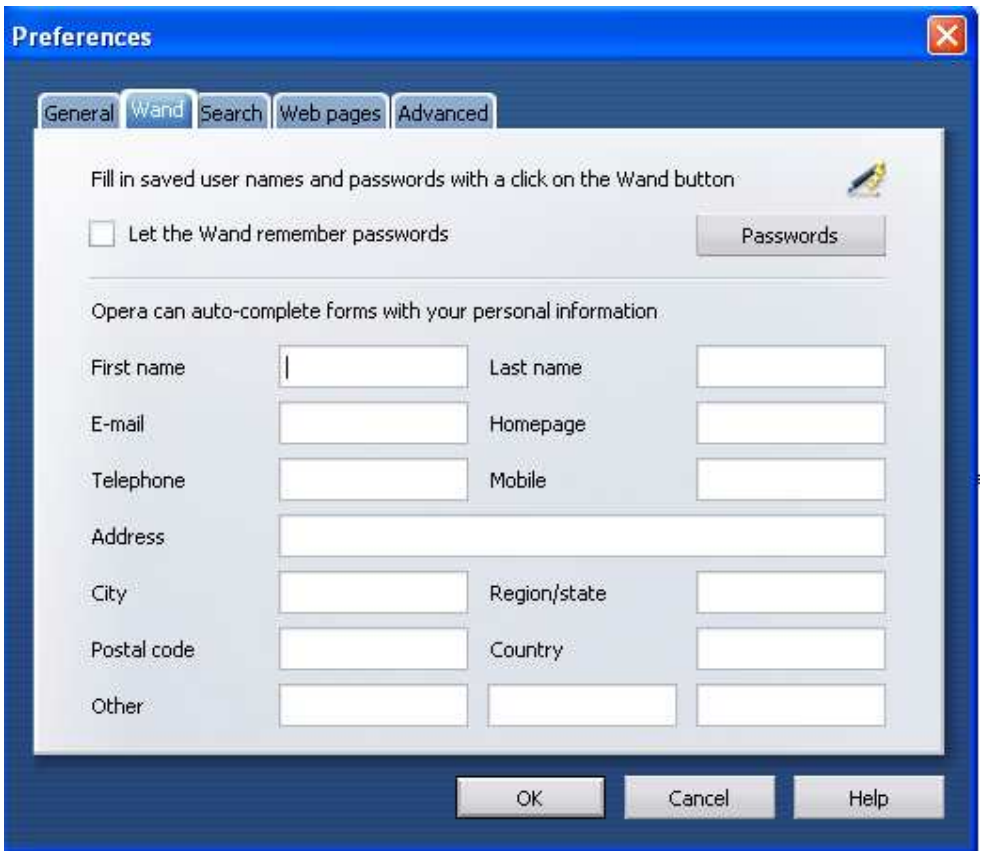
Picture: "Highlighted form"

The user can configure (e.g. add, edit or delete) settings made in the form filler by using "Tools >Preferences > Wand > Passwords".



Picture: Configure passwords

The user of Opera is also facilitated in filling in other forms on websites. The user configure their personal data in “Wand” (Tolls>Preferences>Wand) as in picture below.



Picture: Wand

In the picture above of the Wand window, the user can pre-fill personal data, facilitating fast fill-in forms at websites later.

10.6.3 Summary of usability problems

When walking through Opera the following potential usability problems were found:

Warning not seen

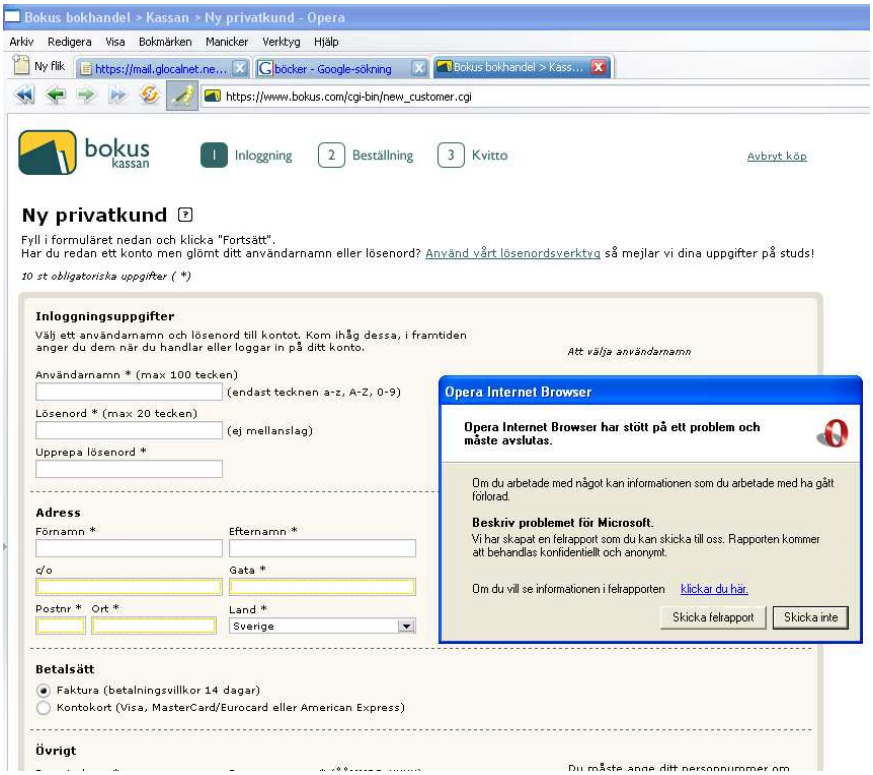
The "warning" icon is too small and not in "eye focus". Below is a picture of an example were the website is not fraud checked. But the problem is that this warning is hard to see (notice) e.g. since the icon is too small and is displayed in the "wrong" place. The users focus is often in the upper left corner or in the centre of the UI. In Opera the warning is placed to the right in the address field.

Warning icon - not warning

The choice of warning icon can be argued. The icon above should warn users that the website is not fraud-checked. Is a question mark really indicating a warning?

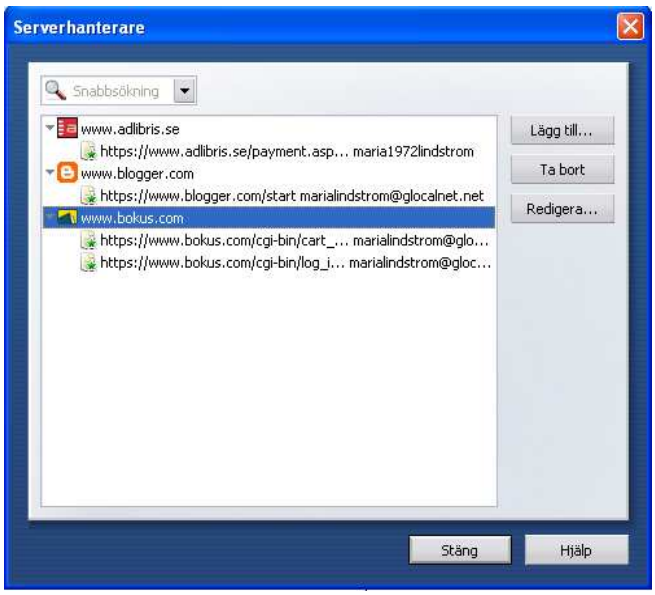
Error when using wand

When using a website where a lot of personal information needed to be revealed (www.bokus.se – an online bookstore) some error occurred – Opera was shut down and an error message was displayed (picture below). This happened many times without any explanations.



UI not resizable

After some usage it was obvious that some UIs were not resizable. Thus it was not possible to see (read) the whole link text.



Confusing verifications

As far as understood the **security status** can be stated either "Verified" or "Not verified". The problem is that there seem to be differences in what these two statuses "verify".

"Verified"

When a website is "Verified" Opera tells the user "This site has been verified by a trusted third party and is not listed on Operas blacklist of suspected fraud sites". Is this one or two verification?

One verification*:*

The third party has checked that the visited website is not on the blacklist (which is verified by a trusted third party).

Two verifications*:*

1. This site has been verified by a trusted third party (first verification?)
2. This site has been checked against a blacklist and is not on the list (second verification?)

Three verifications

The message below can be interpreted as three different messages:

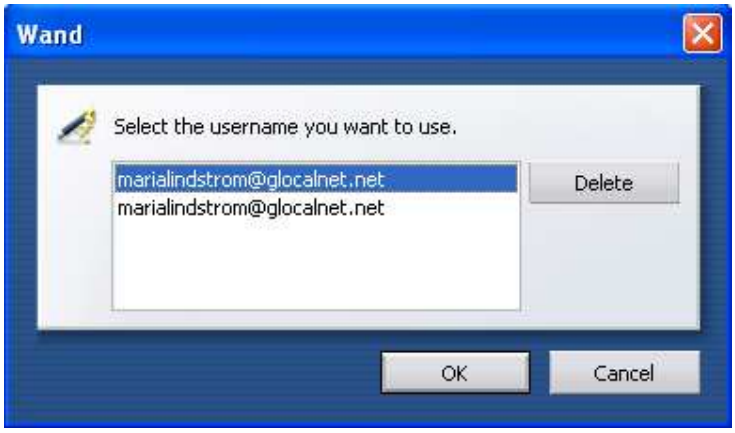
In this message the user gets **three** messages:

1. Opera has no information about the current site
2. The site is not on the blacklist of suspected fraud sites
3. The site has not been verified by a trusted third party

A solution to this problem could be to separate the different messages (verifications) explicitly, both in textual form (differentiate the different verifications in a better language) and in graphics (e.g. by use of different icons for different kinds of "not verified")

Same password – no distinction

When inspecting Opera the same settings (user name and password) were used for all visited websites (to make it easy to remember during inspection). When trying to login by using the "Wand" the problem occurred: it was impossible to distinguish the different logins (e.g. as in picture below).



The problem is that the users is not allowed to "name" or "tag" the different login, e.g. the user would like to name them according to which website they belong, or according to purpose (e.g. shopping, chatting etc) of login.

10.6.4 Usability criteria

In this section some reflections are made about the web browser Opera according to the usability criteria developed in this study.

10.6.4.1 Comprehensibility

Opera is comprehensible after some usage, but at first some things are hard to understand, e.g. to make the "verification" understandable (see "Confusing verification" p.34).

10.6.4.2 Error prevention

Opera is quiet good in giving the user feedback. A lot of messages is given and the "Help" files are rather informative.

10.6.4.3 Task suitability

Opera is suitable for its task. The user is not hindered in its primary purpose, rather assisted e.g. in form filling when buying books.

10.6.4.4 Controllability

The user has control over some things e.g. they gets "certificates" and "security status" from Opera and has the option to set a password for saved data. But the user has no control over their data after data is revealed (e.g. retention, deletion of their data at the service provider).

10.6.4.5 Likeability

Opera is easy to adopt because of the ease of filling in web forms etc. But one could wish for more features e.g. in identity management.

10.7 Safari

Safari (version 3.1) was installed and evaluated the 1th of April 2008.

10.7.1 Short description

According to the website (Apple, 2008) Safaris is: "the fastest, easiest-to-use web browser in the world. With its simple, elegant interface, Safari gets out of your way and lets you enjoy the Web — up to 1.9 times faster than Internet Explorer".

Safari offers "Pop-up blocking", "Tabbed browsing", "Resizable text areas" etc.

10.7.2 Available functions

Safari offers many different functions (some stated above). Interesting for this evaluation is the following:

- **Security**
  "Enjoy worry-free web browsing". Safari uses "robust encryption to ensure that your private information stays that way". A lock icon in upper UI indicates security and offers detailed information about credentials and security certificates.
- **Private Browsing**
  Private Browsing – "to keep your online activities private". With this function turned on nothing about the browsing is stored. If Private Browsing is turned off Safari offers a one-click "cleaner" to empty the cache and clears all browsing data (function called "Privacy Reset").
- **Form AutoFill**
  Form AutoFill "completes online forms with information from the users address book or from previous completed forms". Safari also offers automatic fill-in of account name and passwords – "all stored in a secure encrypted format".

The functions within Safari map **to some extent** into the following IMS functionality (see section 2.3 above):

- **Identity management** – Even thought Safari does not fulfil *all* requirements within the identity management, but to some extent they do. The user is able to manage different digital identities for internet browsing, and these identities can be configured. But Safari does not assist the use of pseudonyms and "Linkability".
- **Authentication management** – A user of Safari can easily manage passwords and other login data. Safari also handles certificates and credentials. The user is also able to fill in web forms assisted by simplified actions within Safari.
- **Context-monitoring –** To some extent Safari also facilitates context monitoring. It identifies the context within Safari (e.g. contextual help files) and also the context on the service side (e.g. the user has been here before…).
- **From filling -** "Form AutoFill" completes online forms from address book or previous forms.

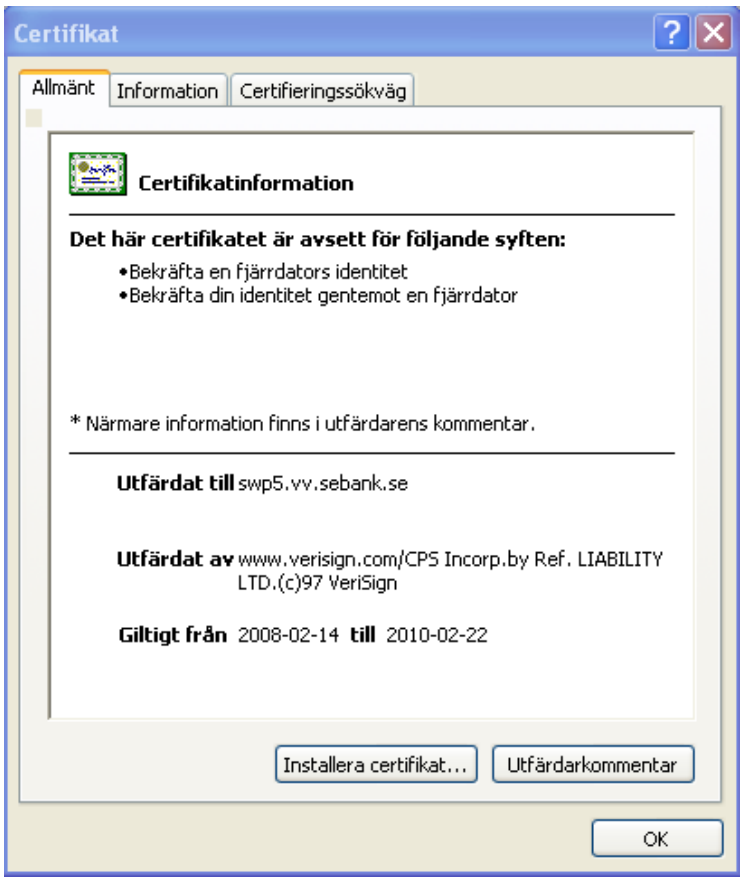
10.7.2.1 Security

When a user enters websites that handles private and economical data one should control that the website is "safe". Safari offers security certificates to ensure safe communications. These certificates are issued by trusted organizations e.g. VeriSign, Inc. and RSA Security, Inc. Safari checks if the visited websites have legal certificates. When a user enters an authorized website (with approved certificate) a small lock icon is displayed (marked in picture below) in the upper right corner of UI. Unfortunately the icon is **too small** and would probably be missed if the user does not know to look for it (more about this in chapter below, called "Usability problems").

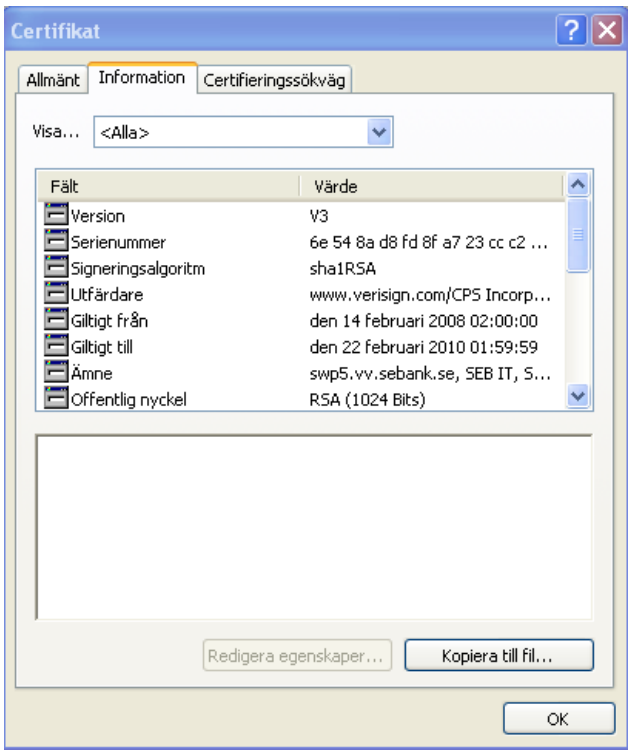
If the user clicks on the lock icon **detailed information** about the certificate is displayed, in three different tabs.

