

Flexible Services

LUCRE - Local and User-CREated services D1.6: User Studies and Evaluation of LUCRE Pilot Services

Research area:	Everyday Services
Project/WP Title:	LUCRE - Local and User-CREated services / WP1
Document Type:	P (Public)

Document Title:	D1.6: User Studies and Evaluation of LUCRE Pilot Services
Document Identifier:	FS_LUCRE_D16
Delivery date:	31.5.2010, due project's extension period delayed to 30.09.2010
Authors:	Sanna Finnberg, Juha Leino, Kari-Jouko Raihä (UTA / TAUCHI) Katarina Segerståhl, Petri Saarikko, Suvi Silfverberg, Ayush Bajaj (HIIT / Uix)
Status/Issue:	version 1.0
Date Last Change:	27.08.2010

Table of contents

Tables and Figures	3
Executive summary.....	4
1. Introduction.....	5
2. LUCRE Pilot Services.....	7
2.1. LUCRE Media Pilot	7
2.1.1. Concept	7
2.1.2. News reading and recommending survey results.....	9
2.1.3. Haavi-Alma Toolbar usability test results	12
2.1.4. Haavi-Alma Toolbar field study	13
2.1.5. Lessons learned.....	14
2.2. Wedding Organizer Service	15
2.2.1. Concept	15
2.2.2. To-do widget usability test results	18
2.2.3. Lessons learned.....	20
3. Work summary of WP1	21
3.1. Studies and documents	21
3.2. Design work.....	27
4. Conclusions.....	29
4.1. LUCRE project tasks and goals for WP1	29
4.3. Look in the future	31
Appendix 1:	32
5. News reading and recommending survey results.....	32
Appendix 2:	47
6. Haavi-Alma Toolbar usability problems and bugs reported in Redmine.....	47
Appendix 3:	49
7. Haavi-Alma Toolbar usability test report	49
Appendix 4:	63
8. To-do widget V1 usability test report	63

Tables and Figures

List of tables

1. <i>Publications of LUCRE WP1 - first two years</i>	21
2. <i>Documents by LUCRE WP1—first two years [table on three pages]</i>	22
3. <i>Studies conducted by LUCRE WP1—first two years [table on two pages]</i>	25
4. <i>Design work for Wedding Organizer Service by LUCRE WP1</i>	27
5. <i>Design work for LMP by LUCRE WP1</i>	28

List of figures

1. <i>Haavi-Alma Toolbar integrated on Aamulehti news site.</i>	8
2. <i>Shared haavi (left), invitation showing on the toolbar (right).</i>	9
3. <i>“Butterfly net” icon shows up when hovering over a news item title.</i>	9
4. <i>Roadmap for widget development for the widget ecosystem.</i>	17
5. <i>Tested version of the To-do widget interface.</i>	18

Executive summary

The purpose of this document is to give last results of user studies related to LUCRE (Local and User-CREated services) project pilot services not reported in previous deliverables (D1.2, D2.1, and D1.4), to summarize the two years of user-centered design (UCD) activities done by Work Package 1 (WP1), and to elaborate lessons learned, both problems and advances.

The document includes brief introduction to the theme. Then the focus is turned to work done with the pilot services. To keep the document short and as coherent as possible also for external readers, only main finding of the user studies are presented. The whole documents of user studies (except from the latest field study) can be found as appendixes. The pilot services are gone through from latest to previous: first LUCRE Media Pilot (LMP) and then Wedding Organizer Service (pilot called EVENT). [In parallel deliverable 2.4: *Report on LUCRE Pilot Services* these are referred as *Social media recommendation pilot* and *Shared todo-list pilot*.] The user studies inside the pilot services are in chronological order: In LUCRE Media Pilot the News reading and recommending online survey, usability testing and field study of the Haavi-Alma Toolbar and in Wedding Organizer Service pilot the To-do widget usability testing.

After this the focus is broadened to summarize the two year work of WP1 in form of documents and studies together with design work presented chronologically in table format. Also academic publications based on WP1 work are listed, even though some still in the writing phase.

In the end the achievements are compared to LUCRE's tasks and goals constituted in the beginning of the project. WP1's goal was to come up with successful services for managing daily activities by studying current practices and reviewing literature to identify promising service areas. In addition, WP1 was to take part in designing pilot services with user-centered design (UCD) methods. When the main project partner dropped out of the project after first year LUCRE had to undergo a shift in focus to take the remaining industrial partners' interests better into consideration. Overall, WP1 managed to rather comprehensively go through the task list for both pilot services. Only the depth of some work suffered from the shortness of time caused by this.

The major findings and lessons learned from the UCD work done in the LUCRE project are not general theories or frameworks but findings rather focused on the field of LUCRE pilot services. So far the work done within WP1 has resulted two academic publications. The research findings on shared lists, cross-medial services, and highly generic and easily integrated widgets can be useful for future project and studies in the same field as these themes seem to be of interest for many these days. Also future of these kinds of services is shortly discussed in the end of this document.

1. Introduction

The goal of LUCRE (Local and User-CREated services) project, part of the larger Flexible Services program, has been to gain understanding of the potential of user-created services by developing and studying service composition tools and technologies. This deliverable concentrates on the work of Work Package 1 (WP1), as the work of Work Package 2 (WP2) will be discussed in their deliverable *D2.4 Report on LUCRE Pilot Services*. However, Work Packages in LUCRE have co-operated closely and there have not been clearly delineated borders between them: The borders have been porous to allow smooth co-operation. The main distinction of the Work Packages has been following: WP1 has concentrated on users' aspect of services and WP2 on technological advances.

This deliverable also closes the first two years of LUCRE project, and thus, in addition to reporting on the work done since the last deliverable *D1.4 Current practices and resulting services*¹ (Section 2) it summarizes briefly the work and results of WP1 during the two years (Section 3). Since LUCRE will not continue as a separate project in reshuffled Flexible Services program, this deliverable also concludes LUCRE for WP1 (Section 4). The main reason for LUCRE not continuing as a project within Flexible Services program is the lack of industry partners. The ever-shifting industry-partner situation has been the major difficulty in LUCRE and has forced the project to re-direct its focus over the course of the project.

WP1 has focused its efforts on studying current user practices to base its development work on actual use, service designing work, and pilot service testing to develop services that users can successfully use to facilitate everyday activities. Initially, the efforts were directed towards creating a service concept for a widget ecosystem where users could create services for themselves by combining widgets that could share information. The goal was to pilot a Wedding Organizer Service (pilot called EVENT)² that would be made by combining different widgets. After the initial work and user studies on wedding organizing and PIM, WP1 produced wireframe designs for three widgets, to-do widget, reminder widget, and map widget, that could be combined together to produce services. The graphical user interface for to-do widget was designed, the widget was implemented in the widget ecosystem built by WP2, and first round of usability testing was conducted.

However, as industry partner situation changed, LUCRE was forced to re-focus on media-related services. The current version To-do widget was adopted by OtaSizzle³ Ubiquitous Social Media for Urban Communities and WP1 focused its efforts on designing social recommending and sharing application. LUCRE Media Pilot (LMP)⁴ was designed to consist of three components: Browser toolbar, Mobile application (for Android platform), and Facebook⁵ application (for integrating with the Browser toolbar). Also, WP1 conducted an online survey on current news reading and

¹ Delivered on November 30, 2009

² In parallel deliverable 2.4: Report on LUCRE Pilot Services this is referred as Shared todo-list pilot.

³ <http://mide.tkk.fi/OtaSizzle>

⁴ In parallel deliverable 2.4: Report on LUCRE Pilot Services this is referred as Social media recommendation pilot.

⁵ <http://www.facebook.com>

recommending practices. After designing the LMP applications, WP1 also conducted usability testing and field study on the Browser toolbar.