The **first tab** is called "General" and contains information about:

- The purpose of the certificate
- Issuer of certificate
- Who the certificate is issued for
- Date of certificates validity etc.

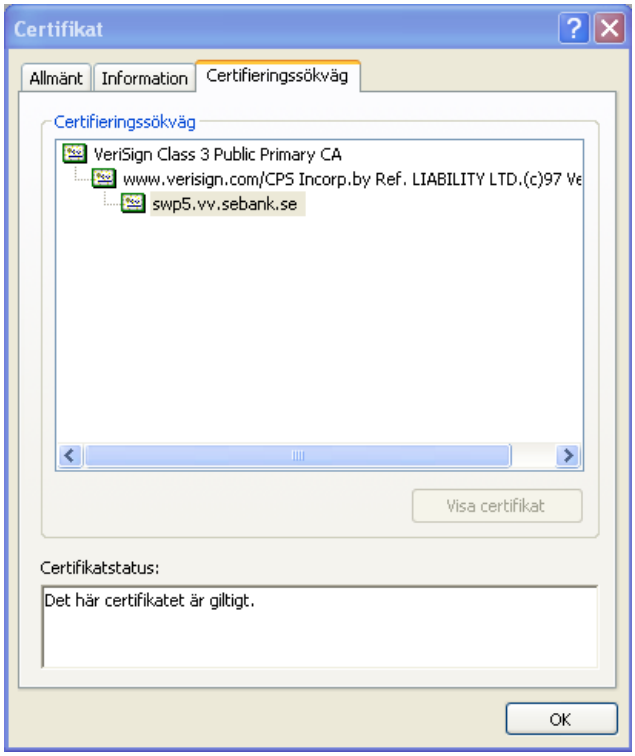


The **second tab** (below) is called “Information” and contains details about the certificate such as version number, serial number, algorithm etc.



Picture "Information" tab

The **third tab** is called "Certificate Search path". It also displays the **status** of the certificate (the certificate below is valid…).



Picture "Certificate Search path" tab

If a website cannot transfer data in a secure way Safari will **notify the user with a message**. Thus the user is able to abort transmission of data. If the user does not want this kind of warnings the "security message" can be turned off.

Another security function is about **saving login data.** When a user enters login data he/she is asked if Safari should save these data to facilitate next login. The user has three options (from picture below):

- Never, at this website
- Not now
- Yes

The first option ("Never at this website") puts the website on a list of sites where login data is "denied" storage. This list is editable by the user, thus the user is able to delete sites from the list whenever the user wants to.

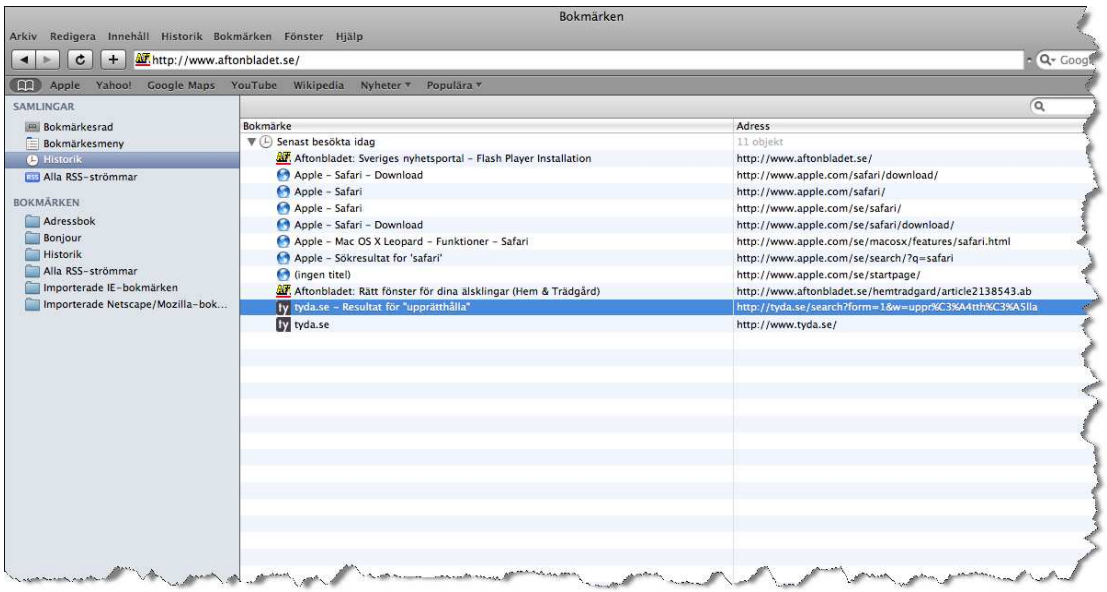
10.7.2.2 Private Browsing

When a user brows the Internet a lot of information is stored e.g. visited websites, its content, user name, password, credit card number etc. If the user doesn't want this information to be stored the "*Private surfing*" feature can be turned on, which means that:

- Visited websites are not stored in "history" list.
- The names of all downloads are erased from the "down load list".
- Information from "auto forms" (e.g. name and password) is not saved.
- Search words are not saved
- Cookies are deleted

This is a really good feature, since the user can be provident and decide to not save history data even prior browsing. Private Surfing must always be turned **on by the user** (not a default setting) and can be done at any time.

If the user hasn't activated "Private surfing" data is stored e.g. about visited websites (the **history of browsing**). But the user is able to delete this information afterwards. This can be done either by deleting all data at once or by manually selecting what history to delete (manual deletion in picture below). Safari can also delete this information automatically.

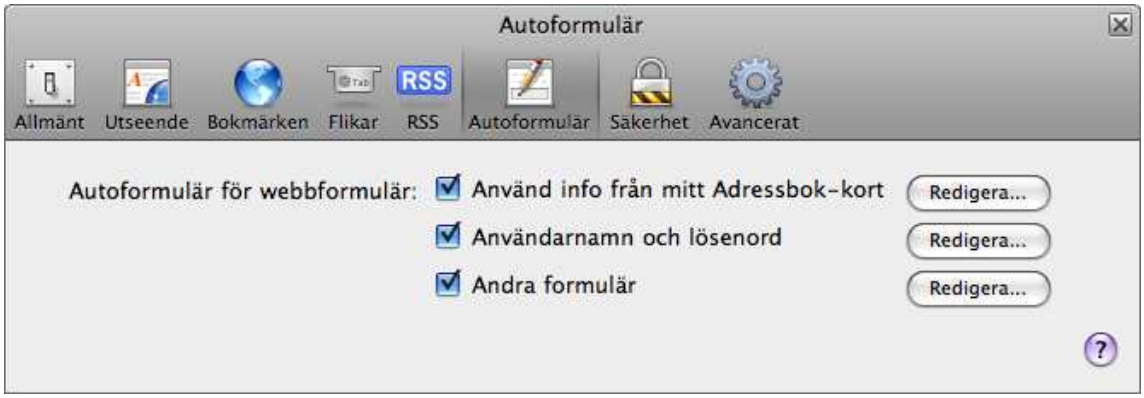


Picture: Manually deletion of history

Safari also has a simple-click "Reset" function. By clicking this *all* information is deleted within the Safari. This can be useful if you use a shared (public) computer and not want anyone else to see this information.

10.7.2.3 Form AutoFill

Safari can help the user to fill in forms at websites, e.g. your name, address and account information when buying things on Internet. This "AutoFill" can be done in different ways by using this "menu":



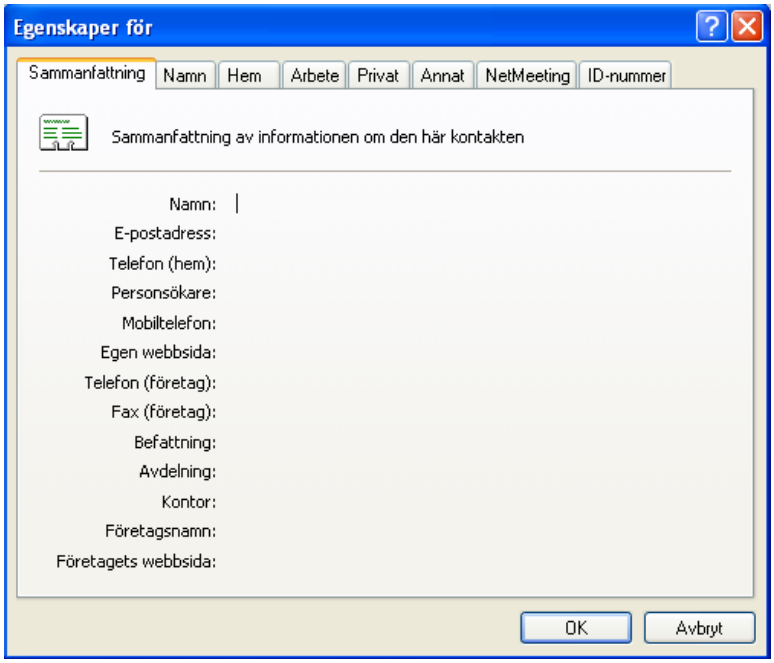
Picture "AutoForm menu"

The options are:

- **Use info from address book** – Safari uses the information from the address book (info needs to be pre-filled in by the user) to fill in forms at websites. If the address book is empty, then *no* data can be automatically filled in (picture 1, below).
- **Username and password** – Safari can help the user to fill in username and passwords at websites that demands login data. When entering data for the first time the user gets a question if Safari should remember login data. And the next time the user enters this website login data is already filled in (picture 2, below).

- **Other forms** – If this feature is activated Safari saves the information the user enters in different web forms at Internet. This data can later be used in automatic form filling (picture 3, below).

The following three UIs are activated by the three different **"Edit" buttons** (Swedish "Redigera") in picture above ("AutoForm menu"):

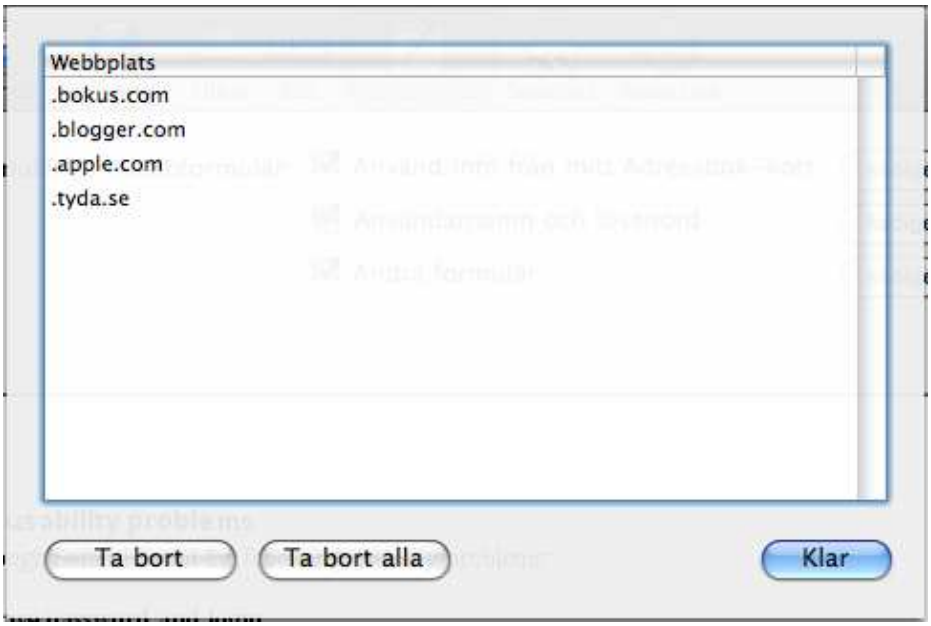


Picture1, activated by "Use info from address book"

The address book was empty. To make this feature work the user have to pre-define data in the address book.

Picture2, activated by "Username and password"

This list present different websites were I have revealed login data. Unfortunately the user is not allowed to *name* these identities (e.g. use pseudonyms).



Picture3, activated by "Other forms"

This list consist different websites where web forms were filled in during inspection (with personal data).

10.7.3 Summary of usability problems

When walking through Safari the following potential usability problems were found:

Private Surfing – not "Private"

Safari offers a function called "Private surfing". From the help files of Safari one can read that "*Private surfing*" means that:

- Visited websites are not stored in "history" list.
- The names of all downloads are erased from the "down load list".
- Information from "auto forms" (e.g. name and password) is not saved.
- Search words are not saved
- Cookies are deleted

To an "ordinary user" the name of this feature can indicate that he is browsing the internet without leaving any traces. This is not really true. The user of Safari is not "private" to the website they visit, only on local computer, since Safari users are not hiding their IP-address and the data transferred is not sent anonymous.

The function "Private surfing" is only private to the *local* computer. E.g. if a public (shared) computer is used for browsing the "Private surfing" feature is good, since no other user of that particular computer can observe former browsing and internet activities. But the browsing is register at the service providers (websites they visit).

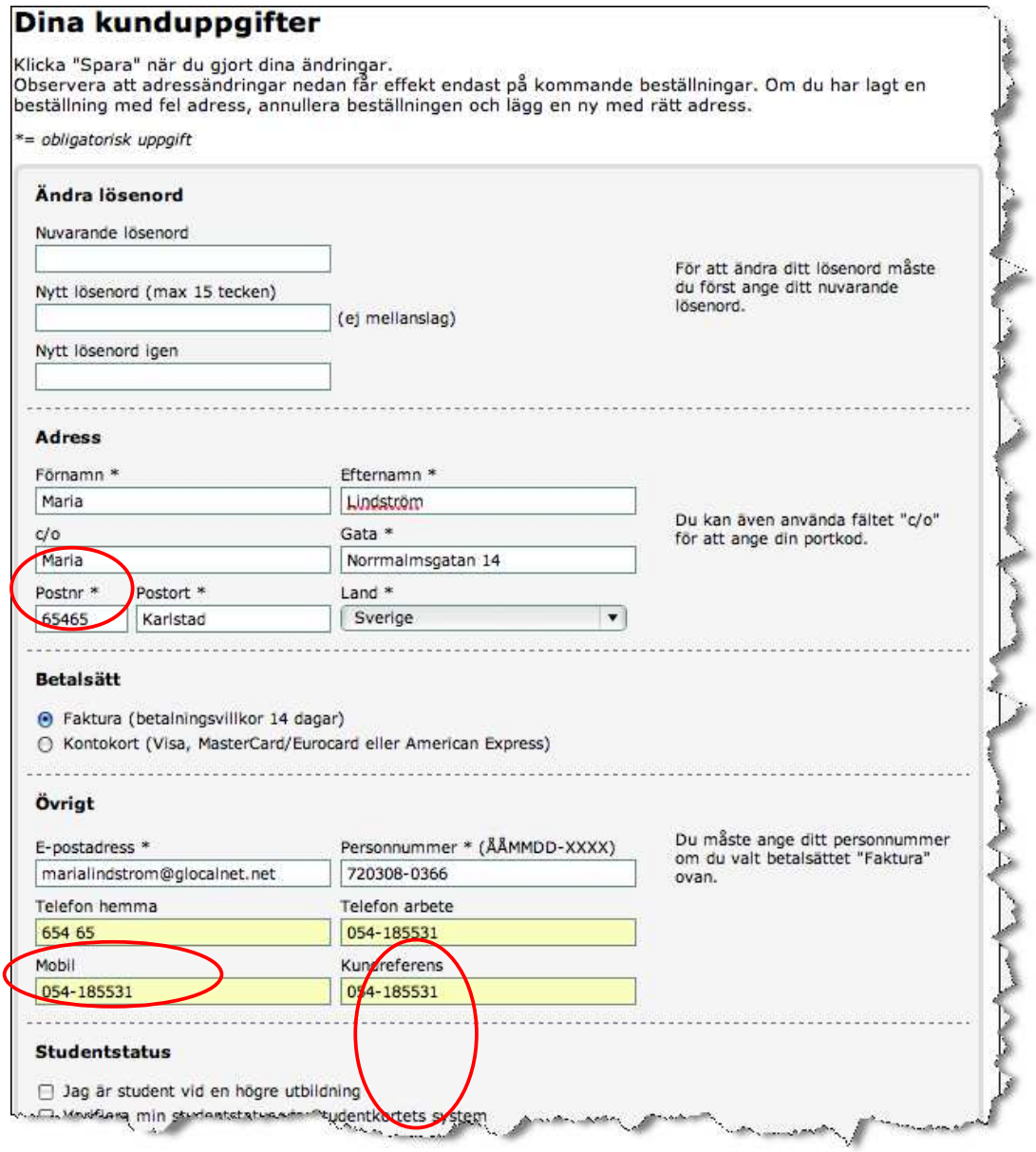
Thus the option to surf "private" on *local* computer is a good feature, but must be **named** understandable (elucidate that the privacy is only at the local computer – not towards internet providers.).

Address book revealed wrong data in web form

An error occurred during inspection of Safari. When trying to automatically fill in a form (using pre-filled data from address book) data was filled in at *wrong* places (see picture below).