As WP1 focused its efforts on LMP in the fall of 2009, thus already before the previous deliverable, this deliverable discusses mainly the work done on it in addition to summarizing the work of WP1 as a whole. No other work will be discussed in detail except for the usability test report on the To-do widget, which will be officially published here.

We first look at the LUCRE pilot services (Section 2): the LUCRE Media Pilot we designed, discuss background work for designing, a survey study (n=147) of news reading and recommending practices in Finland. Then we look at the user testing process, first presenting the main findings of the usability testing (n=6+1) and finally the field study (n=10). Then we briefly introduce the Wedding Organizer Service pilot (EVENT) and main findings from user testing of the only implemented widget for the widget ecosystem, the To-do widget (n=10+1). The user study reports except from the field study can be found as appendixes. That concludes our discussion of our work since the previous deliverable and we turn to summarizing the work done by WP1 during the first two years of LUCRE (Section 3). As longer reports and documentation are available elsewhere, we only summarize the work without going into great detail. In Conclusions (Section 4), we summarize the achievements from EVENT and LMP and peek at the future of such solutions based on our findings.

2. LUCRE Pilot Services

WP1 conducted user studies to two LUCRE pilot services: The Wedding Organizer Service (EVENT) and LUCRE Media Pilot (LPM). This Section outlines first LMP and then EVENT by concept, related user studies that have not been reported in earlier deliverables, and main lessons learned.

2.1. LUCRE Media Pilot

The Media Pilot (LPM) was the latest pilot of the LUCRE project. The first LMP related user study⁶, and the roadmap are reported in *D1.4A: Current practices and resulting services*.

2.1.1. Concept

One of the major challenges in the web environment is in finding and screening out relevant content. In order to respond to this challenge, service providers need to incorporate value-adding browsing mechanisms, such as recommendations, and other mechanisms to attract people in the first place and to make sure that they return to the service.

The LUCRE Media Pilot (LMP) is constructed around the domain of digital news and media. The pilot combines ideas from social navigation and crossmedial widgetting to improve the browsing experience and to focus traffic on particular news and media sites. Social navigation refers to the practice of navigating in an (online) environment based on the recommendations or actions of other users. In LMP, social navigation is based on social recommendations. Social recommendation may be very effective as users are typically drawn to the content that has been accepted, noted and "screened" by other users—preferably ones that they know personally—as opposed to the service providers or automatic recommendation systems doing the screening. Crossmedial widgetting refers to a widget-based approach to providing functionality that can be extended across media, such as services, platforms and technologies, e.g. Facebook, Twitter⁷, widget engines, mobile devices and public displays.

Social recommendations constitute one of the most prominent forms of social interaction in the online environment for supporting social navigation. For example, IRC (internet-relay chat) has a tradition of users pasting interesting links to channels for other users to see. Discussion forums are another popular place for sharing interesting links as recommendations for other users on what is worth checking out. In addition, a number of links travel through emails and other messages among more restricted user groups. The main value of these practices is in that they screen content for particular interest groups or individuals and thus make it more useful. The user perception of the value of a media service may be significantly increased if this kind of collaborative

⁶ news reading and recommending online survey

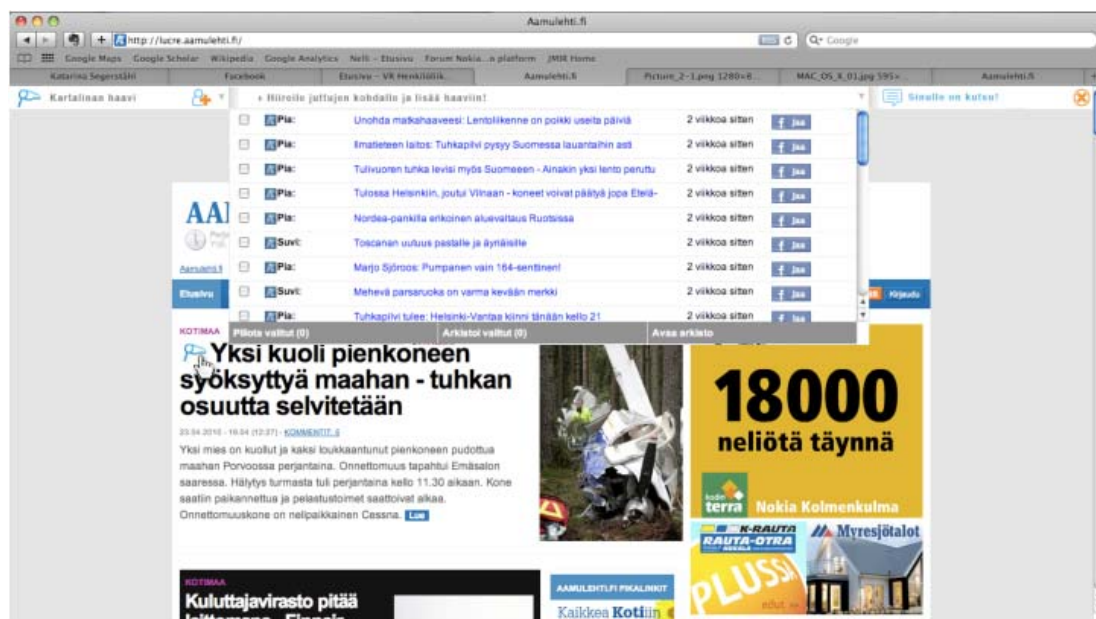
⁷ <http://www.twitter.com>

screening can be supported in a non-intrusive way as part of the normal browsing experience.

In LMP, the solution was first implemented on one news service, the Aamulehti website⁸, for piloting and proof of concept. This enabled the designers and implementers explore the basic technical and user experience aspects associated with the solution and its initial implementation. The initial solution is called Haavi-Alma.

Haavi-Alma consists of a toolbar that hosts socially constructed lists of (news) items. It was integrated in the Aamulehti news site for piloting. The main purposes of Haavi-Alma are to 1) provide users with a non-intrusive tool (as opposed to for example email) to browse and share interesting links and 2) explore the potential of micro-social service user experience (as opposed to for example Facebook expanding out of social comfort zones with increasing privacy concerns and open, macro-social solutions, such as public discussion forums with issues associated with anonymity).

The Figure 1 illustrates the implementation of the LMP widget in the form of a toolbar on the Aamulehti website (<http://lucre.aamulehti.fi>). The left section of the toolbar is for managing haavis (“butterfly nets”) that can be created for collecting items around specific themes, purposes or groups. The center part of the toolbar is allocated for presenting actual haavi content, i.e. the list that contains collaboratively collected items. The list can be opened and closed and its top row serves as a status bar, showing the item added last and who added it. The right section is for displaying and reacting to announcements, such as when someone has invited you to participate in their haavi.

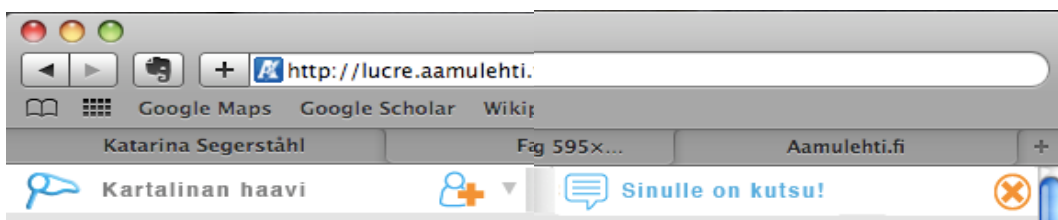


1. Haavi-Alma Toolbar integrated on Aamulehti news site.

⁸ <http://www.aamulehti.fi>

Haavis can be created and managed for collecting items around a specific theme or purpose. For example, to collect the latest news around the “Ash Break” to keep a family member who is traveling updated about what the Finnish media (in this case, Aamulehti) is publishing around the theme.

Haavis can be shared with one’s contacts (Figure 2). A shared haavi may be simultaneously viewed and updated by any or all of its users. Contacts can be invited via email or they can be searched among registered Haavi-Alma users. One can also log into Haavi-Alma with their Facebook user name and password. In this case, they will be able to browse through contacts associated with their Facebook account and invite them to share haavis.



2. Shared haavi (left), invitation showing on the toolbar (right).

News items can be added to haavis by simply clicking on the non-intrusive “butterfly net” icon that appears when hovering over news titles (Figure 3).



3. “Butterfly net” icon shows up when hovering over a news item title.

2.1.2. News reading and recommending survey results

The news reading and recommending study was carried out as an online questionnaire, and it gauged the current news reading and recommending practices to advise the designing of the social news reading toolbar, the axis of LUCRE media pilot (LMP). The whole report can be found as Appendix 1 but here the main findings are summarized.

The results showed that the most popular method of reading news was to access some news site directly, that is, using the service one knows and likes. However,

nearly half the participants (70 people or 48% of the 147 participants) typically also used recommendations from people to locate interesting news items. In fact, 22% received recommendations several times a day and 53% received them at least several times a week.

While recommender sites, such as Slashdot⁹, Digg¹⁰, and Reddit¹¹, aggregation services, such as Ampparit¹² and Google News¹³, and RSS feeds, were all used to locate news by some participants, their share pales in comparison to the recommendations from people.

The popular channels for giving and receiving recommendations were social media, such as Facebook and Twitter, email, IRC, and instant messaging. Surprising many also mentioned face-to-face conversations and even phone conversations and SMS. In addition, discussion forums still do play a small part in the equation, even if their importance is likely to have decreased from their past glory.

Recommending news appears to be part of social behavior in modern times. People like to stay in touch and maintain relationships, and recommending news that we expect the receiver to be interested in is a way not only to give the receiver something interesting to read but also to maintain relationship and to indicate that I know what you like. The fact that the recommendations are, at least occasionally, continuations of conversations and occasionally invitations to conversations are all part of humans being social beings.

Part of this is that news items are read based on who the person making the recommendation is. When there is a close relationship between the recommender and the receiver, the recommended item is likely to be read. Although the receiver reading the item is not so important to the sender, the receiver appears to feel compelled to read the item when the relationship is close.

On the other hand, the rise of social media has moved the recommendations to the realm of such social media sites as Facebook and Twitter. While Facebook ruled supreme among our participants, Twitter also plays a role in how news finds their readers. However, a recommendation over a social media site is not as compelling as a personal, direct recommendation, and other factors, such as the title of the article, come into play in deciding which items to read.

Thus, in designing ways to recommend news through such services, giving information about the distal item is important to help them make up their mind about reading it. Furthermore, being able to give such items personalized

⁹ <http://slashdot.org/>

¹⁰ <http://digg.com/>

¹¹ <http://www.reddit.com/>

¹² <http://www.ampparit.com/>

¹³ <http://news.google.com/>

feeling could also help. A news recommendation on a Facebook wall is not personal but one shared personally with one through Facebook can be a personal recommendation.

In making recommendations, being able to comment on the news item one was recommending was appreciated, as it made it possible to make one's stance in relation to the item clear and help receiver to see the important part. Responses from the receivers were not seen as essential but there was evidence of some sort of give and take taking place. For instance, lack of response had caused 2 participants quit recommending news some people. It appear that occasional response is appreciated, and thus enabling it is important, but it does not need to be a constant barrage of responses, and being forced to respond every time did not appeal to the participants.

In receiving recommendations, especially when not involving the social give and take, there was a clear preference to recommendations that took one's preferences into account. Flooding people with news recommendations that do not take their preferences into account is seen as spam and is likely to lead to reduced reading and acting (that is, going to read the news). Fewer but good quality recommendations help dealing with the endless number of news while too many non-personalized recommendations are not appreciated.

The only group of people the participants expressed interest in following for recommendations was that of experts in a given fields. If recommendation channels are built around people, then the people around whom they are built need to establish their specialization in that field.

In general, people do not recommend news that they have not at least glanced, which is in keeping with the goal of keeping up relationships—bad recommendations are unlikely to improve our standing in the eyes of others. All in all, people would like to recommend whole articles or individual pictures in articles. Being able to recommend parts was not in the most participants' wish list.

In closing, this discussion only went through the replies of the participants as a group, and it is likely that there are distinct groups of users within the participants that have different needs and goals. One system rarely can serve everybody, and thus needs to be designed for a group of users with certain shared characteristics. For instance, the users of social media could be analyzed in separation of other participants for developing applications to be used in the context of the social media sites. The values and ideas of people who do most of their recommending by email, for example, would not greatly contribute to designing an application for social media, and consequently, should not be considered at all or at least with reduced weight in relation to those of the social media users.

2.1.3. Haavi-Alma Toolbar usability test results

The usability testing of Haavi-Alma toolbar was conducted to confirm the usability of the application for the field study. This was done only one week before the field study began, so most crucial usability problems and bugs were reported to the design and development teams directly via Redmine¹⁴, a flexible project management web application while testing was still ongoing. References to these findings can be found in Appendix 2. The usability test report that lists general findings about the current usability and ideas on the possible next development level of Haavi-Alma can be found as Appendix 3 but here the main findings are summarized.

Executive summary

System Haavi-Alma Toolbar (<http://lucre.aamulehti.fi>)

Method Usability test (prior Expert walkthroughs)

Setup The Haavi-Alma Toolbar integrated on Aamulehti.fi was tested with 6+1 users in laboratory conditions. The tests lasted 1-2 hours including interview related to Haavi-Alma concept (Facebook/ social media integration, Mobile application). The main aim of the testing was to debug and improve the application for the upcoming field study. The most urgent findings were reported via discussions and Redmine directly to design and development teams to guide the fast iteration. This reporting will focus on findings (from testing and walkthroughs) that were not possible to integrate for the field study but were left for possible next level.

Tested functionality of the Haavi-Alma Toolbar included: creating an account, logging in, adding items to haavis ("butterfly nets") from article ingress page and details page, adding haavis (default haavi created automatically), archiving and hiding items (but no way to see those items again), deleting items, deleting haavis, inviting friends to one's haavis (users of Haavi-Alma via system and other users with their email address), accepting/declining invitation to friend's haavis and reassigning from friend's haavis.

Conclusions 21 problems were found (direct reporting not included), six of them considered to be critical. The main problems were related to features that were had to be left unfinished (Kaverilista, Hiding/archiving, Adding comment) and insufficient sharing information.

¹⁴ <https://redmine.media.tkk.fi/login>

Based on the use participant gave the To-do widget grade of 7,5 (at school grading scale 4-10) ranging from 6 to 8,5. SUS-score

-presenting participants' subjective assessments of usability- was in average 65 (0-100) ranging from 38 to 85.

The positive findings from the test included that adding items to haavi was easy. Participants also appreciated that it could be done from the ingress page without first needing to read the article. The overall usability of the Haavi-Alma Toolbar was considered to be quite good as such. Some participant even preferred this system over bookmarks in private use but for using it actively in their every day lives would demand possibility to comment the items and the system to incorporate also other papers.

2.1.4. Haavi-Alma Toolbar field study

Haavi-Alma Toolbar was improved based on the findings from the usability testing and Facebook integration was added. Focus of the study was to find out underlying meanings behind sharing news items in participants' everyday life.

The field study took place in late May 2010 and lasted for ten days. Ten participants were required: couples, friends and a small group of co-workers. Every participant had a pair who they knew in real life to share news items with in the study. The participants were asked to invite others to share news with.