Examples of errors:

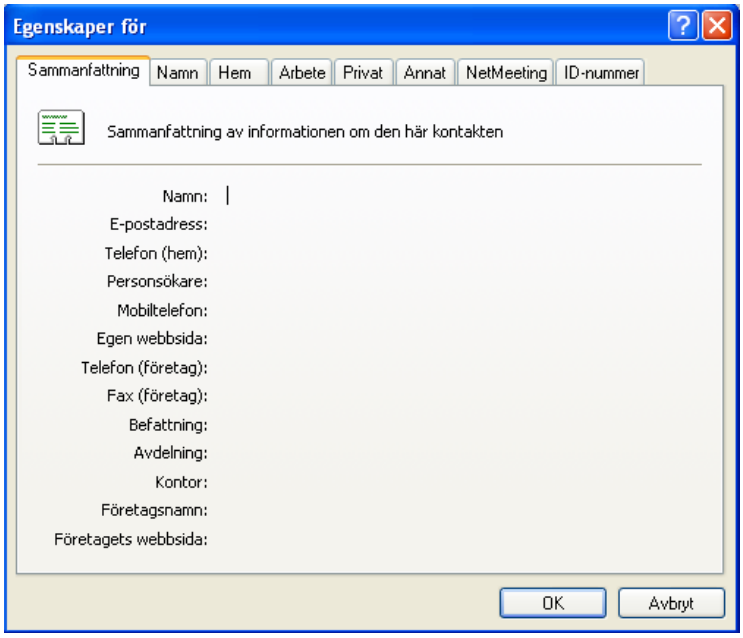
- at "c/o" "first name" was inserted
- at "Home phone number" the area code of Karlstad was inserted.
- Home phone number was inserted at three different places (at "Mobile no", "Work number" and "customer credential")



Hidden address book

Another problem is that the address book is hidden deep down (low hierarchical) in Safari and is hard to find. To be able to fill in information in the address book the user has to do a lot of clicking – Edit> Settings>AutoForm>“Use info from my Address book”>Edit…

After these steps the following UI is displayed:



Picture "Address book"

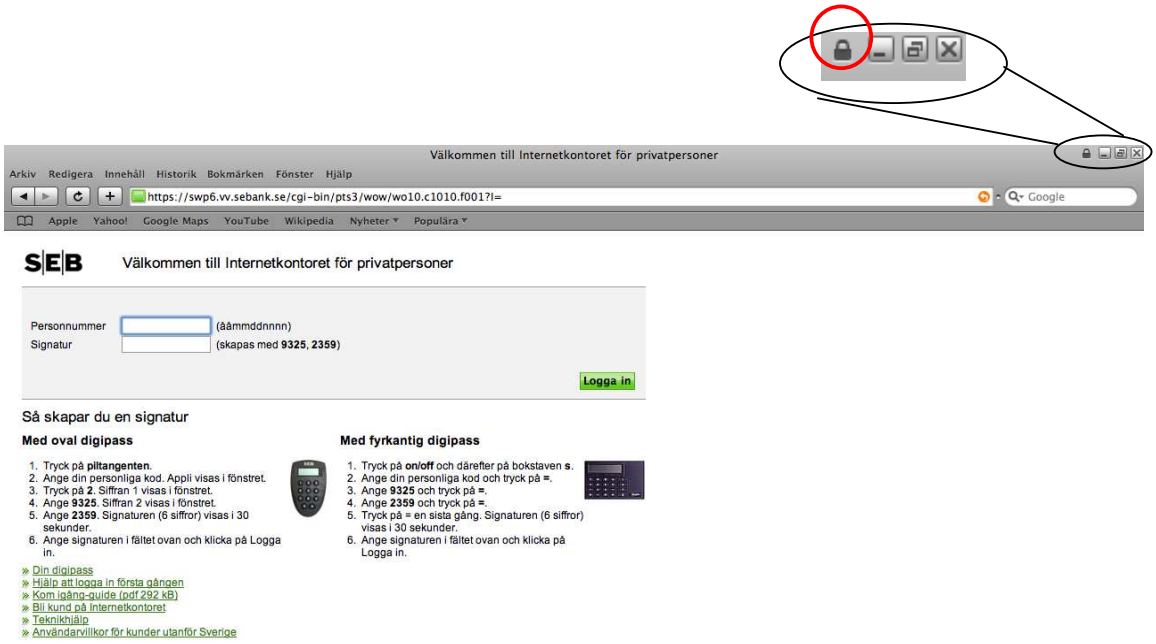
To be usable the address book needs to be present higher up in the menu hierarchy.

AutoForm could not fill in form

During inspection a message was telling that "data is **not available** in the address book" even thought at least *some* of the data was. At this point the one would have hoped that Safari at least should have filled in these data, thus the user only have to complement with data missing in address book..

Too small icons

Security certificates are displayed by a small lock icon in Safari (in the upper right corner of UI below). Unfortunately the icon is **too small** and would probably be missed if the user does not know to look for it. A better solution may be to put this icon into the address field.



Not able to name identities

Safari offered a list of websites where the user have revealed login data. Unfortunately the user is not allowed to name these identities (e.g. use pseudonyms).

Sensitive data in address book

Since the address book contains sensitive data it would be good to protect these data with a "master" password, which also supports privacy on the local computer. This is not facilitated within Safari today.

In some cases the user may want to delete all data within the address book (e.g. the usability inspector wanted this feature after finished inspection). This is not facilitated in Safari today. This is probably an unusual way of usage, thus probably not a common problem. The function "Reset Safari" was deleting all data except data in the address book.

10.7.4 Usability criteria

In this section some reflections are made about Safari according to the usability criteria selected for this study.

10.7.4.1 Comprehensibility

Safari is comprehensible. The help files are really good. The text is short, informative and written in a simple language helping the user to understand things that otherwise could be hard to interpret.

10.7.4.2 Error prevention

Safari is really good in error prevention. E.g. they offer “Privacy Reset” which is a function that erases tracks of the users browsing. Safari also facilitates messages if something is wrong. E.g. if a website can’t maintain a secure transition of data it notifies the user prior sending data so the user can abort transition if he/she not want to “risk” anything.

10.7.4.3 Task suitability

Safari is also suitable for its task, especially filling in password and to use for “Private Surfing”. The user is not hindered in its primary purpose, rather assisted.

10.7.4.4 Controllability

Safari offers detailed information about security certificates and credentials. The function “Privacy Reset” is also implying control, since the user can erase browsing data if he wants to.

The user of Safari has many options (extra finesses) that easily can be turned off if the users want to.

Safari facilitates a simple-click “Reset” function, which gives the user control over his/hers data. By clicking this all (absolutely all) information is deleted within the Safari.

10.7.4.5 Likeability

Safari should be likable to most Internet users. It assisted in filling in passwords, and it also allowed “private” surfing (very easy one-click).

10.8 Firefox

Firefox (version 2.0.0.13, released 25 march 2008) was installed and evaluated the 3th of April 2008.

10.8.1 Short description

According to Firefox website (Mozilla, 2008) "Firefox 2 is full of helpful features to make your time online more productive".

Firefox offers "Tabbed browsing", "Spell checking", "Pop-up Blocker", "Handle cookies", etc.

10.8.2 Available functions

Firefox offers many different functions (stated above). Interesting for this evaluation is the following:

- Secrecy
- Security
- Certificate

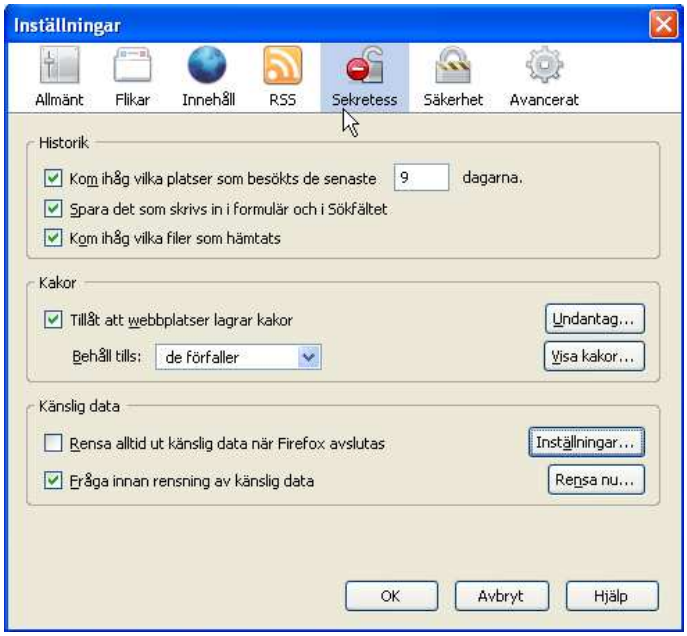
These functions map into the following IMS functionality (see section 2.3):

- **Authentication management**
- **Context monitoring**
- **Form Filling**

10.8.2.1 Secrecy

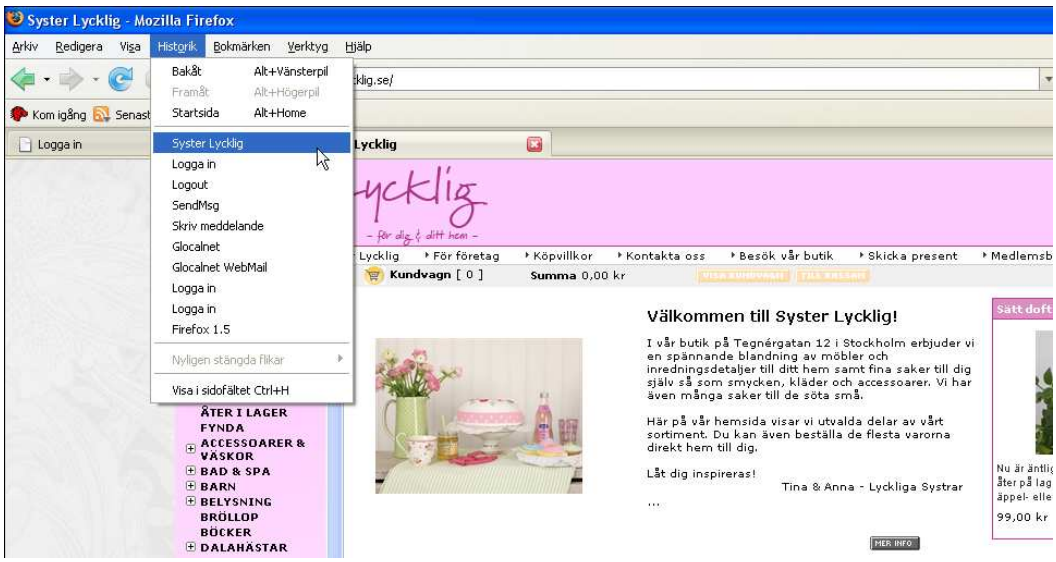
In the secrecy folder the user can set conditions for:

- History
- Cookies
- Sensitive data



History

In History the user can set for how long visited web addresses will be saved, if fill in data should be saved or not, and if Firefox should remember what files have been downloaded. The history list (picture below) only shows what websites the user has visited. Thus this is not a history management according to requirements (section 2.3.6 above) of an IMS (e.g. "review all data one have transferred to a particular service provider").



UI "History list"

Cookies

The user is allowed to set conditions for cookies.

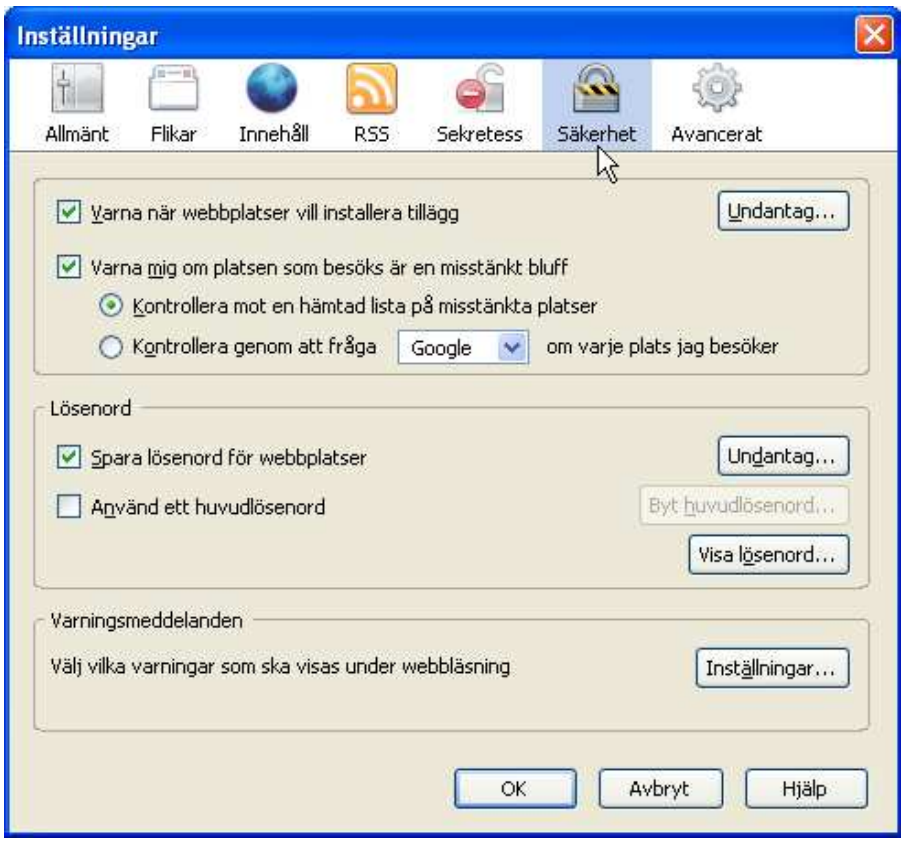
Private Data

The last secrecy feature is about "Private data". Firefox gives the option to "always erase sensitive data when Firefox is closed" and "Ask before erasing sensitive data". Firefox also offers the ability to **clear all your private** Web browsing data with just one click (Tools) - with the "Clear Private Data" tool (snapshot below).

10.8.2.2 Security

In the "Security folder" the user can set conditions for:

- Warnings
- Password
- Phishing Protection



UI "Security"

Warnings

The user can be warned about the following:

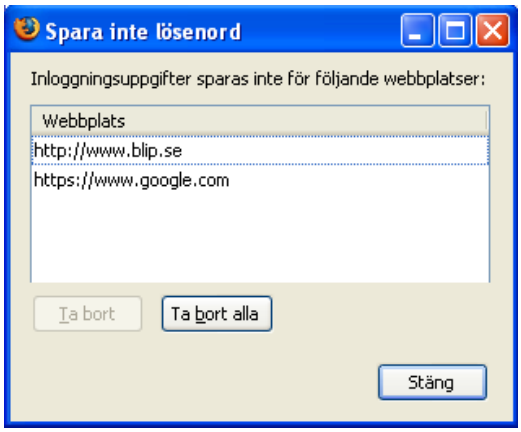
- Warn before installation
- Warn if website is suspicious (e.g. phising protection)
  - Control against list of suspicious sites
  - Control by asking…at every site I visit

Password

The user can set the following passwords conditions:

- Save password for websites
- Don't save password for these sites
- Use a main password

UI "Saved password"



UI "Don't save password at these sites"

Password facilities

When login data is saved the users is facilitated with either **pre-filled** login data, or the user gets **alternatives in drop-down list** (se picture below).

Another good thing about Firefox is that the user can enter a **"master" password** to hide all his login data within his local computer. This means that other users of the same computer can't observe others login.

Password **can be erased**, but this feature is hard to find and too low in hierarchy. (Tools>Settings>View password>Delete all data)

Phishing Protection

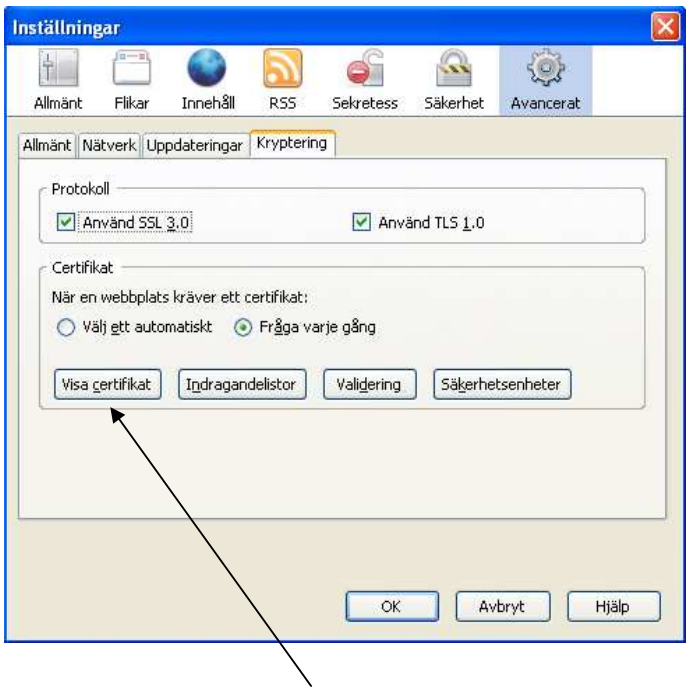
According to Firefox website phishing protection "takes security to a new level, helping to safeguard your financial information and protect you from identity theft. When you encounter a Web site that is a suspected forgery (known as a "phishing" site) Firefox will warn you and offer to take you to a search page so you can find the real Web site you were looking for" (Mozilla, 2008).