Introductions to participants were given via email and phone. Participants were also provided a user manual including somewhat detailed instructions of the use and they were able to contact their own researcher (omatutukija) at any time via phone.

Participants had no major problems with the interface except some who did not get the idea on the whole. All participants had several haavis and the number of items in the varied from one to 30. Invitation to share links with outsider were unsuccessful but within the pairs and small group participants shared items of their interest, items that they knew other(s) were interested and work-related items.

The first findings were similar to the usability testing: To actually use this kind of toolbar in one's every day life would require the service to integrate more than one paper. There should also be possibility for commenting (basic interaction already designed but left out from the study because of the time constraints).

Haavi-Alma field study is reported here only briefly. More detailed findings (e.g. 1) identifying core functions for sharing news in personal communications and 2) evaluating the functionality of our solution in supporting this type of communication and sharing.) will be reported via different academic forums.

2.1.5. Lessons learned

Crossmedial services are highly potential in supporting tasks that take place across different contextual settings (stationary-mobile) and across different services. For example sports, hospital work, real-estate activities or information seeking and sharing. The initial idea was to incorporate Haavi-Alma on several Alma Media sites, including news sites on daily news and the economy as well as portals for selling items and searching for real estates. There were two scenarios that were initially on the table: 1) collaborative browsing and sharing of news and 2) collaborative browsing of real estate adds and managing the task of “real estate hopping”.

Here is an example of how the service might benefit the real-estate hopping scenario:

Jack and Jill are looking for a new apartment in down town Helsinki and are browsing through various websites that advertise real estates. They are using Haavi-Alma to create a haavi for collecting adds that are of particular interest to them (assuming Haavi-Alma is integrated with the sites they're browsing). Jill is browsing adds from home and Jack from work, when he has a free slot. Both of them have a real time view to their shared haavi and can see, when the other has added an item to it. This way they can share the browsing and search experience through Haavi-Alma. Individual haavis can be accessed on the mobile and utilizing the mobile phone's positioning capabilities, items in haavi's can be mapped based on their location tags. In this case for example, when Jack and Jill actually go out to see open-houses they're interested in, they can map the items in their haavi based on where they're located as well as based on their reported showtimes.

This is an example of how different interaction resources in this case complement each other within a use scenario, forming a crossmedial service.

However, we were only able to implement the service on one media within this project, the Aamulehti online news site. This enabled us to study the basic interactions that the service affords, and how the haavi concept applies to scenario 1, news recommendation tasks and social browsing within a single media. The initial aim of bridging gaps between different media (crossmedial browsing) was not accomplished and should be studied in the future. We were also not able to conduct extensive studies on the mobile application because we could not implement the Haavi-Alma on the Etuovi.com website (housing portal). This scenario is highly potential and should be taken further in future projects and studies.

Relevant business models need to be better understood. How to get different media providers involved and how to support integrating services in ways that will benefit the participating service providers as well as end users?

The strength of the Haavi-Alma concept is not only in its crossmedial aspects, but in that it is highly generic. Haavi as a concept is not restricted to shopping, information search or any specific scenario, but can be adapted for a variety of purposes. The idea of a shared list (in the form of a haavi in this case) may be useful in many cases and it would be interesting to research what kind of varied uses users may find for this concept within the online service ecology. According to our experience of the development process and the user studies, we believe that when the service is actually extended across media, it could be very powerful. Users' practices in the online world are diverse and fluctuating. They do not stick to one single site or service, but browse and navigate across several different sites within the contexts of their real world activities. This multicontextual browsing practice could be supported with a generic solution (such as Haavi-Alma) that helps to bridge and share online content.

2.2. *Wedding Organizer Service*

The Wedding Organizer Service (EVENT) was the first pilot of the LUCRE project. The first two EVENT related user studies¹⁵ and the service concept including design process details are initially reported in *D1.2: Report on domain areas and initial services*, the next user studies¹⁶ in *D1.4A: Current practices and resulting services*, and the background research¹⁷ also partly in *D2.1: Report on the state of the art- Part 1: Service composition - Part A: Service composition*.

2.2.1. **Concept**

EVENT was to be a service, consisting of a collection of tools that allow for organizing different types of events, be that event a wedding, business seminar, or football club's get-together. As a wedding scenario was considered the most challenging, it was chosen to be the pilot service on the grounds that if the tools allow organizing weddings, which can even include sub-events, such as bachelor parties, they can easily be used for many less complex events, too.

To get user feedback as early as possible we decided to build the Wedding Organizer Service with a component-by-component approach where the first components built would have provided tangible standalone services, and then, in a step-by-step fashion, other components could be added to them to provide a more complete event organizer service. Gradually, these components together could be used to organize various events of increasing complexity, and eventually we would have had a Wedding Organizer Service each part of which would have been tested by actual users.

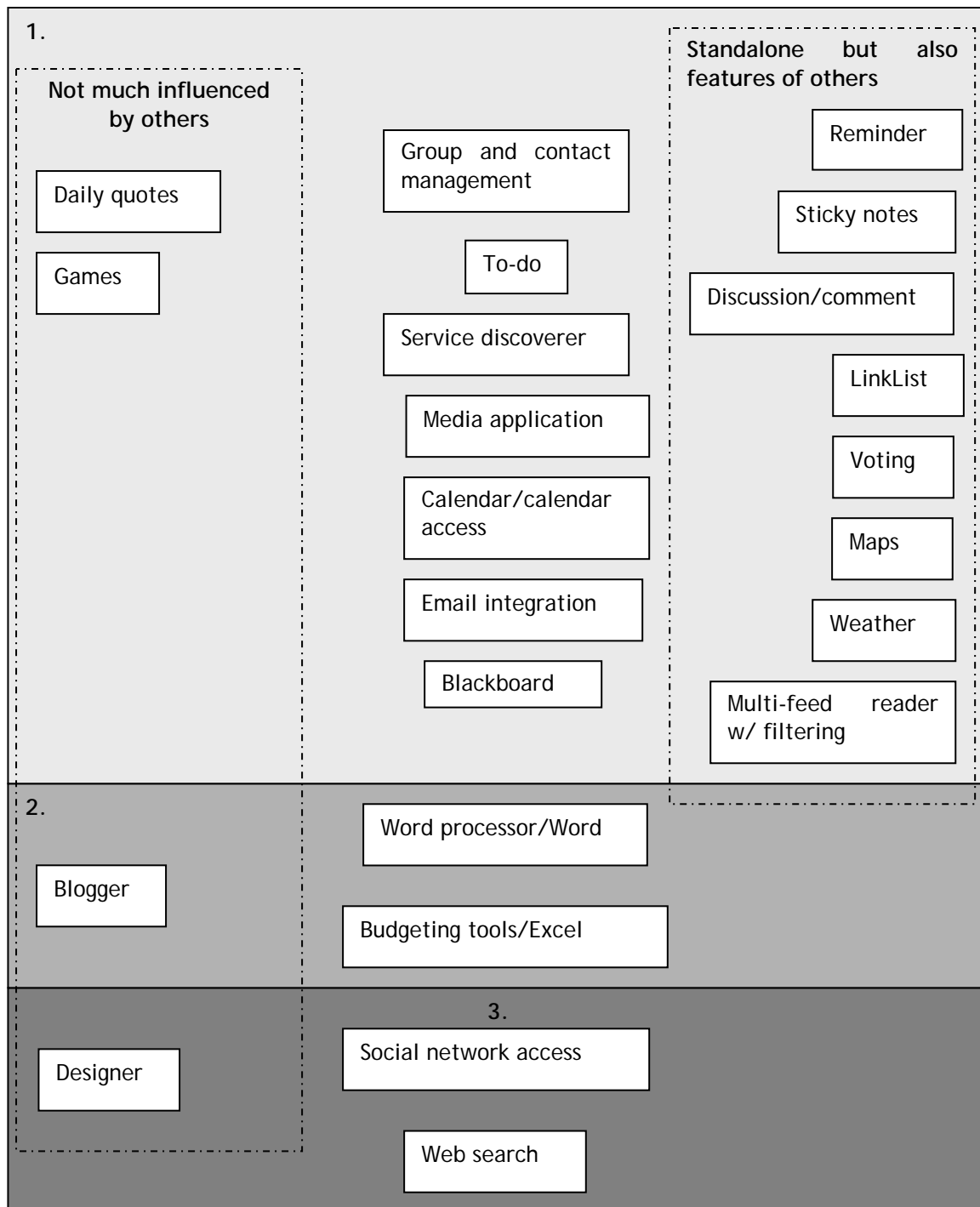
¹⁵ 1) first findings of the to-do list feature survey and 2) wedding organizing interviews