UI phishing protection

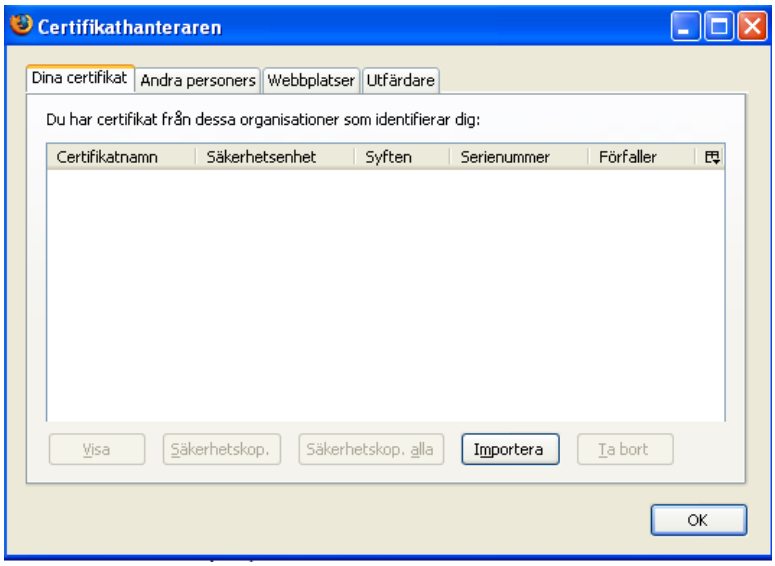
10.8.2.3 Certificate

Certificates help the user establish safe connections to websites. When a website demands a certificate the user can choose between

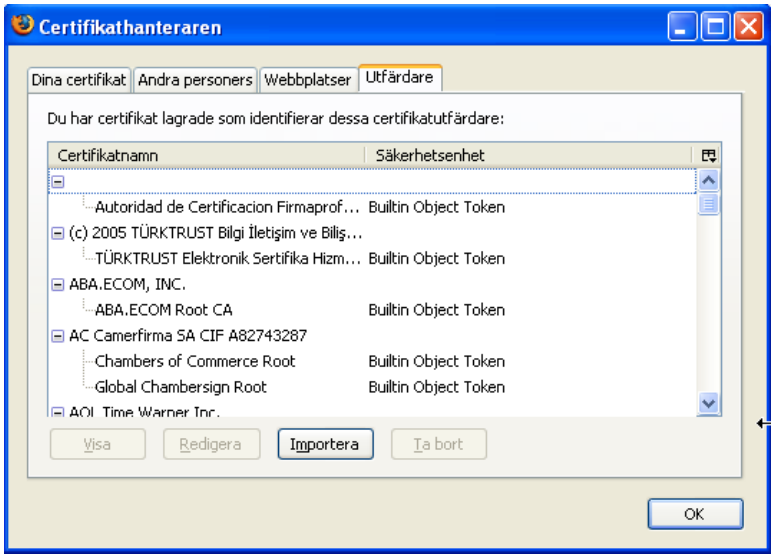
- choose a certificate automatically
- ask each time



The user can also "View certificate" (picture below) - view details such as "your certificates", "Other persons certificates", "Websites" and "issuer".

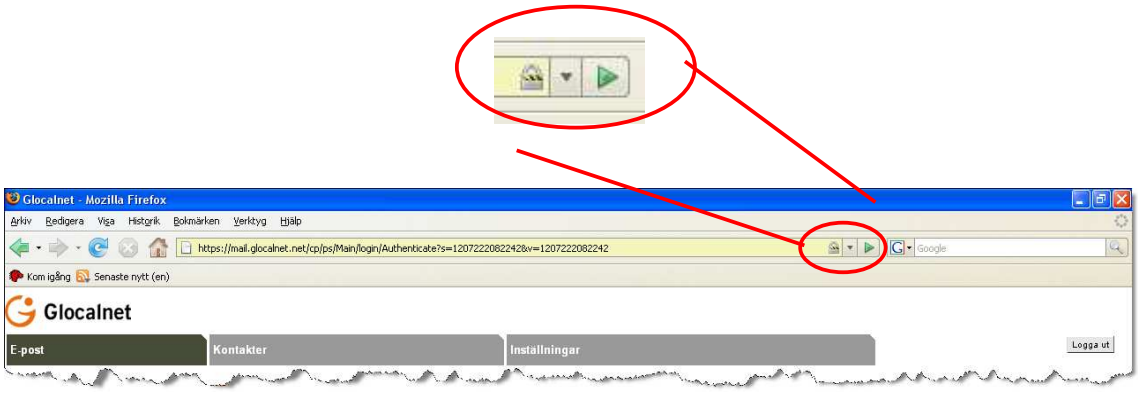


UI "View certificate – Your certificate"



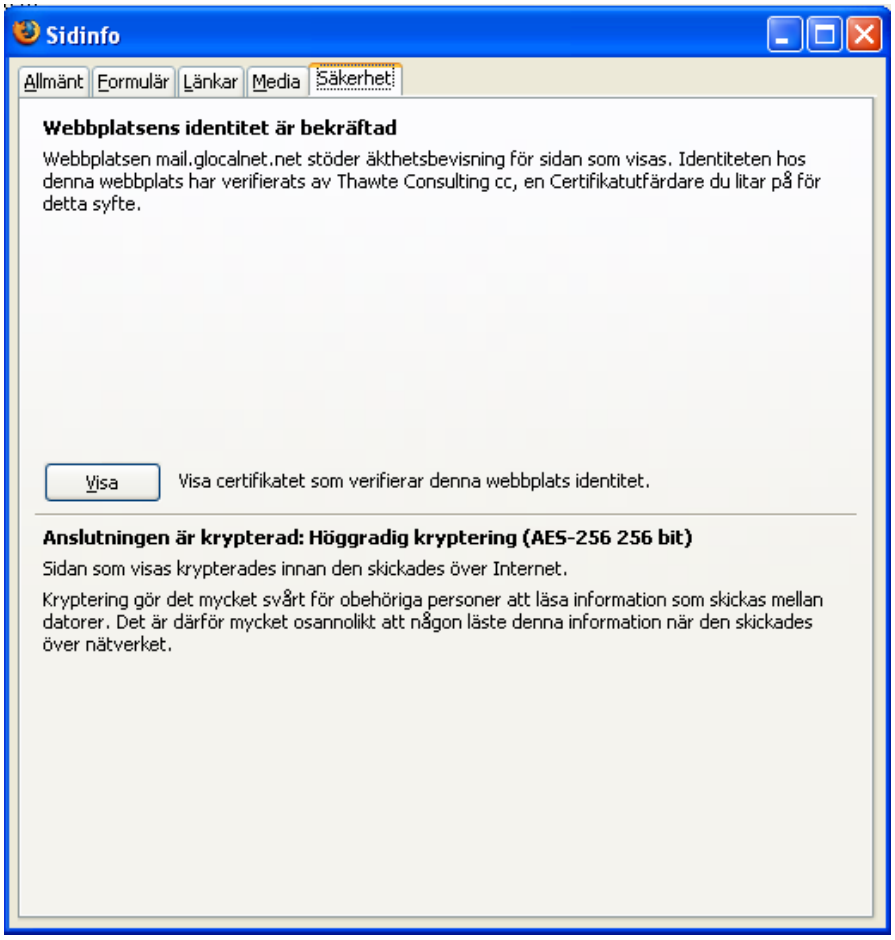
UI "Issuer"

When entering a website where a safe connection is established the address field turns **yellow** and a **lock icon** is displayed, which means that Firefox uses double indication of safe connection.



UI "Yellow address field"

Clicking on the lock icon (picture above) activates a "security confirmation" (picture below), which tells the user that the identity of the visited website is confirmed.

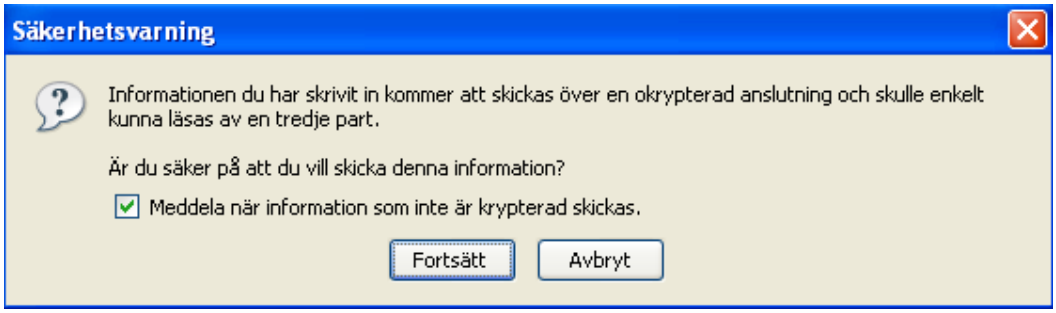


UI "Security confirmation"

When testing Firefox the usability inspector tried to log in at different websites. At one website wrong login data was entered (login data was forgotten and correct login data was asked for in an email). Then this message was displayed – a security warning:

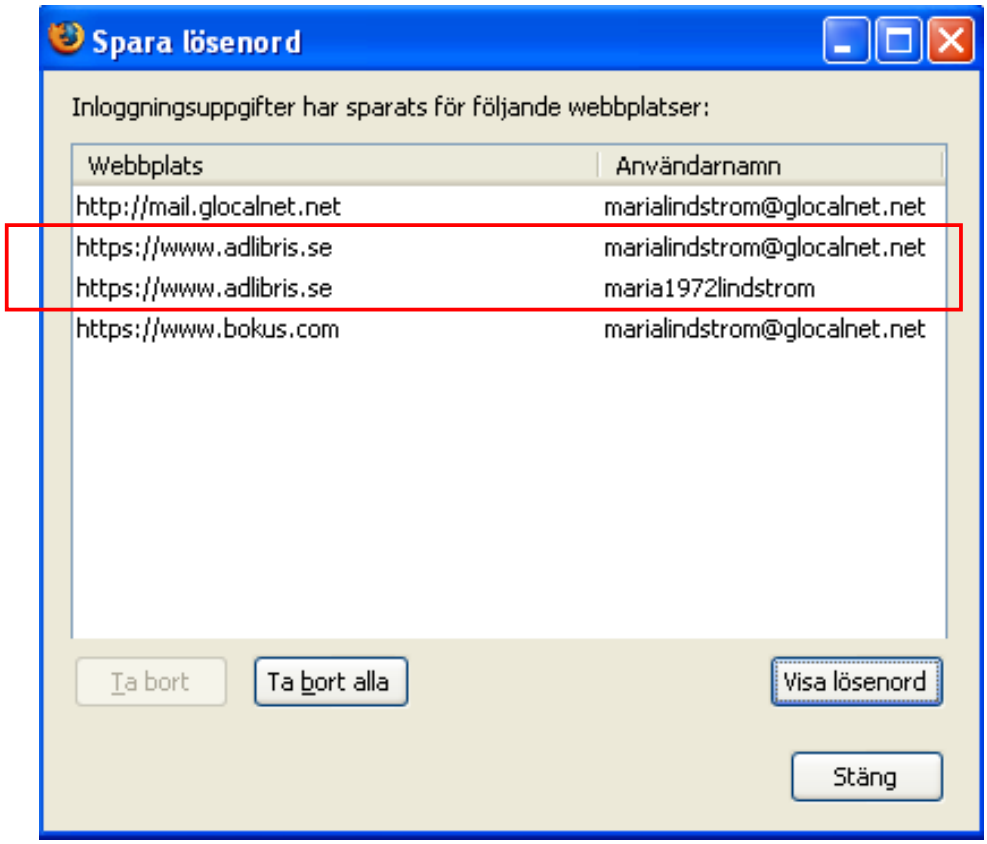
The information you have entered will be sent by a decrypted connection and could therefore easily be read by a third party.

Do you really want to send this information?



UI "Security warning"

After receiving the correct login data the usability inspector returned and tried to log in again. And again a question came up asking if login data should be remembered. After answering yes the interesting was to see what had happened to the wrong login data – was it saved or deleted?



UI "Saved passwords"

Surprisingly Firefox had **saved the incorrect login data** too, which lead to having **two different login to the same website** (picture above), where one of the login is incorrect.

10.8.3 Summary of usability problems

When walking through Safari the following usability problems were found:

Session Restore

Firefox offers a function called "Session Restore", which means that "if Firefox has to restart or closes when it comes back you'll pick up exactly where you left off. The windows and tabs you were using, the text you typed into forms, […] will all be restored". This is a good feature, but it can also be a threat to one's privacy. Imagine if it is *another* user switching on the computer…then this user can "restore" and observe everything the previous user did. If the user don't understand this underlying consequence this is also a usability problem.

Wrong data was saved

Once wrong login data was entered at a website leading to refused enter, but at the same time the incorrect login data was saved. This also implied that the user had two login data for the same website…

Hard to find important functions

In Firefox the user can save passwords in a list, and if he/she wants to these passwords can be erased. But this feature is really hard to find since it is placed too low in hierarchy. (Tools>Settings>View password>Delete all data)

Similar functions separated in UI

As one can see in the UI below there are three areas with options, separated by a thin line. The upper area is about different warnings. By some reason the third area (lowest) is also about warnings. To make an UI easy to use it is important to group elements. To separate similar elements can confuse the user and makes it harder to find the element.

Missing headline

Another usability problem in UI above is that the upper area misses a “headline” or a “tag” telling the user what’s in it. A usable UI is also consistent – thus *all* areas should be named.

10.8.4 Usability criteria

In this section some reflections are made about Firefox according to the usability criteria selected for this study.

10.8.4.1 Comprehensibility

Firefox is OK to use. The help files are good, but Firefox does not imply trust. And it was not helping the user with much (just filling in password).

10.8.4.2 Error prevention

Just like Safari, Firefox offer the user to erase tracks of browsing. In Safari this was called “Privacy Reset” and in Firefox it is called “Delete Private data”. Firefox also gives the user feedback and displays warnings in an explicit way (double indication by using both lock icon and yellow coloured address field) which is error preventive.

10.8.4.3 Task suitability

Firefox is suitable for its task. It is facilitating features like form filling and is not hindering users’ primary purpose, rather assistive.

10.8.4.4 Controllability

The user can make some settings about how Firefox should behave. If the user wants to he/she can also get warnings if something harmful is happening.

One feature in Firefox is threatening the user’s privacy. It is called “Session restore”, were an earlier web session can be restored afterwards. This means that another user (on local computer) can restore and observe another users web session. A user must be able to erase his session is he wants to – to be able to control his own privacy.

Firefox offers information about security certificates and credentials. The function “Clear Private Data” is also implying control, since the user can erase browsing data (NOTE! only erase data from *local* computer) if he wants to.

10.8.4.5 Likeability

Some features were really good in Firefox (e.g. main password). Firefox assisted in filling in passwords, and one good feature was the double indication of security status when entering a website with safe connection (both icon and colour-indication).

Some negative comments are that it was impossible to find the personal form filler data. It is stored somewhere (if the user wants to configure it) but could not be found.

10.9 Internet Explorer

Internet Explorer (version 7.0) was installed and evaluated the 4th of April 2008.

10.9.1 Short description

According to Microsoft website Internet Explorer offers "new security features [that] make you safer than ever". "Internet Explorer offers new ways to better protect against the **theft of personal data** from fraudulent websites, a practice known as phishing" (Microsoft, 2008).

Internet Explorer 7.0 offers "Tabbed browsing", "Advanced printing", "Instant search box", "Favorites center" etc.

10.9.2 Available functions

Internet Explorer offers many different functions (e.g. as stated above). Of interest for this evaluation are the following functions:

- **Security Status bar**
- **Phishing filter**
- **Cross domain barriers**
- **Delete browsing history**
- **Address bar protection**
- **International domain name anti-spoofing**
- **URL handling security**
- **Fix my settings**

Information about the functions (below) is taken from Internet Explorer website (Microsoft, 2008).

Security status bar

"Color-coded notifications appear next to the address bar to make you aware of website security and privacy settings. The Address Bar changes to **green** for websites bearing new High Assurance certificates, indicating the site owner has completed extensive identity verification checks. Phishing Filter notifications, certificate names, and the **gold padlock icon** also appear next to the address bar for better visibility. You can easily display certificate and privacy detail information with a single click on the Security Status bar".

The colour of the Security Status bar tells you whether the certificate is valid, and it displays the level of validation that was performed by the certifying organization.

The following table describes what the Security Status bar colours mean.

Colour	What it means
Red	The certificate is out of date, invalid, or has an error. For more information, see "About Certificate Errors" in Related Topics.
Yellow	The authenticity of the certificate or certification authority that issued it cannot be verified. This might indicate a problem with the certification authority's website.

White	The certificate has normal validation. This means that communication between your browser and the website is encrypted. The certification authority makes no assertion about the business practices of the website.
Green	The certificate uses extended validation. This means that communication between your browser and website is encrypted and that the certification authority has confirmed the website is owned or operated by a business that is legally organized under the jurisdiction shown in the certificate and on the Security Status bar. The certification authority makes no assertion about the business practices of the website.

NOTE!

When evaluating Internet Explorer only “white” colour of the status field were found.

Phishing filter

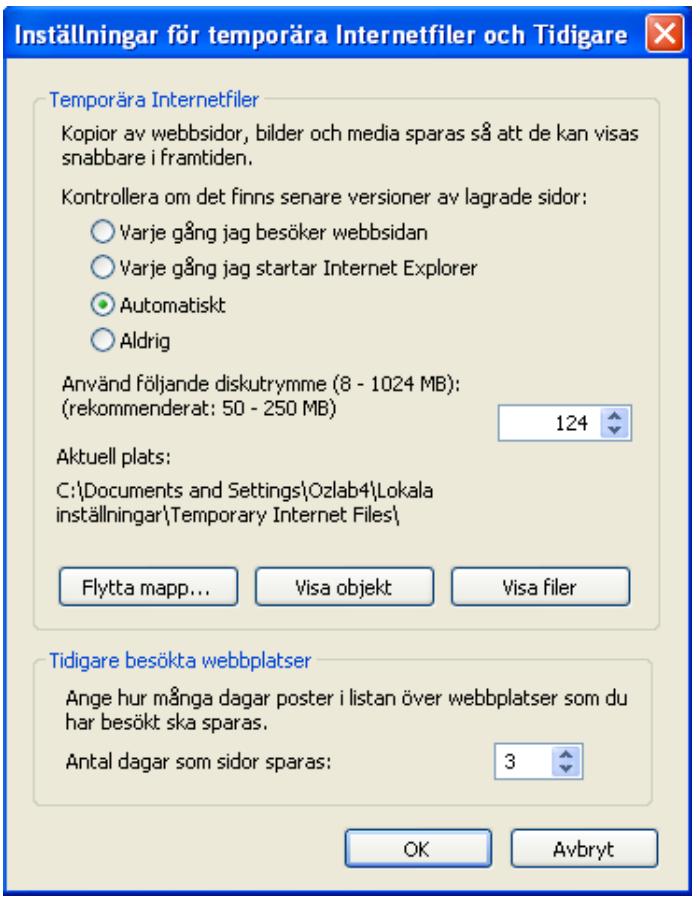
“This filter **warns you** about and helps to protect you against potential or known fraudulent websites, and **blocks** the sites if appropriate. This opt-in filter is updated several times per hour using the latest security information from Microsoft and several industry partners.”

Cross domain barriers

“Internet Explorer 7 helps to prevent the script on webpages from interacting with content from other domains or windows. This enhanced safeguard gives you additional protection against malware by helping to prevent malicious websites from manipulating flaws in other websites or causing you to download undesired content or software”.

Delete browsing history

The user can select how temporary internet files should be handled and for how long (days) visited websites should be saved.



UI "Browsing history" (temporary internet files)

Address bar protection

"Every window, whether it's a pop-up or standard window, will show you an address bar, helping to block malicious sites from emulating trusted sites".

International domain name anti-spoofing

"In addition to adding support for International Domain Names in URLs, Internet Explorer also notifies you when visually similar characters in the URL are not expressed in the same language—protecting you against sites that could otherwise appear as known, trustworthy sites".

URL handling security

"Redesigned URL parsing ensures consistent processing and minimizes possible exploits. The new URL handler helps centralize critical data parsing and increases data consistency throughout the application".

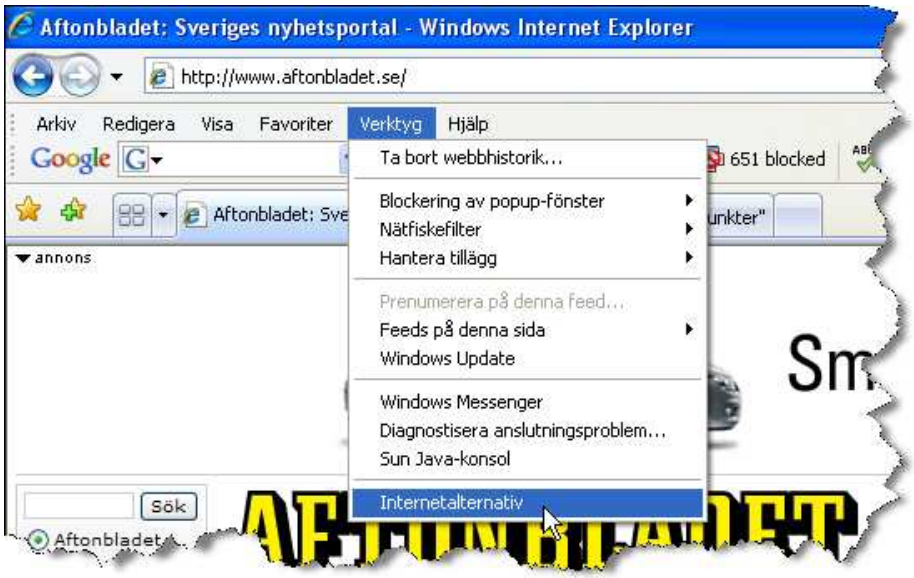
Fix my settings

"To help protect you from browsing with unsafe settings, Internet Explorer 7 warns you with an Information Bar when current security settings may put you at risk. Within the Internet Control Panel, you will see certain critical items highlighted in red when they are unsafely configured. The Information Bar will continue to remind you as long as the settings remain

unsafe. You can instantly reset Internet security settings to the "Medium-High" default level by clicking the "Fix My Settings" option in the Information Bar".

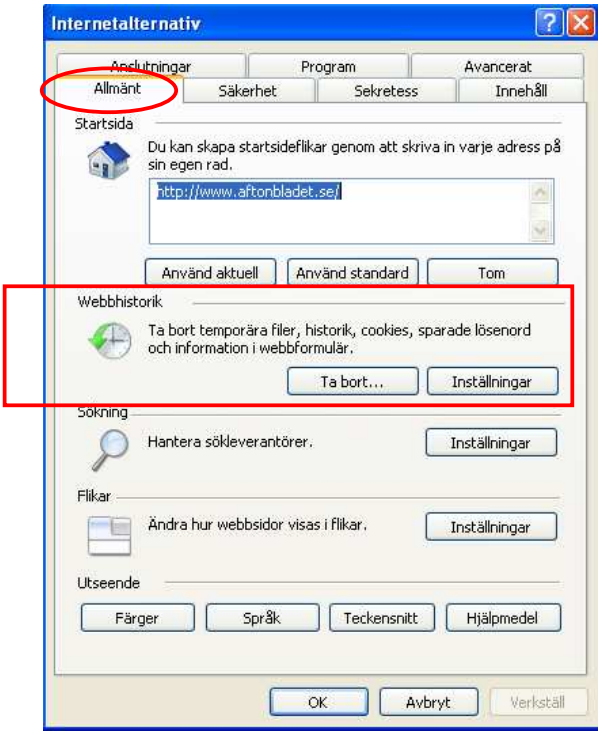
10.9.2.1 Usability inspection of Internet Explorer

Since some of the stated functions above (from Microsoft website) were hard to recognize during inspection, this report will be structured according to menu (e.g. in picture below) within IE rather than the stated available function.

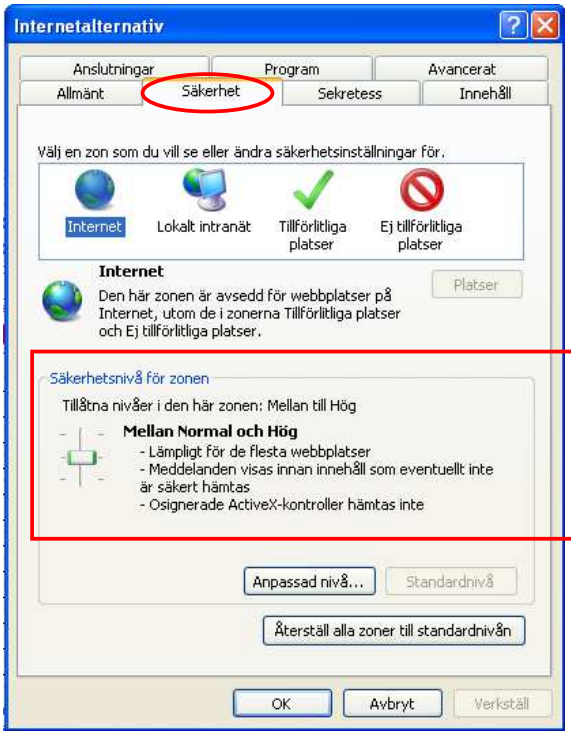


In the "Tool" menu (picture above) the user can "Delete web history", "Block pop-ups", "Handle phishing", "Handle extensions", "Feeds" etc. The last option (blue in picture above) is "Internet alternatives" a kind of "menu" (picture below). The user can make a lot of settings in this UI. But in this report only functions of interest for this study is inspected:

"General"-tab



UI "Internet alternatives – Tab **"General"**



UI "Internet alternatives – Tab **"Security"**

History:

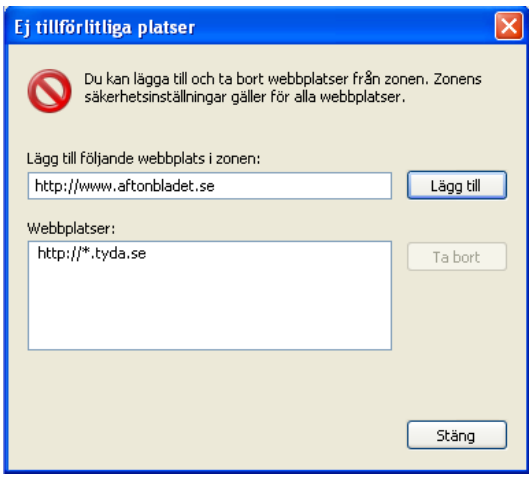
Marked in "General tab" (left picture above) is the option to manage "History". Here the user can "delete" ("Ta bort") or make "Settings" ("Inställningar") about web history.

"Security"-tab

The "Security tab" (in right picture above) handles security. In the lower part of UI (marked) the user can set level of security with a control knob; "high", "medium" or "low" security. Explanations of what the different level implies are displayed next to the control knob.

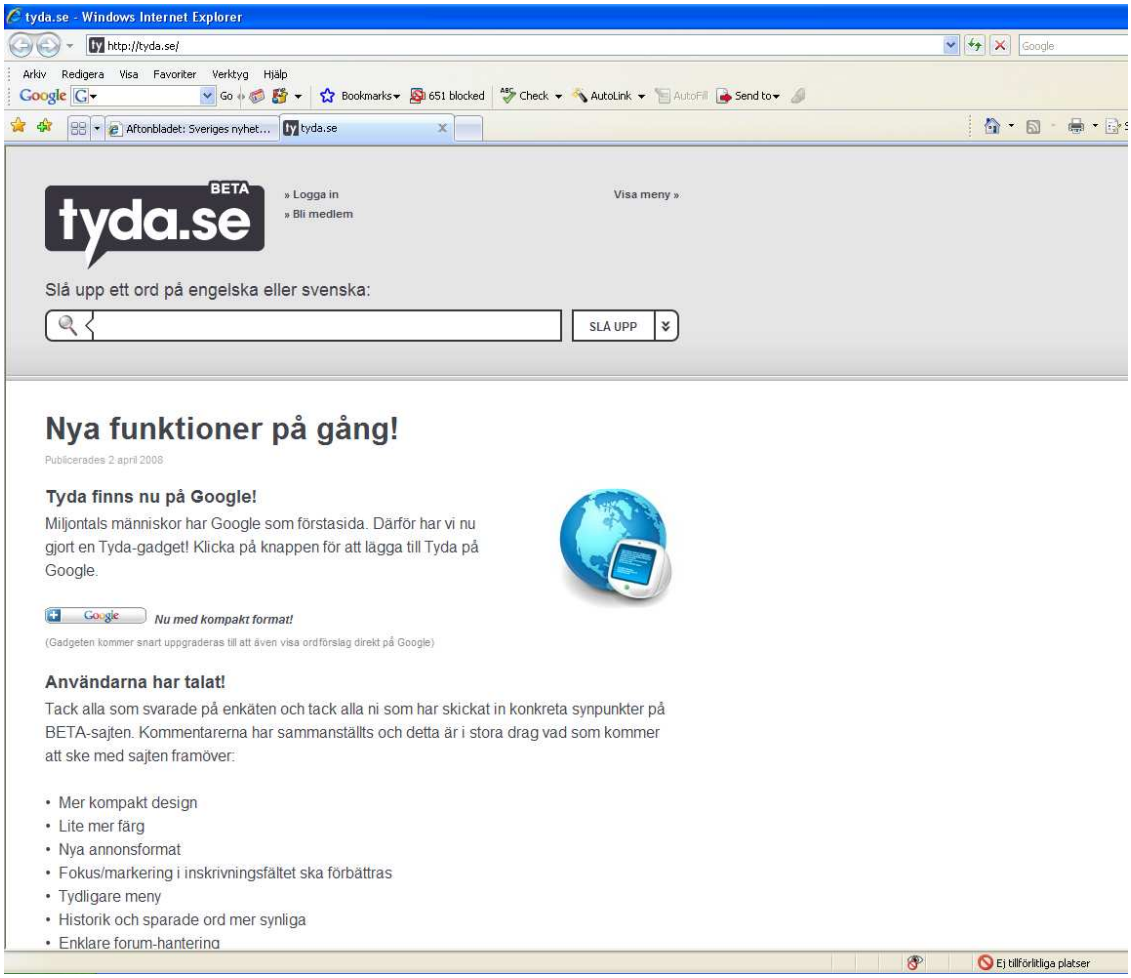
In the upper part of “Security” tab the user has four icons:

“Internet” and “Local network” handles “zones”. The other two icons handle “trusted websites” (green check icon) and “Not trusted sites” (red icon as in picture below). In picture below the usability inspector has “blacklisted” one website, just to try the function.



UI Not trusted websites

But after blacklisting the website confusion took place. First it seemed as nothing had happened (above). But after awhile a very, very small icon had appeared in lower right corner of the window (cf. picture below), a part of UI seldom seen by users. To put a warning at this place in the UI is a threat to both usability and security.



UI An example of a website I have blacklisted. But the warning from Internet Explorer is insufficient.

"Secrecy"-tab

In the third tab, called "secrecy" (left UI below), users can make settings for level of cookie handling (blocking) and block pop-up windows.

UI "Secrecy"-tab labelled "Settings" **UI "Content"-tab**

"Content"-tab

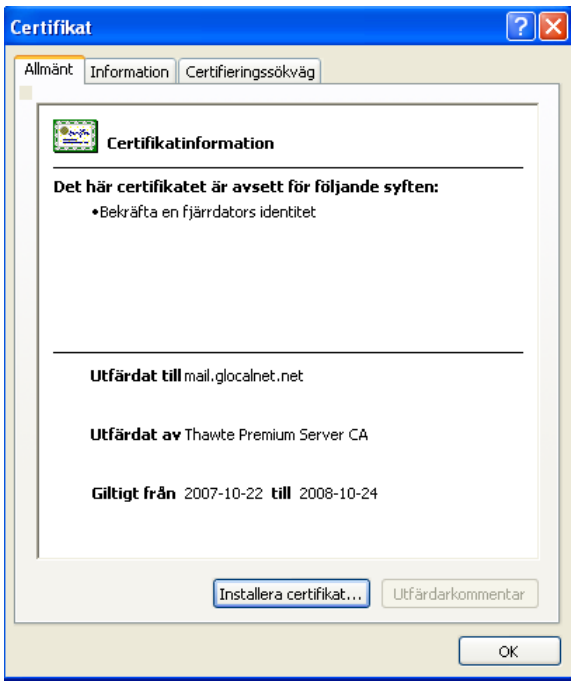
In the fourth tab called "Content" (Swedish: "Innehåll", marked in the right UI above) the users can:

- "Classify content" (control what content can be shown on this computer)
- **Handle "Certificates"**
- **Handle login data**
- Handle feeds

Below only "Certificate" and "Login data" is described.