¹⁶ 3) detailed findings of the to-do list feature survey and 4) electronic PIM among smart phone owners interviews

¹⁷ review on the current service composition practices also from users point of view

This component-by-component approach was to help us design better the following components so that as more complex use scenarios become possible, we could trust the individual components to provide the correct functionalities and be usable. We could then focus on the use process and user experience sides of the service rather than worry about individual components.

Figure 4 traces out a roadmap to having a complete wedding organizer service. The central line follows the development of major widgets. The box on the right, titled "Standalone but also features of others," lists widgets that while being complete standalone widgets as such, are as often also used as features of other widgets. The box on the left, titled "Not much influenced by others," lists widgets that are typically not used in connection with other applications or as their features, and are thus typically used as standalone widgets.



4. Roadmap for widget development for the widget ecosystem.

The very first component that was designed and implemented was a to-do list widget (Figure 5, tested version). Our to-do list was planned to provide all the basic to-do list functionalities, including multi-user and multiple to-do list (work and home, for instance) scenarios, and different views to the lists (for instance, all my tasks from several multi-user lists in one view). We never got this far but basic sharing of multiple to-do lists with all users was made possible.



5. Tested version of the To-do widget interface.

The next development design steps were to create requirements and design for widgets that can be used as a feature in a to-do list: Reminder widget was selected because it can function as a standalone widget, generate reminders for the whole to-do list (for instance, tasks that are nearing their deadline), or for individual tasks. Reminders also tied us closer to the mobile environment, as they can also be delivered to mobile devices.

Reminder widget needs map widget for marking areas where reminders are delivered and for building location bookmarks that are then available for all widgets. Consequently, map widget followed reminder widget in design.

We were at this point when the project's focus changed. Some background research had been made to pick up another widget from the central line for developing. A good candidate was Media widget, a very complex widget that required certain amount of thorough background work. While media widget would not strongly integrate with to-do widget, it could have used the map widget for showing photo locations, and is thus related to the existing pieces.

2.2.2. To-do widget usability test results

Usability testing of the To-do widget was conducted to confirm the usability of the widget and to get ideas for further development and widget integration. The test report can be found as Appendix 4 but here the main findings are summarized.

Executive summary

System To-do widget V1 (<http://flexi.tml.hut.fi/todo/index.xformsdb>)

Method Usability test (prior Expert walkthroughs)

Setup The To-do widget prototype was tested with 10+1 participants in laboratory conditions. The tests lasted 60-120 minutes including interview related to diary study of the use of mobile phones in PIM (reported in D1.4). The reporting here focuses on guiding the possible next development level of the To-do widget (mobile use requirements and widget environment integration left for later).

Tested functionality of the To-do widget V1 included: adding tasks and folders, moving tasks and folders on the list, tasks also into folders, changing task details (priority, status, deadline, start time, reservation), making tasks starred, viewing starred tasks, sorting tasks (priority, status, deadline, start time, reservation), and searching tasks.

Conclusions 39 usability problems were found, nine of them being critical. The main problems of the widget were related to drag-and-drop feature, unfamiliar conventions (e.g. American style date marking), and missing details (e.g. folder reference). Also some icons were found problematic but it was already a known fact because there was no time to finish them before the testing.

Based on the use participant gave the To-do widget grade of 2,9 (at university grading scale 0-5) ranging from 2 to 4. SUS-score -presenting participants' subjective assessments of usability- was in average 56 (0-100) ranging from 43 to 83.

Participants mentioned that one of the advances of the widget was that adding items is clear and simple and that completed items are strike through in the item list. Dynamic search was also appreciated. The overall usability of the To-do widget was considered to be quite good and the scope also useful but revisiting some design issues would make it even more usable.

Simplified interface could solve many problems but on the other hand versatility was one of the widgets advances. One design principle was to make the widget as suitable as possible for both desktop and mobile use but on the other hand provide desktop users advancements like drag-and-drop. The vacuum from desktop to mobile still needs its' own design decisions.

2.2.3. Lessons learned

Paper and electronic to-dos questionnaire showed us how the replies that we get from people are colored by who those people are. Specifically, we found that the sex and the level of experience and exposure affect how people evaluate the importance of features for an application. Women turned out to place greater importance on larger number of features than men, and experienced users did the same in comparison to inexperienced users. Consequently, we have to be careful about whom we ask lest it biases the replies we get and misdirects our development work.

Diary-interview study gave us insight on how lead users, i.e. users with Internet-enabled smartphones, use their phones in PIM and in managing their timetables. Our participants were giving up paper calendars and moving towards all-in-one-device future. The phones will carry our to-do lists in addition to functioning as our calendars that alert us to different tasks. Furthermore, the selfsame phones function as cameras and amusement centers, enabling listening to music, watching videos, and browsing the net. However, our study also showed that we are not yet there, and the devices need to develop further for that picture to become reality.

User studies also showed us how differently people understand the to-do concept. To-do means different things for different people and needs related to to-dos evolve over time. That is why we need to allow customization and offer different templates (different basic sets of features) for users to get quickly one that works for their needs rather than force them to first learn how to customize a complex tool. Once they are settled in, let them customize, not so that customizing is the starting point. Sharing is not a default need in spite of it being a hot research topic - what people need and what interests the research community can be two different things.

The modular approach to Wedding Organizer Service appeared to be a good way of building the complex service into a widget ecosystem. We just never reached the point where the widget ecosystem could have shown its full potential. Still the design and testing the stand-alone To-do widget gave a good starting point. Widget integrations both inside and outside the ecosystem are left for future projects and studies.

3. Work summary of WP1

This Section collects all major studies, documents and design work related to two pilots: Wedding Organizer Service (EVENT) and LUCRE Media Pilot (LMP).

3.1. *Studies and documents*

WP1 produced quite versatile user studies related to user created services and the use of location information within the two pilots EVENT and LMP. Table 1 presents the academic publications, Table 2 summarizes the documents produced Table 3 the user studies conducted by WP1 during the first two years.

1. *Publications of LUCRE WP1 - first two years*

Publication name	Release date	Forum	Status	Content
Review of Service Composition Interfaces (D-2009-7)	October, 2009	The electronic report series of the University of Tampere	Accepted	Based on the D2.1: Report on the state of the art- Part 1: Service composition - Part A: Service composition.
	February, 2010	DIS 2010	Rejected	Based on the literature review for to-do, findings of the to-do survey and wedding interviews: design guidelines for PIM
The Times They Are A-Changin: Mobile PIM is Leaving the Paper Trail Behind	September, 2010	British HCI 2010	Accepted	Based on the findings of the smart phone lead user diary/interview study and the to-do survey: electronic PIM replacing paper based tools.
Mobile Lifestyle or Life in the Mobile: Lead User Study	September, 2010	Social Mobile Web -workshop, MobileCHI 2010	Accepted	Based on the findings of the smart phone lead user diary/interview study.
	September, 2010	NordiCHI 2010	Waiting for results	Based on the news reading and recommending survey: finding news online
	(August, 2010)	CSCW 2011	Waiting for results	Based on the findings of the Haavi-Alma field study: how does the application help socially constructed knowledge.
	(2010 - 2011)		Writing in progress	Description of the Haavi-Alma design case.

2. Documents by LUCRE WP1—first two years [table on three pages]

Document name ¹	Release date	Document type ²	Purpose	Content
FS Action plan [together with WPO and WP2]	December, 2008	Other	Required project document	Documents main objectives, definition of work and expected results for LUCRE.
LUCRE: Definition of UTA's Work, first research year	January, 2009	LUCRE Internal, Presentation	Document and present the possible actions on UTA's behalf.	Different ideas for UDC actions depending on the pilots to be selected in written form with visualizations.
Pilot service presentation: EVENT - The Wedding Organizer service	January, 2009	Presentation	Present The Event organizer (Wedding organizer service) pilot and VIPU pilot at higher level.	What is involved in a wedding organizing and how we plan to facilitate it with LUCRE pilot (Event organizer: Wedding organizer service) and (more briefly) VIPU, the travel-assisting pilot (later on dropped).
LUCRE Project plan 1 st year [with WPO and WP2]	February, 2009	Other	Required project document	Documents detailed focus, goals, and workplan for the first year of LUCRE.
LUCRE pilots: EVENT and VIPU	February, 2009	Presentation	Present the LUCRE pilot ideas to UDOI and other vertical projects in FS.	What, why and how related to pilot services.
The Wedding Organizer Service: Modular approach / Roadmap	March, 2009	LUCRE internal, Presentation	Document the proposal for approaching Event organizer development modularly.	How an event organizer naturally consists of different interconnected services that put together form an event organizer, and thus the development work can be approached modularly.
FS Case Description: EVENT	March, 2009	Other	Required project document	Documents the pilot details to FS level use.
LUCRE: SOAR, User-created services: Building and disseminating today	March, 2009	Other	Telling project findings to FS level, WP1 leader / project manager	Studied ecosystems and three approaches to building mini applications, forms, building by blocks and programming (by example) -- -themes and finding mini applications for use

¹ Document names might have some inconsistency because the pilot service terminology evolved during the two years.

² LUCRE deliverable, LUCRE Internal (documents produced for within-the-project use and not released), Presentation (presentations for internal use and within FS, e.g. UDOI Booster), or Other.

Document name ¹	Release date	Document type ²	Purpose	Content
D1.2: Report on domain areas and initial services	May, 2009	LUCRE deliverable	Required project deliverable	Domain areas and services for the pilot service based on a literature review on existing practices, a review of the features in existing services, a questionnaire study, and a round of user interviews.
D2.1: Report on the state of the art- Part 1: Service composition [with WP2]	May, 2009	LUCRE deliverable	Required project deliverable	The state of the art in service composition by end users (approaches, technologies, tools, user interaction, and services composed).
FS/UDOI- WP2 - Initial Research Framework V1.0	June, 2010	Other	Co-operation with other FS projects	Review of the FS co-project's deliverable.
LUCRE project plan 2 nd year [with WPO and WP2]	June, 2010	Other	Required project document	Documents detailed focus, goals, and workplan for the first year of LUCRE.
LUCRE Widgets Current planning state	August, 2009	Presentation	Document the widgetting approach: Interconnected and intercommunicating widgets form services.	Show the concept of how interconnected and intercommunicating widgets connect to form services within the uSpace widget ecosystem.
Pilot scenarios	Oct. 1, 2009	LUCRE Internal	Collection of possible scenarios for LMP	Collection of use scenarios for different pilot ideas.
LUCRE pilot	Oct. 7, 2009	Presentation	To decide upon the direction of the next pilot. Presentation of potential concepts.	Presentation of potential scenarios in monthly meeting for the next pilot including coaching, shopping and media browsing. Media browsing was chosen for further development.
LUCRE: Cybercom pilot (Suggestion: Heimoselailu/ Yhteisöseläminen)	Oct. 14, 2009	LUCRE Internal	Document an early suggestion about LMP.	A short review of currently available services, benefits to different parties and some technical considerations.
LUCRE Media Pilot	Oct. 19, 2009	LUCRE Internal	Documenting the LMP approaches.	Definition of LMP to be adopted for development.
LUCRE widget for media sharing and social recommendations	October, 2009	Presentation	LMP toolbar design with pictures.	Early design of the LMP toolbar documented for the project internal use and for development discussions.

Document name ¹	Release date	Document type ²	Purpose	Content
FS Case Description: LMP	November, 2009	Other	Required project document	Documents the pilot details to FS level use.
D1.4A Current practices and resulting services	Nov 30, 2009	LUCRE deliverable	Required project deliverable	Current practices in to-do use and lessons learned in the to-do survey study (2nd part), Lead user (smart phone) study results, and road-map and milestones for LMP.
LUCRE LMP roadmap	Nov. 25, Dec. 2, 2009	LUCRE Internal	A roadmap for the LUCRE Media Pilot	Planning the design and implementation process and timetable for LMP.
TimesPeople: Activities, Following and Followers	December, 2009	LUCRE Internal	Background study for LPM (what has already been done, what good etc.)	A study of how social recommending is done in TimesPeople of NYTimes to guide LMP development.
FS2: Everyday Services and Communication Tools	Mar. 12, 2010	Other	Overview of the themes studied in LUCRE project and to suggest some ideas for FS2 vertical Everyday services and Communication tools	What are "Everyday services and communication tools", why the theme is interesting and how it can be studied.
LUCRE pilots and POCs overview"	March, 2010	Other	Presenting LUCRE pilots to FS level, project manager.	Information on design, technologies and user studies on LUCRE pilots and POCs: Haavi-Alma Toolbar and Mobile, USpace, Calendar integration (WP2), GeoFeed (WP2).
Results of the first two years	April, 2010	Other	Presenting project findings to FS level, project manager.	The top three results of LUCRE assessed on novelty potential and technology maturity.
Haavi-Alma user guide (käyttäjähje)	May 14, 2010	LUCRE internal	Guide for the field study participants	Details of the study, registration and use of Haavi-Alma.
FS/UDOI- WP2 - Research Framework V2.1	May, 2010	Other	Co-operation with other FS projects	Review of the co-project's deliverable.
D1.6: User studies and evaluation of LUCRE tools	August, 2010	LUCRE deliverable	Required project deliverable	Give last results of user studies, to summarize the two years of UCD activities, and to elaborate lessons learned.