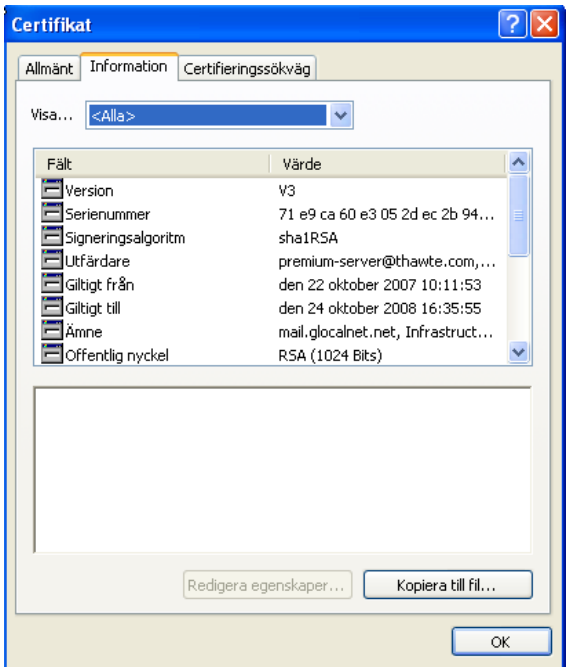
Certificate

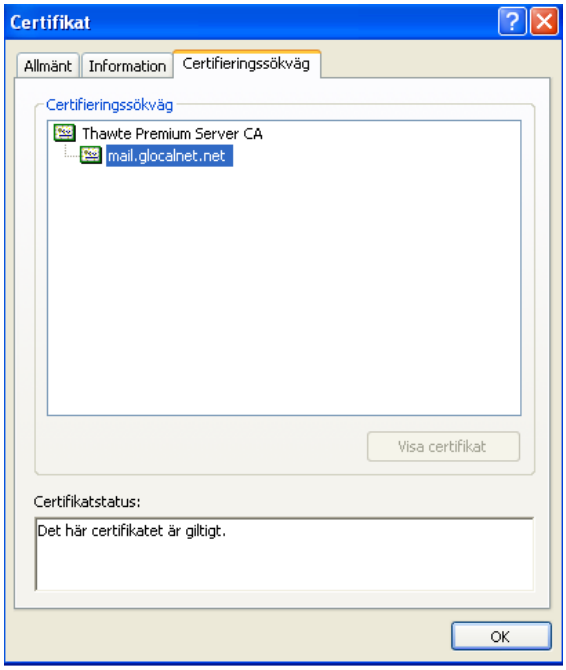
The user can view the certificate (UI below) and get information about purposes of the certificate, information about issuer and date of issue.



UI "Certificate" Tab "General"

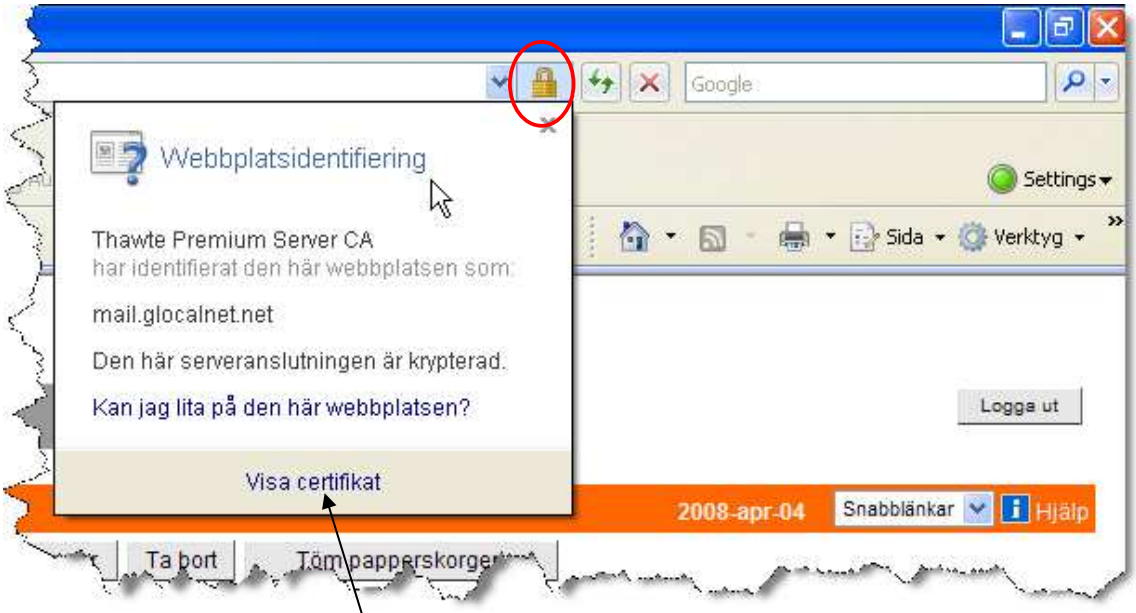
Below are examples of the other two tabs in "Certificate". The "Information" tab has details about version, serial number etc. The other UI is about "Certificate search path".





UI "Certificate" Tab "Information" **UI "Certificate" Tab "Certificate Search path"**

Also when surfing on the Internet the user can view a certificate. A lock icon is displayed in the address bar (indicating a safe internet connection). When *clicking* on the lock icon, **website identification** is displayed (picture below).



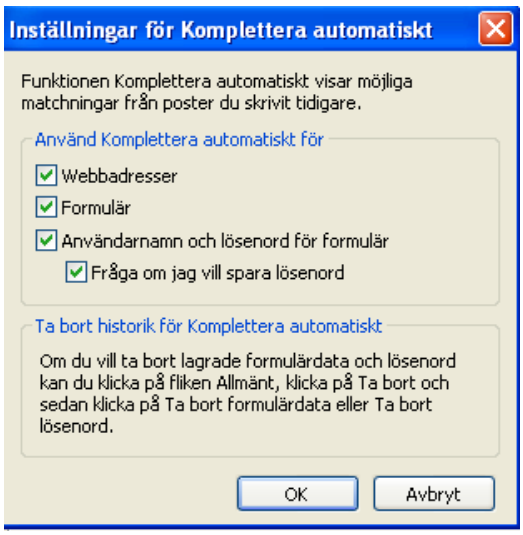
UI – "Website identification"

The message is that "this server connection is encrypted". It also tells that a trusted part (Thawte Premium Server CA) has identified the website.

The user has the option to "view certificate".

Handle login data

In the "Content"-tab (p.63) one area is named (labelled) "Automatically Completion" (Swedish: "Komplettera automatiskt"). This label can be confusing for users, since the content rather is about handling passwords and login data (e.g. for later use). When clicking on the button "Settings" in this area the following UI is displayed:

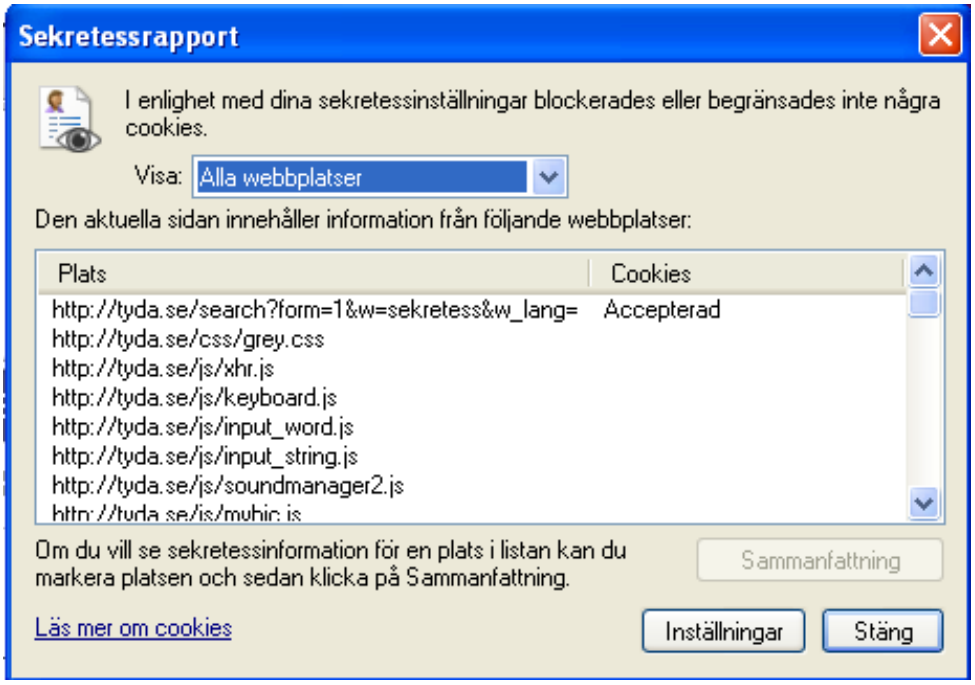


The user can tell Internet Explorer to remember Web addresses, forms, username and passwords for later use. The user is also allowed to delete all history data. (Unfortunately the user doesn't get direct access to "delete history" here, only and textual instruction of where to find it…

Secrecy policy

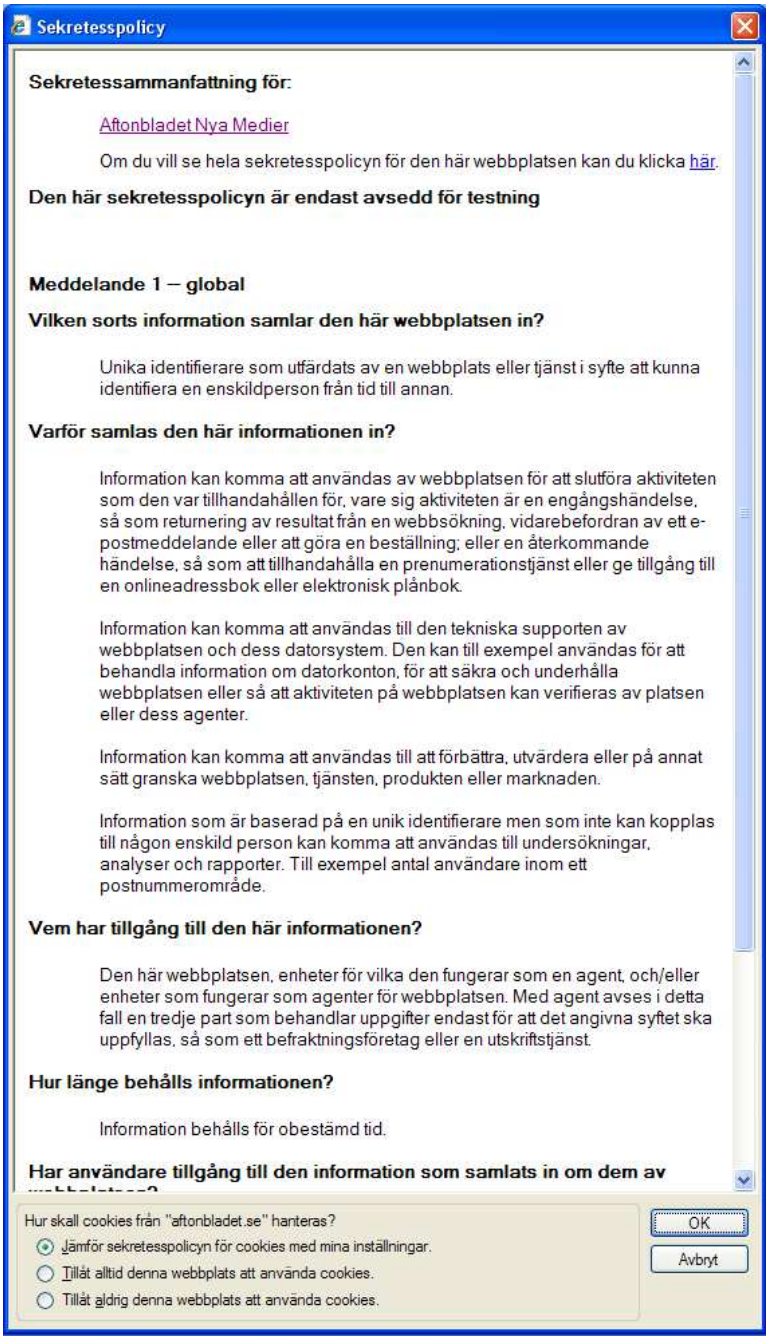
Of all security and privacy functions one function was not reachable from the "menu" ("Internet alternatives"). The option "View websites secrecy policy" is only reachable from the "View" menu (picture below).

The UI below is activated and here the user is able to look at a summary of visited websites secrecy policy.



Unfortunately no policy summary was found at many websites (e.g. the Swedish bank SEB and book site "Adlibris". Thus the user is referred to contact the website themselves to read their policy.

Below is an example of a policy summary found (Aftonbladet, 2008).



UI "Policy summary" found at www.aftonbladet.se (News site)

Very long text – for a summary…

10.9.3 Summary of usability problems

When walking through Internet Explorer the following potential usability problems were found:

Too long "summary"

In Internet Explorer the user can request a "Policy summary", where the website states e.g. how they handle personal data. But this "summary" was too long (too much text – to be a summary). This may depend on that the original policy was extensive. Maybe a point list of links to more extensive information would be a solution?

Different meaning of colour

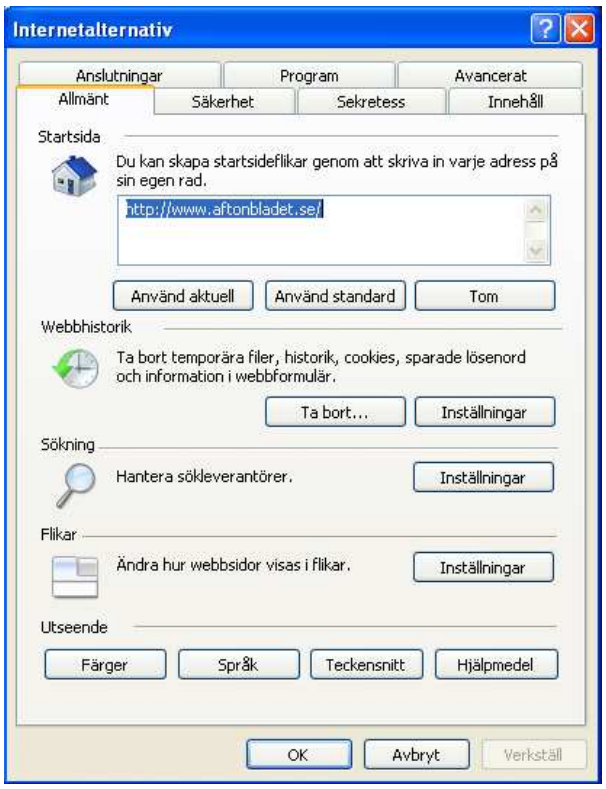
Different from other inspected IMS Internet Explorer used yellow colour in address field to indicate "authority cannot be verified". All other browser IMS used yellow colour to indicate "good" privacy.

Too small lock icon

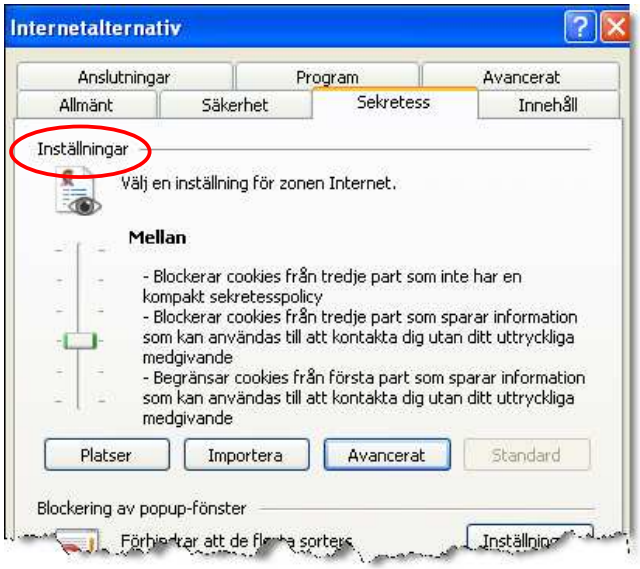
The lock icon (placed in address bar indicating a safe internet connection) is too small and placed in an area not in focus of user's eyes, thus it can be hard to see it. In other browsers (e.g. Opera and Firefox) the lock icon is **combined** with making the address field yellow, which improves usability and is helpful to make users observe a safe connection.

Unclear naming of functions

A general feeling after using Internet Explorer is that it has failed in **naming** different functions. The labelling (naming) should help the user to understand content before clicking on it, but several examples have shown that it doesn't. E.g. the UI below is called "Internet alternatives" (translation of the Swedish word "Internetalternativ"). But since it is some kind of menu for all user settings in Internet Explorer I would rather call it "**Menu**".