3. Studies conducted by LUCRE WP1—first two years [table on two pages]

Study	Date	Reported in	Purpose	Content
To-do feature survey (N=61) Conducted by UTA	May, 2009	D1.2 (first findings) and D1.4 (detailed findings)	To support to-do widget development for EVENT pilot.	Research plan consisted of features for the respondents to assess and comment in together with closed and open ended questions about the current to-do list use.
Wedding organizing interviews (N=10) Conducted by UTA	April - May, 2009	D1.2	To provide background understanding on wedding organizing for designing EVENT pilot.	Research plan for semi-structured interview (1-1,5 hours) with three recently married couples, three maids of honor and one best man consisted of questions about the wedding organizing process, e.g. materials and communication.
To-do widget usability testing (N=10+1) Conducted by UTA	Oct -Nov, 2009	D1.6	To ensure usability of the to-do widget V1 and to get base for further development and integration to other widgets of EVENT pilot.	Research plan consisted of predefined test tasks with private (first time use) and shared conditions in together with semi structured interview about the functionalities and the theme more generally.
Lead user (smart phone) diary and interview study (N=10+1) Conducted by UTA	Oct -Nov, 2009	D1.4	To support designing PIM for mobiles and provide background understanding for designing mobile application in LMP.	Research plan for semi-structured interview (1 hour) consisted of questions on how lead users use mobiles to support PIM and remembering today. Interview was partly based on a 5-day diary of mobile use the participants filled out prior.
News reading and recommending survey (N=147) Conducted by UTA	November-December, 2009	D1.6	To provide background understanding on news reading practices and support designing Haavi-Alma Toolbar in LMP.	Research plan consisted of questions on how respondents currently read, recommend and get recommendations for interesting news.
Haavi-Alma toolbar usability testing (N=6+1) Conducted by UTA	May, 2010	D1.6	To ensure usability of the Haavi-Alma toolbar before the field study and to get ideas for further development of LMP.	Research plan consisted of predefined test tasks within the theme of news reading for Aamulehti.fi. in together with semi-structured interview about the functionalities and the theme more generally. Separate research plan was also made for usability testing with Mikko.fi (selling and buying artifacts) but the integration was not possible within the project.

Study	Date	Reported in	Purpose	Content
<p>Haavi-Alma Mobile application usability testing</p> <p>Planned to be conducted by UTA but CANCELLED due to integration problems</p>	---	---	---	Initial research plan was made for the mobile application testing with Mikko.fi but the integration was not possible within the project.
<p>Haavi-Alma Toolbar field study (N=10)</p> <p>Conducted by HIIT</p>	May, 2010	D1.6 (focus of the study and first finding, more thoroughly in academic publications)	The provide understanding on underlying meaning behind sharing news items in participants' everyday life and to get ideas for further development of LMP.	Research plan consisted of questions on the interactions with Haavi-Alma for Aamulehti.fi in the wild.

3.2. Design work

WP1 conducted design work for both pilots. Table 4 presents designs for Wedding Organizer Service and Table 5 for LUCRE Media Pilot.

4. Design work for Wedding Organizer Service by LUCRE WP1

Design task	Date	Type ¹	Purpose	Description
To-do widget requirement [UTA]	May, 2009	LUCRE internal	To organize the main functionality of the to-do widget and prepare concept for specification.	Includes a description of the to-do widget functionality.
To-do widget wireframe [UTA]	May, 2009	LUCRE internal, Presentation	Design the interactions of the to-do widget.	Definition and visualization of main views and initial interaction flow.
Reminder widget requirement [UTA]	June, 2009	LUCRE internal	To organize the main functionality of the reminder widget and prepare concept for specification.	Includes a description of the reminder widget functionality.
Reminder widget wireframe [UTA]	June, 2009	LUCRE internal, Presentation	Design the interactions of the reminder widget	Definition and visualization of main views and initial interaction flow.
Map widget requirement [UTA]	August, 2009	LUCRE internal	To organize the main functionality of the map widget and prepare concept for specification	Includes a description of the map widget functionality.
Map widget wireframe [UTA]	August, 2009	LUCRE internal, Presentation	Design the interactions of the map widget	Definition and visualization of main views and initial interaction flow.
USpace / Spaces initial requirements / wireframes [HIIT/UTA]	August, 2009	LUCRE internal, Presentation	To organize the main functionality of the widget ecosystem and prepare concept for specification.	Definition and visualization of main views and initial interaction flow.
To-do widget interface design [HIIT]	June - September, 2009	LUCRE internal	Visual design for the to-do widget.	Layouts, icons, color scheme, look n' feel.
Media widget initial requirements [UTA]	June - August, 2009	LUCRE internal	To begin organizing the main functionality of the media widget and preparing concept for specification.	Includes an initial description of functionalities based on background research of current practices.

¹ LUCRE Internal (documents produced for within-the-project use and not released), Presentation (visual presentations for internal use).

5. Design work for LMP by LUCRE WP1

Design task	Date	Type ¹	Purpose	Description
Concept design for the media pilot [HIIT]	October, 2009	Presentation	To organize the main functionality of the service and prepare concept for specification.	Includes a description of toolbar, Facebook and mobile functionality including interlinkages between.
Visual design for toolbar and mobile application [HIIT]	Oct. 1, 2009	LUCRE Internal	Visual design for the LMP toolbar and mobile application.	Layouts, icons, color scheme, look n' feel.
Interaction design for toolbar [HIIT]	Oct. 7, 2009	LUCRE Internal	To decide upon the direction of the next piloting project. Presentation of potential concepts.	Definition and visualization of main views and initial interaction flow.
Interaction design for mobile application [HIIT]	Oct. 14, 2009	LUCRE Internal	Document and early suggestion about LMP.	Definition and visualization of main views and initial interaction flow.
Interaction design for Facebook-interaction [HIIT]	Oct. 19	LUCRE Internal	Designing the layout and process of Facebook interaction	Definition and visualization of main views and initial interaction flow.
LUCRE-LMP-mobile-specs-0.3-091222 [HIIT]	Dec. 23, 2009	LUCRE internal	Design of the Haavi-Alma Mobile Application	Detailed design of LMP Mobile Application [iterations also after]
Visual design for Haavi-Alma for Aamulehti (revision) [HIIT]	April, 2010	LUCRE Internal	To refine the look n' feel of the application for user testing.	Re-designing the color-scheme and logos for the media pilot.

¹ LUCRE Internal (documents produced for within-the-project use and not released), Presentation (visual presentations for internal use).

4. Conclusions

This Section concludes the both the document and the LUCRE project by first assessing the UCD activities done in contrast to project tasks and goals constituted in the beginning of the project and revisiting the UCD process and finally looking at the future of this type of services.

4.1. *LUCRE project tasks and goals for WP1*

WP1's goal was to come up with successful services for managing daily activities by studying current practices and reviewing literature to identify promising service areas. In addition, WP1 was to take part in designing pilot services with user-centered design (UCD) methods. The pilot services were to bootstrap the service ecosystem platform developed by WP2. In addition, WP1 was to test and study the implemented services to further develop them and to guide the development of other services.

Because the decision of LUCRE's financing came six months after the official start date, the work started in earnest the same six months late (In December 2009). This affected WP1's work, as there was no time to start identifying promising areas through user-centered methods. Thus, the service areas were decided by WPs of LUCRE together, and the current practices and literature review work was directed to support the decided areas. The eventual decision was to focus on an event planning with Wedding Organizer Service as a pilot (EVENT).

However, when main project partner dropped out of the project one year after being in, being out for the first year—their internal reshuffling had broken down the unit that was to take part in LUCRE originally—LUCRE had to undergo a shift in focus to take the remaining industrial partners' interests better into consideration. The new focus was in media and supporting social recommending. Again, the new focus was decided at project level, and WP1's work continued by studying current practices prevalent today in social recommending in online media and designing LUCRE Media Pilot (LMP).

WP1's task, taken from LUCRE Project plan (version 1.5, Updated activities for the second year application), were as follows:

- **Task 1.1: Initial pilot services:** Design of selected candidate services, from PoC's to adaptive and scalable interfaces to create requirements for the service composer.
- **Task 1.2: Generalized service-creation patterns:** Implementation of service templates that are applicable in many different contexts, and wizards that codify useful patterns in the service creation processes.
- **Task 1.3: Research on existing practices:** Lightweight ethnography, self-reports, interviews, and/or literature reviews for qualitative understanding on information and communication needs related to everyday activities

- **Task 1.4: Service releases and empirical trials:** The creation of releases, publication, and field trials with selected user populations.
- **Task 1.5: Evaluation of services:** Field trials and evaluation of the services created with the service platform by developers and end-users.