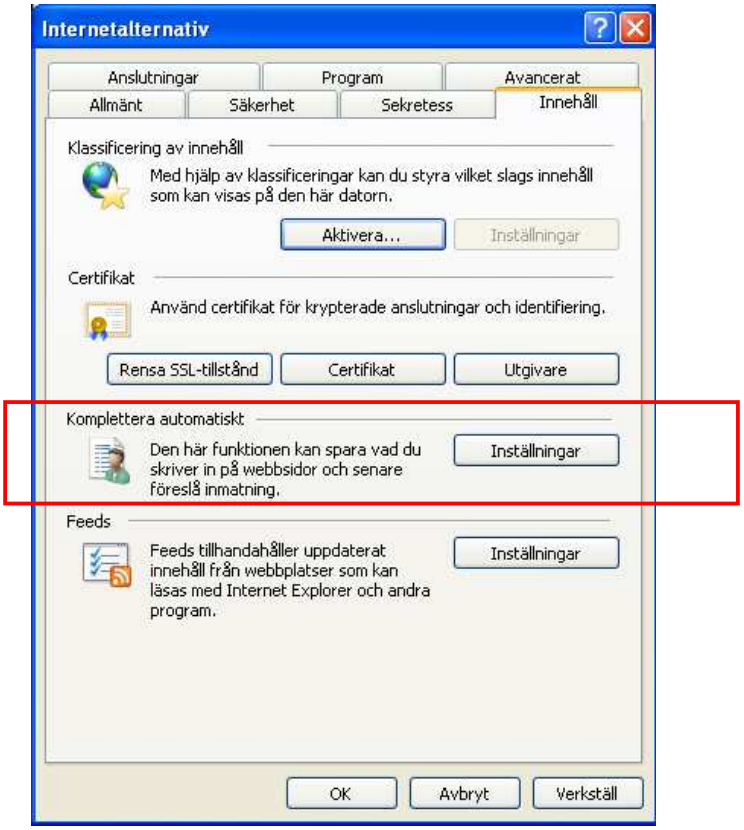
Another example is from the tab "Secrecy" in picture below. The label in upper part of tab tells the user that this UI handles "Settings" (Swedich "Inställningar").



But the user has to read the text to understand what kind of settings they mean. A better naming would be "**Cookies**" (or "Cookie settings") since it is cookie settings displayed in UI.

Another problem in the same UI (above) is the naming of the **three buttons** (named "Places", "Import" and "Advanced").

The **worse example** of unclear names (below).



Poor labeling (naming)

”Komplettera automatiskt”

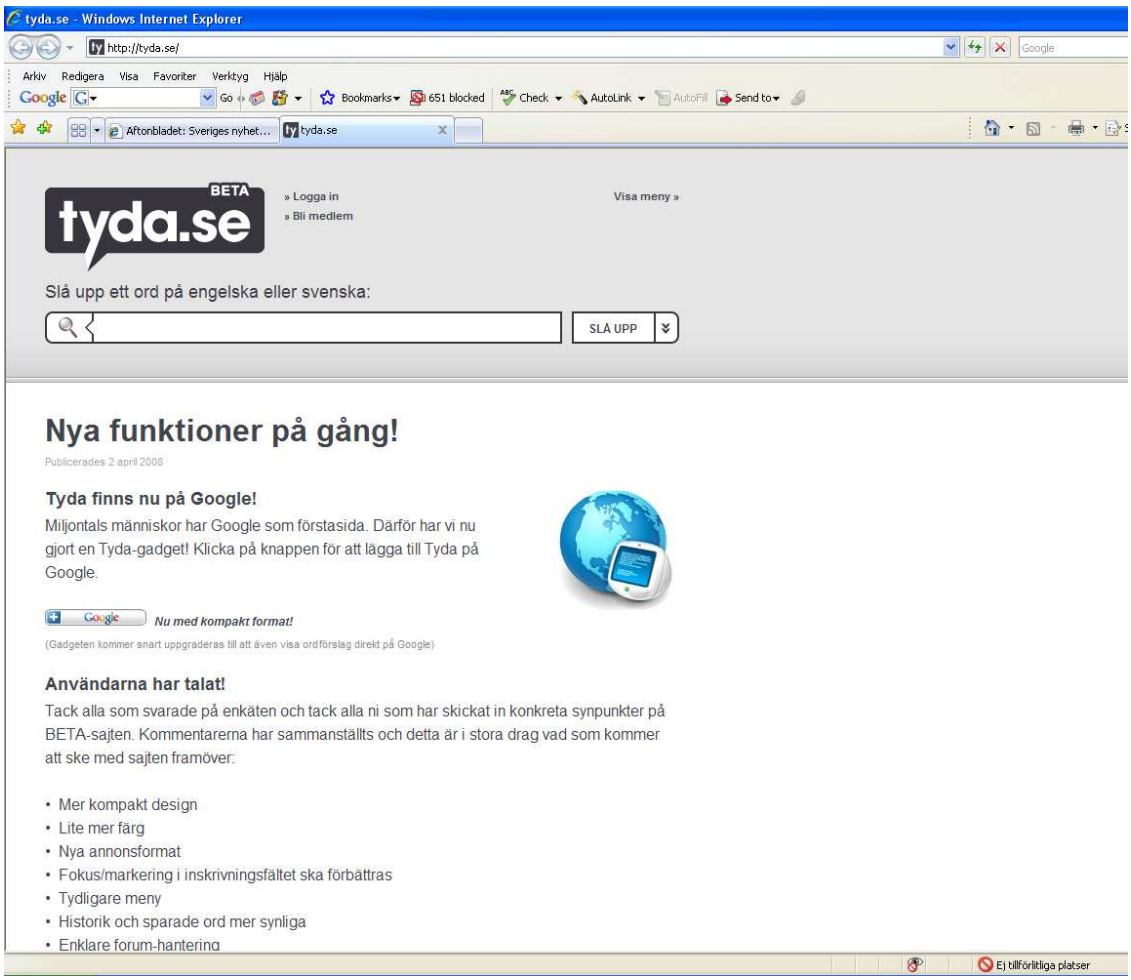
(English: ”Automatically Complement”)

In this UI the user can make Internet Explorer to remember password and other login data for later use. This is done by clicking on ”Inställningar” (Eng: ”Settings”).

A better labeling (naming) could be ”Password settings” or ”Login settings”.

Too small icon for warnings

A very, very small icon was displayed in the lower right corner of the browser window as an indication of a warning (in picture below). This icon should warn the user that the visited website was earlier put on a “black list” of unreliable websites. To put a warning at this place in UI is a threat to both usability and security.



Hard to find functions

In Internet Explorer the user is allowed to make settings about e.g. how data should be handled. But unfortunately it is too hard to find were to make these settings. And when the user finds them it is hard to remember were they were.

10.9.4 Usability criteria

In this section some reflections are made about Internet Explorer according to the usability criteria selected for this study.

10.9.4.1 Comprehensibility

Internet Explorer is quiet easy to use and also quiet comprehensible. Microsoft website was informative and explained a lot of functionality and features. But some of the stated functions (taken from Microsoft website) were hard to recognize during evaluation.

10.9.4.2 Error prevention

Internet Explorer offered a lot of warnings, if the user wanted these. The help files were also informative offering error preventive information. But when using the product some things were harmful for users. E.g. the warning icon was too small and placed in "non-focus area" which is not error preventive.

10.9.4.3 Task suitability

Internet Explorer was not hindering users in there primary purpose and therefore quiet suitable for its task.

10.9.4.4 Controllability

If the user wants to (and is familiar with all available features) they can be in control of some settings (functions). But the user has no control over their revealed data…

10.9.4.5 Likeability

The overall impression of this IMS is good. But it is important to know that Internet Explorer is the ordinary[8] browser used every day by the usability inspector, thus this may lead to easy of use and too positive reflections than other inspected IMS.

[8] "Ordinary" – I use Internet Explorer as my default browser (every day use). But i have never used the features evaluated in this study (from filling etc.).

10.10 CardSpace

The history functions of CardSpace was evaluated by Nina Rönntorp in September 2007 for her BA thesis in Information Systems at Karlstad university, Sweden. Parts of her description have been used below.

10.10.1 Short description

CardSpace is a product from Microsoft Corporation. The goal with CardSpace for the future is: "to let people on any machine, running any operating system, use digital identities as easily, as effectively, and as securely as they today use their identities in the physical world" (Chappell, 2006).

According to Microsoft website CardSpace is: "a new technology […] that simplifies and improves the safety of accessing resources and sharing personal information on the Internet […]. Windows CardSpace helps alleviate the problems of traditional online security mechanisms by reducing reliance on user names and passwords and instead using a separate desktop and cryptographically strong claims-based authentication. By helping users better manage their personal information and control how it is released and to whom, Windows CardSpace can facilitate more secure online experiences such as online shopping, banking, and bill payment" (Microsoft, 2008b).

10.10.2 Available functions

The CardSpace system has many different functions. Microsoft website mentions these:

- Simplify Online Authentication.
- Increase Web Site Security.
- Help Protect Against Digital Identity Theft.

According to Chappell (2006) four aspects of what CardSpace provides stand out as most important:

- **Support for any digital identity system**
- **Consistent user control of digital identity**
- **Replacement of password-based Web login**
- **Improved user confidence in the identity of remote applications**

In this study **only the history function** is evaluated.

CardSpace offers the user to see the history of each information card. To see this, the user chooses a card, and then chooses "view card history". A list is shown, that shows which sites the card has been sent to and the date of the latest transaction to each site. According to Windows CardSpace Help, the help function in CardSpace, the information that is stored when the user sends a card is site information about the site to which the user sends a card, time and data when the user sends it and the type of data that was sent to the site but not the data itself, e.g., "first name" is stored but not "John".

The purpose of storing these details is for CardSpace to be able to alert the user when a new site is requesting a card, the identity information or privacy statement of a site to which the user has sent a card has changed or when a site to which the user has sent a card is requesting new data.

The card history can be deleted by the user.

10.10.3 Summary of usability problems

Since CardSpace history functions evaluation was made before the preliminary analysis of IMSs, no walk-through of CardSpace was made to assess usability problems.

10.10.4 Usability criteria

In this section some reflections are made about CardSpace according to the usability criteria selected for this study.

10.10.4.1 Comprehensibility

The card metaphor seems granted to be understood by large user groups.

10.10.4.2 Error prevention

Specific notifications such as that the user has not visited a site before helps preventing users from sending information to fishing sites.

10.10.4.3 Task suitability

This is probably the greatest advantage of CardSpace but presumes web services are using the same protocols to spark the CardSpace into action. Ideally, user should not have to fill in web forms but select a card instead. This way of disclosing data takes the user away from the main web page, but on the other hand, such form filling could only be a step to reach the main task which is logging in, registration, or purchasing.

10.10.4.4 Controllability

CardSpace nicely presents its functions, but as the usability study on history functions revealed (see section 6.3) it was less easy to scan the history. It is also hard to know what information of a card was included in different disclosures as it is possible to suppress or add information when using a card for sending data.

10.10.4.5 Likeability

The graphic design is nice but likeability depends on if web services will use the protocol.

10.11 PRIME History function

PRIME History function (a mockup made during the development of IPV2 in 2006) was evaluated by Nina Rönntorp in September 2007.

PRIME (Privacy and Identity Management for Europe) is a European **research project** with participants from all over Europe, e.g. IBM in Zurich, Technische Universität Dresden in Germany and Karlstad University in Sweden. The project owner is IBM Belgium.

10.11.1 Short description

The *vision* of PRIME is user-controlled identity management systems, in which the user has control of personal information and negotiates its disclosure in return for access to a service (PRIME (2), 2007).

According to Nina Rönntorp (2008) PRIME has developed a system for identity management with key components as **access control**, advanced **anonymous credential** systems, and **automated reasoning**. "There are three parties involved in this system. The first party is the *user*, which has certificates, data and policies regarding his/her data. The user engages in transactions with service providers, in which he/she can choose to release data. The second party involved is the *service provider*, which offers services and resources to customers through transactions. *Certifying authority*, the third party involved, is a special type of service provider that issues certificates, i.e. digitally signed statements. When issuing a certificate, the certifying authority vouches for the truthfulness of the statement (Camenisch et al., 2005)."

10.11.2 Available functions

The PRIME prototype consists of many different functions (some stated above), but in this study **only the "history function" is evaluated**. It is also important to know that PRIME is not a working program yet (only a prototype).

The history function within PRIME-prototype is called "Data Track". In Data Track every transaction is saved e.g.

- Information about data sent
- Date and time of transaction
- Website to which data is sent
- Preset (user role) used during transaction (settings of personal data)

In the PRIME-prototype a **search function** is available to search in Data Track for earlier transactions.

10.11.3 Summary of usability problems

The evaluation of the history function in CardSpace included also an evaluation of an html mockup of Data Track. Therefore, the present analysis does not include any fault finding as concerns Data Track (it should be noted that the mockup is not identical with the implemented Data Track of the final PRIME prototype called IPV3).

10.11.4 Usability criteria

In this section some reflections are made about PRIME according to the usability criteria selected for this study. These reflections are based on the Rönntorp study (Rönntorp, 2008)

10.11.4.1 Comprehensibility

According to Rönntorp's study PRIME history function was easy to use and most users liked it. Thus the users also seemed to think the history function was quite comprehensible.

10.11.4.2 Error prevention

No explicit information found about this in Rönntorp's report.

10.11.4.3 Task suitability

According to the Rönntorp study is seems like the users thought PRIME history function was suitable for its task.

10.11.4.4 Controllability

Since PRIME history function saves a lot of information about a data transaction the user has control over what data has been sent and who has received it. The user is also in control over the history function since it has a search function with in the history function.

10.11.4.5 Likeability

The graphic design varies quite a lot between different PRIME mockups of the Data Track. It is thus hard to make any guesses as to the immediate impression people will have, but the functionality, with one search function covering all data releases regardless of what data disclosure preferences have been used at different times was clearly liked by the test subjects in the Rönntorp study (see section 6.3). As for Microsoft's CardSpace, the success of the PRIME overall concept is dependant on web services accepting a protocol for immediate data requests machine-to-machine, although within the PRIME project also a web form-filling function has been tried, which could pave the way for using PRIME and Data Track by structuring data releases in a format possible to store in Data Track.

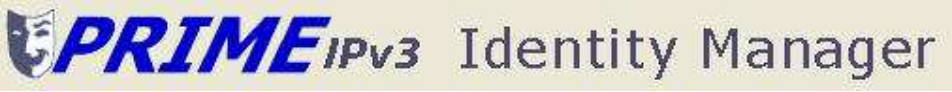
10.12 PRIME IPV3

PRIME IPV3 (version 3) was installed and evaluated the 20-21st of April 2008. PRIME (Privacy and Identity Management for Europe) is a European **research project** with participants from all over Europe, e.g. IBM in Zurich, Technische Universität Dresden in Germany and Karlstad University in Sweden. The project owner is IBM Belgium.

It should be noted that the usability inspector had limited prior knowledge about the PRIME-prototype and PRIME-technical concepts. Unfortunately the last PRIME-report on the prototype deliverable was not read prior this inspection, which could have excluded some of the reported problems.

10.12.1 Short description

According to the website (PRIME, 2008) "PRIME aims to develop a working prototype of a privacy-enhancing Identity Management System. To foster market adoption, novel solutions for managing identities will be demonstrated in challenging real-world scenarios, e.g., from Internet Communication, Airline and Airport Passenger Processes, Location-Based Services and Collaborative e-Learning."



PRIME IPV3 has a "PRIME console", a kind of menu where all available functions are gathered. Some of the functions are "Personal Data" (data that the user has entered into the system, data that the PRIME system should handle), "Data Track" (a history function), "PRIME settings" (where the user can customize the PRIME system) etc.

10.12.2 Available functions

Within the usability test of PRIME IPV3 only some of the PRIME functions were evaluated.

- Role and (partial) identity management
- Authentication management
- Form filling
- Context monitoring

10.12.3 Summary of usability problems

When walking through PRIME IPV3 some potential usability problems were found, presented below.

First some comments about things connected to the functionality of the prototype (IPV3), not really a usability problem (well, perhaps it is too…).

- **"Send Personal Data" pops up too fast**
  When the user enters a scenario (e.g. Buy Harry) the "Send Personal Data" window pops up too fast, thus the user is not able to read the text at the PRIME playground. Since this ***information is important to be able to understand*** why the Send data" pops up it is a hinder.

- **Links not working**
  When inspecting IPV3 the inspector noticed that many links were not working. This may confuse the participants during a user test!

10.12.3.1 Usability problems in "Send Personal Data" – Login tab

Most of the noticed usability problem is in the "Send Personal Data" UI. One impression is that this UI is straggly. It is hard to get a quick overview of what to do, and hard to see what information is important or not. But first problems connected to the "Login" tab is presented.

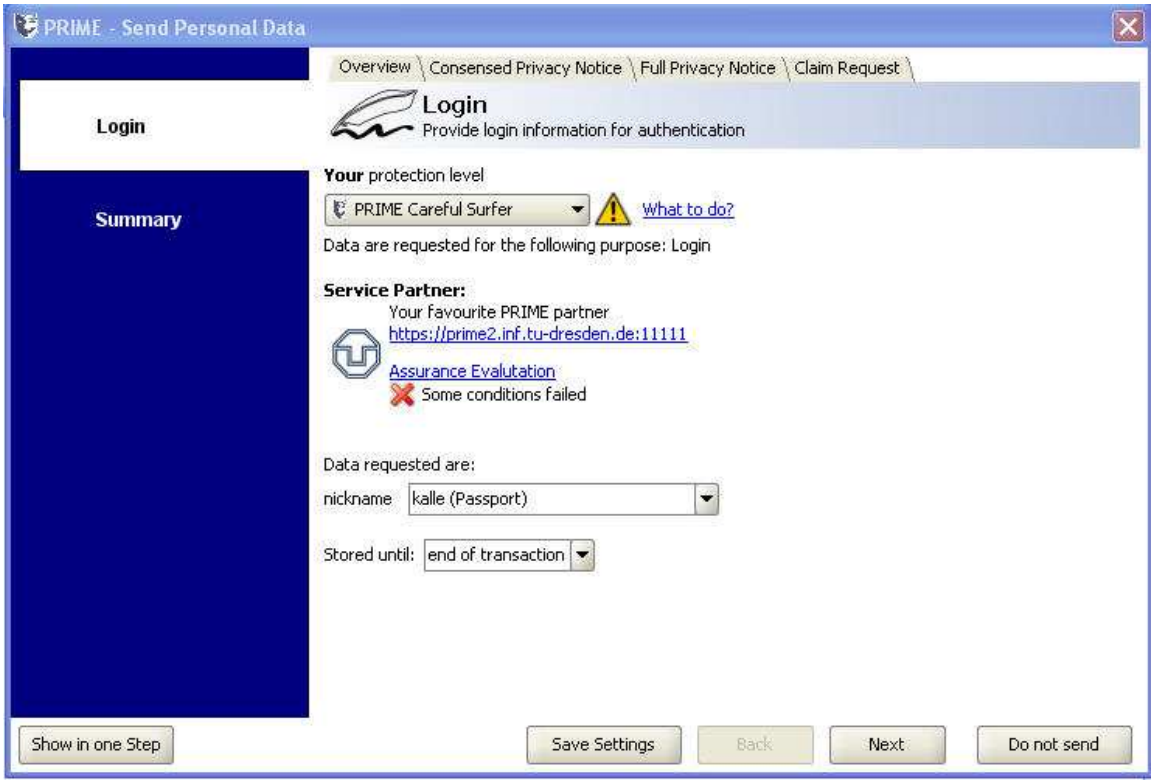


Figure: Send Personal Data – "Login" tab

Login icon

First of all – the login icon is not good. It is hard to see what it represents and the icon is taking too much focus from the user. Either the icon need to be changed or taken away. And if it is kept, it should appear twice (also in the left column).

Important text not in focus

E.g. the text "Provide login information for authentication" should be in focus, since it is telling the user *what to do*. But the text is not in focus – the icon is…

Icons not clickable

None of the icons in this UI is clickable, and no text is given when hovering the mouse over them. At least the warning icon should have been clickable.

Links not clickable

Unfortunately only one link ("Assurance evaluation") is working (clickable) in the prototype. This may hinder users and confuse them during a usability test.

Hard to understand level of protection

First of all the names (labels) of available protection levels are hard to understand. E.g. what does "unworried" or "careful" means? The answer to this will probably differ depending on whom you ask. Another problem is that nothing in the labels (name of the protection level) indicates if data will be "remembered" or not. And as seen in the four pictures above the hint text (in blue box) differ between different protection levels.

10.12.3.2 Usability problems in "Send Personal Data" – Summary tab

Also the tab "Summary" has some usability problems.

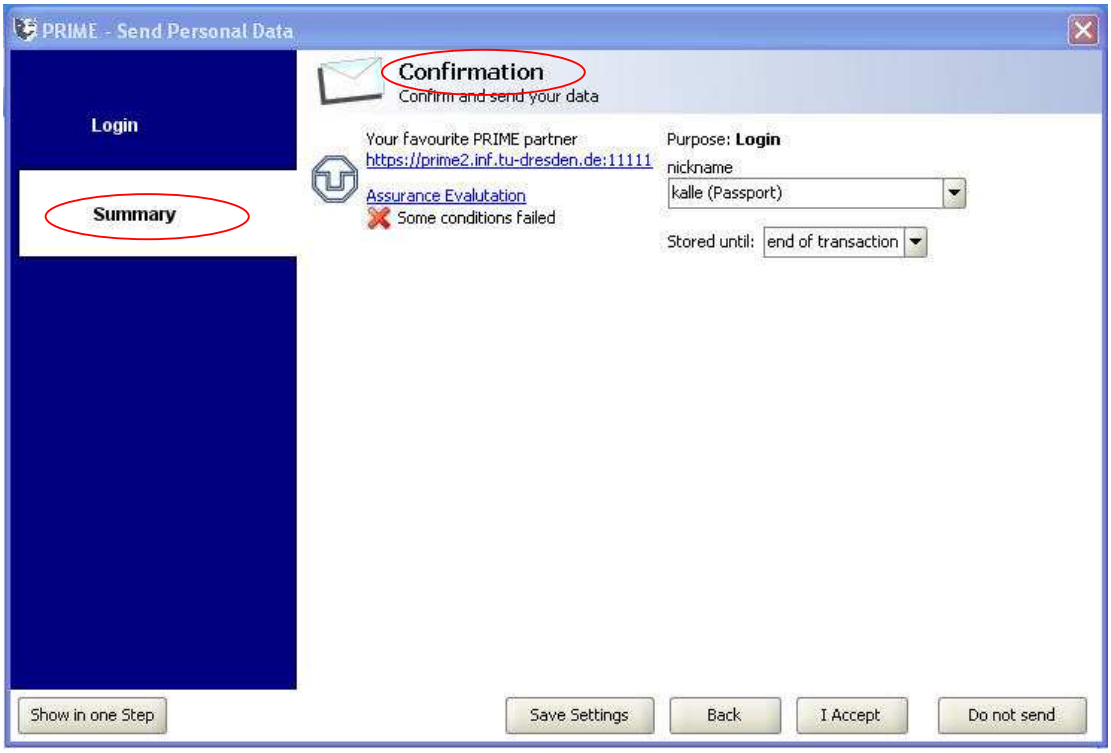


Figure: Send Personal Data – "Summary" tab

Inconsistent names (labels)

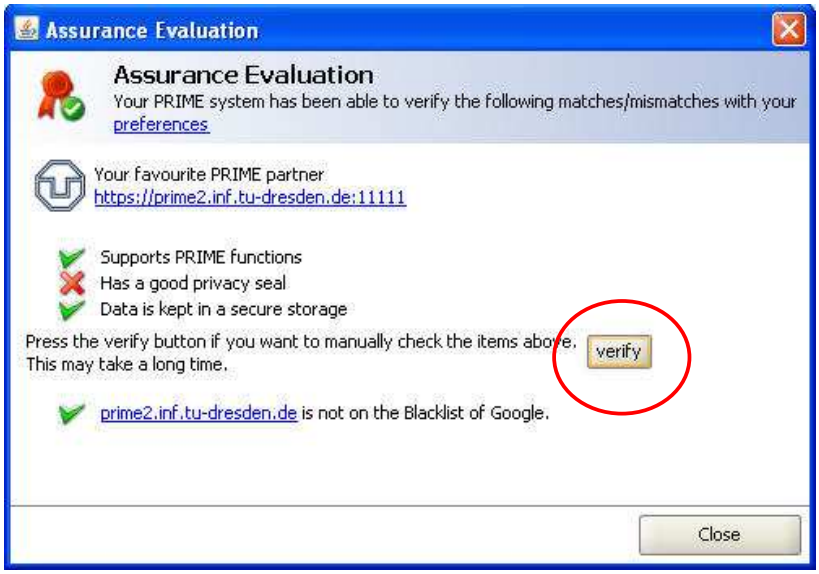
The labels (marked with circles above) have different names. E.g. the label in upper UI is called "Confirmation" while the label in the left menu is named "Summary". To not confuse the user it is important to be coherent and use the same names/labels.

Inconsistent structure of objects

When the user enters the "summary" tab the information is rearranged (moved). Why? This is only confusing the user. The information given in the "Summary" tab should be arranged in a similar way as the other tabs (e.g. Login), thus make it easier for users to recognize it.

10.12.3.3 Usability problems in "Assurance Evaluation"

The only working link in "Send Personal Data" is "Assurance Evaluation", triggering this UI:



Unfortunately none of the links in this UI worked either. The button "Verify" worked and when clicking on this button I trigged the verification process indicated by flickering text in the UI.

Figure: Result of "Verification"

Contradictive messages to the user

The result of the verification is not corresponding to the displayed icons. E.g. the last row (marked in picture above) "Data is kept in a secure…" has a green check icon (indicating OK) on the left hand side, but a text on the right hand side tells the user that the verification failed…

Also the example with the text "Has a good privacy seal" is contradictive. The icon (a red cross) is indicating some kind of warning or mismatch, while the result of verification is telling the user that this is "not claimed". I think a better solution here would be to skip the icon, and just tell the user in text that this is "not claimed". The icon (red cross) should only be used when something is mismatched or wrong.

No explanation of controls

As far as understood the "Assurance Evaluation" is controlling (checking) different things. But it is hard to understand what controls (checks) are performed…Are there one control or many (perhaps as many as **five** controls)? And there is no explanation of what these results will mean to the user. E.g. what does it mean to have "a good privacy seal"?

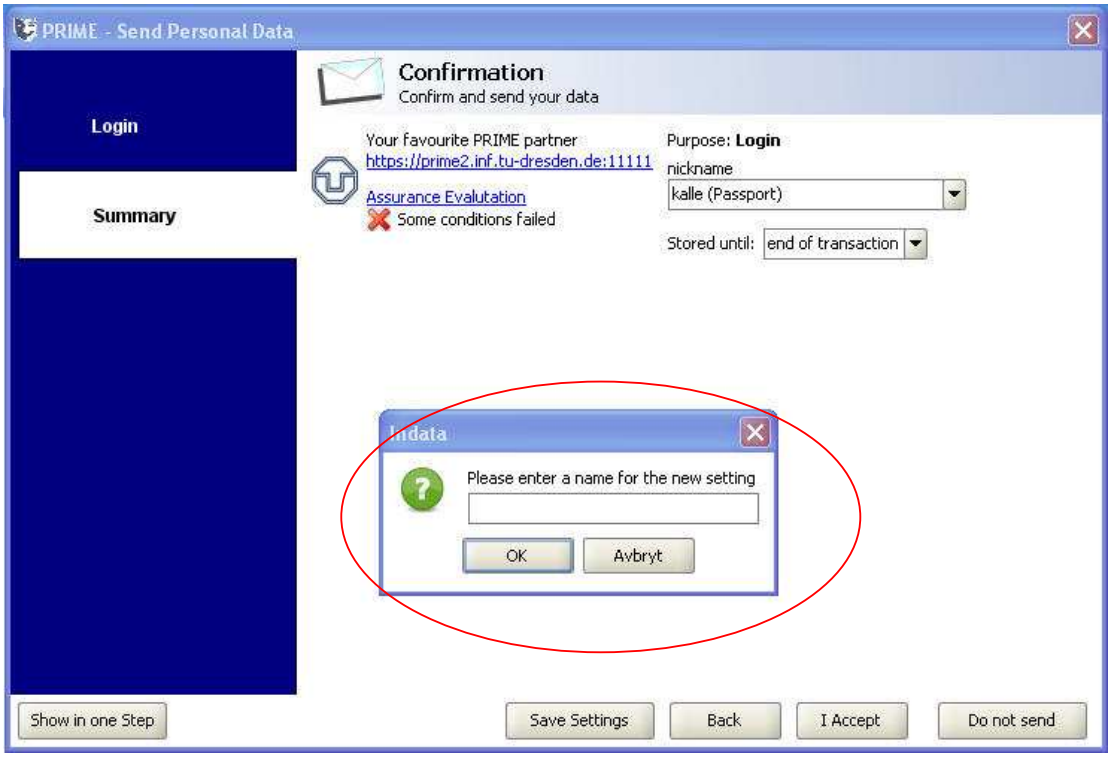
10.12.3.4 Usability problems connected to buttons

When evaluating the different buttons in lower UI of "Send Personal Data" I found the following usability problems:

Hard to understand what "settings" is saved

When clicking the **button** "Save Settings" it was not clear what was going to happen. *What settings* are saved? And for what purpose is these settings saved? This is really not clear.

Is this a way to create a new "Protection level" (former "Role" or "PrivPref"?) The pop-up window (below) that appears is not telling me anything more. The user is only told to enter a name for the new setting.



After entering a name and clicking "OK" a nonsense UI is displayed (below). Why? The user do not get any feedback of what have happened.

10.12.3.5 Usability problems in "Buy Harry" scenario

In the "Buy Harry" scenario the user should prove his/hers age. The PRIME Playground (the faked website) offers three options to proof the age of the user – "Credit card number", "Post ident ID" or "Identity card".

UI hidden too fast

The "Playground" hidden too fast (by the pop-up window "Send Personal Data") – the user does not attain to read the text. This may affect the users comprehension of what is happening when using the product.

Bad example of identification

Why is credit card number an example of proof of *age*? If this option shall be present it should be put as the last option (thus the users probably will not read it…)

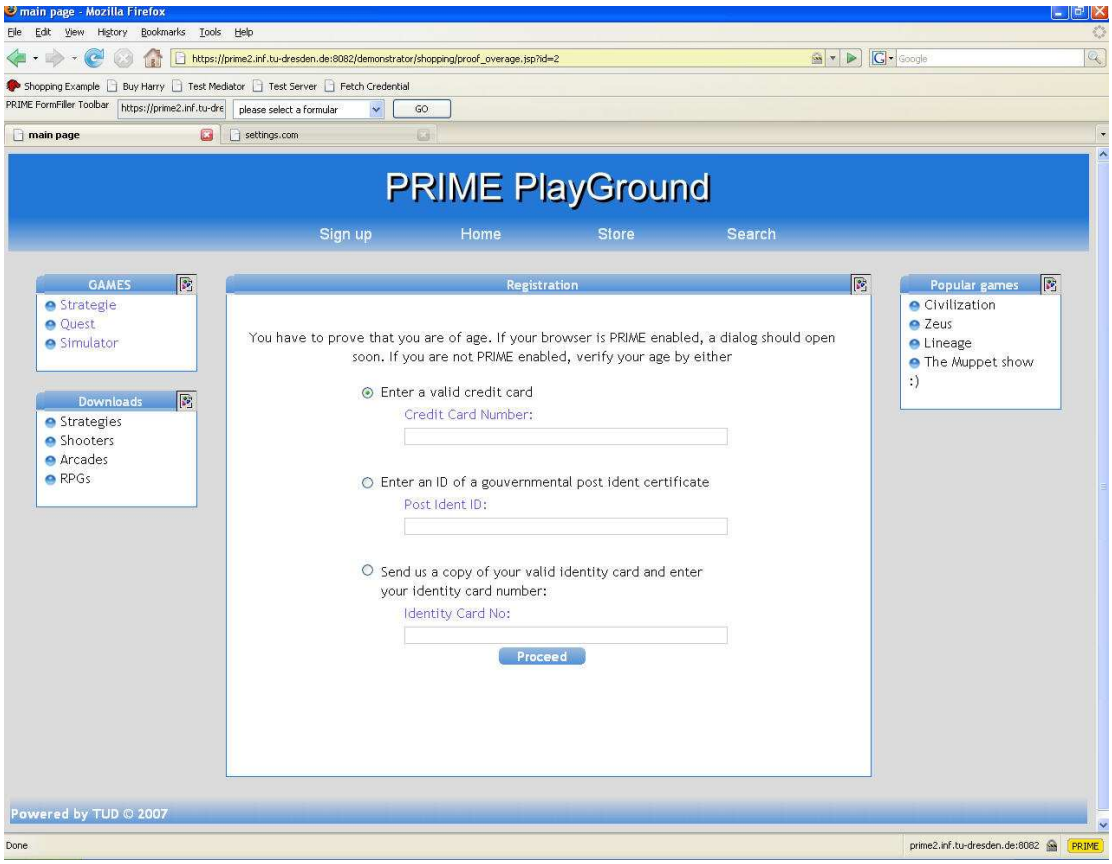


Figure: PRIME "playground" with options to verify age

10.12.4 Usability criteria

In this section some reflections are made about PRIME IPV3 according to the usability criteria selected for this study.

10.12.4.1 Comprehensibility

Some things in IPV3 were too hard to understand to be able to call it comprehensible. E.g. concepts such as "Assurance evaluation" were not explained enough to the user. Other examples of things too hard to understand are the different levels of protection's e.g. "unworried customer" can mean different things to different people.

10.12.4.2 Error prevention

First of all IPV3 only warns the users, not hinders them. And sometimes there also are conflicting warnings displayed to the user. Thus I don't think IPV3 is error preventive enough. The warnings need to be more explicit and informative.

10.12.4.3 Task suitability

Yes, I think IPV3 is suitable for primary task, since the user have the option to "confirm and send" data in *one step action*.

10.12.4.4 Controllability

Due to the fact that the user can decide if personal data should be sent or not, and under which conditions these data will be sent, the user is partly in "control". But if the user *feels* in control when using IPV3 is another question? Thus I would say that this criterion is only partially fulfilled.

10.12.4.5 Likeability

I think most of the functions in IPV3 are accurately thought through, but poorly visualized. And unfortunately users mostly see the visual part.