We will discuss the tasks in relation to the two focuses of LUCRE. For the Wedding Organizer Service pilot, the major undertakings of WP1 were:

- State of the art study of similar ecosystems, such as Popfly, Yahoo! Widgets, and iGoogle, to that planned in LUCRE. The study was done together with WP2, and WP1 focused on user interfaces and conceptualization of services. [Published in D2.1.]
- Literature review of current practices in PIM in relation to other people gave background information to user studies and candidates for individual widget to be implemented into ecosystem. [Published in D1.2.]
- User studies (to-do feature survey and wedding interviews) produced more detailed information on current practices and user needs to help designing the widgets. [Published in D1.2 and in D1.4.]
- Widget wireframe requirements and designs were completed for three widgets (To-do, Reminder, and Map) and on draft state for two more (uSpace and Media). [LUCRE Internal documents]
- Design process continued with detailed graphical user interface design for to-do. [LUCRE Internal documents]
- Usability testing for implemented To-do widget gave information on current usability and ideas for further development. [Published in D1.6.]
- Yet another user study (diary/interviews on users with Internet-enabled smartphones, use their phones in PIM and in managing timetables) was conducted along usability testing to gather information about life in the mobile advances and obstacles of electronic PIM. [Published in D1.4.]

Thus, WP1 went through the task list quite comprehensively even for the Wedding Organizer Service pilot. The widgets used most commonly in event organizing situations were identified, current user practices were studied and literature was reviewed, and the first widget underwent usability testing. The To-do widget was implemented in OtaSizzle when LUCRE's focus changed, and so field study was not carried out.

For the LMP, the major undertakings of WP1 were:

- Current practices review for background of the features and functionalities in similar service New York Times' TimesPeople to be used in the user study and system design [LUCRE Internal documents]
- User study (news reading and recommending survey) gave detailed information on current practices and user needs to help designing the system. [Published in D1.6.]

- Roadmap gave a structure to the pilot and first design versions of the Haavi-Alma Toolbar and Mobile application were produced. [LUCRE Internal documents, partly published in D1.4]
- Usability testing of the implemented Haavi-Alma Toolbar gave information on current usability and ideas for further development. [Published in D1.6.]
- Field study with gave the ultimate test for the system and possibility to study real life interactions [Partly published in D1.6, more thoroughly via different academic forums.]

Overall, WP1 managed to rather comprehensively go through the task list for both focuses of LUCRE in spite of the difficulties that the change in the focus caused. However, the depth of some work suffered from the shortness of time caused by this.

4.2. *Look in the future*

All services are moving to web: Microsoft moving Office step-by-step online to counter Google challenge to its money-cow, Facebook is becoming the platform for finding things on the Internet, and Google's approach is to provide all services online so that they can be used independent of the OS and device. Information and access to it are separated and people will access information with wide array of devices ranging from mobile phones and iPads to desktop computers with 22'' screens, and providing constant and continuous user experience requires designing the interfaces for all these devices (taking into consideration their processing power, screens size and other features) is the challenge. Forcing users to have their information in only on place and access it only with certain device is a recipe for disaster, as the competition will do better. Facebook and Google Wave¹⁸ are showing the way.

Our studies show that today, more and more of PIM is managed with mobile devices and users expect to have information at their fingertips independent of location and device. Portable devices are the way most things are done in the future but for hardcore working, desktops and laptops will still be with us for long time to come. Portables are used for viewing and light editing, PC for heavy inputting and processing.

Everything new needs to be easily accessed but also easily combined to the existing service ecosystem users already have because real world activities do not happen in individual sites and services. This demands integration between services to form a meaningful and usable experience instead of needing to hop between millions of separate services. The LUCRE project's underlining idea of generic shared lists could help bridge and share online content in various use contexts and ways when crossmedial browsing is made possible.

¹⁸ <http://wave.google.com/about.html>

Appendix 1:

5. News reading and recommending survey results

Table of contents

5.1. Introduction	33
5.2. Method and participants.....	33
5.2.1. Questionnaire	33
5.2.2. Participants	33
5.3. Questionnaire results summary	34
5.3.1. News reading habits	34
5.3.2. Finding news	36
5.3.3. Recommending news: Giving and getting	37
5.3.4. Recommendations from online sites	42
5.3.5. The impact of recommendations	42
5.4. Conclusions	44

List of tables

6. <i>Participants by gender and age</i>	33
7. <i>How important are different media for participants in following the news</i>	34
8. <i>Participants who ordered paper newspapers by age and gender</i>	35
9. <i>Participants who both had Internet connection on their mobile phones and used it for reading news online</i>	35
10. <i>How participants typically found online news to read</i>	36
11. <i>How often participants recommended and got recommended news</i>	37
12. <i>What channels participants used to recommend news items</i>	38
13. <i>Through which medium participants received recommendations</i>	41
14. <i>Recommender sites participants used</i>	42
15. <i>How many recommended items participants read typically</i>	43
16. <i>How important the recommendations coming from the different channels are to participants</i>	44
17. <i>How participants would like to get recommendations</i>	44

5.1. Introduction

The news reading and recommending study was carried out as an online questionnaire, and it gauged the current news reading and recommending practices to advise the designing of the social news reading toolbar, the axis of LUCRE media pilot (LMP). This Section outlines the results of the questionnaire replies.

5.2. Method and participants

5.2.1. Questionnaire

The survey form was built so that the respondents needed to answer only the relevant questions. For instance, if the “Yes” radio button was ticked on a question, the form opened to display more detailed questions, but if “No” was ticked, no further questions concerning the topic were asked. Consequently, we occasionally speak not of 147 people (the whole sample) but of a subsample of the relevant people.

The form had both open questions with text fields for answering and closed questions (multiple choice and evaluative ratings). All multiple choice questions had an *Other* option with a text field for entering details. Evaluative ratings were done with a 7-point interval scale. All questions had a text field for commenting.

5.2.2. Participants

Table 6 summarizes the participants by gender and age as reposted by the participants. As many participants found the survey through email list postings at the University of Tampere, University of Helsinki, and Helsinki University of Technology, it comes as no surprise that large number of the participants were either students or working in academia: students numbered 58 (inc. 4 graduate students) and other university personnel numbered 30 (25 researchers, 2 assistant researchers, and 2 professors), thus bringing the number of academia-related participants to 88¹⁹.

6. Participants by gender and age

	Males	Females	No gender given	Altogether
< 20		4		4
20-29	40	37		77
30-39	32	15	1	48
40-49	7	5		12
50-59	2			2
60 >	1	2		3
No age given	1			1
	83	63	1	147

¹⁹ Since only the profession was asked and not if working in academia, this number is likely to be slightly low.

All in all, however, there is no reason to expect that people in the academia are that different in their news reading and recommending habits from the population in general. Also, although the below 40s are disproportionately represented, the young represent the way things are going to be, and so this does not lessen the usefulness of the responses, either.

5.3. Questionnaire results summary

5.3.1. News reading habits

The participants were asked to rate different media for how important they were for them for following news. While online TV and radio are not that important yet, online papers were clearly central for our participants as far as following news was concerned: Online newspapers were considered the most important medium and TV only slightly managed to get to the second place past the online yellow papers (Table 7). Taken together, online newspapers and online yellow papers dominate the news following.

7. How important are different media for participants in following the news

	TV	Online-TV	Radio	Online radio	Newspaper (paper)	Newspaper (Internet)	Yellow papers (paper)	Yellow papers (Internet)
AVG	4.13	2.46	3.03	2.00	4.56	5.99	2.32	4.12
STDEV	1.96	1.67	1.79	1.51	1.96	1.28	1.49	2.10
Females AVG	4.78	2.25	3.46	1.95	4.87	6.02	2.57	4.54
Females STDEV	2.00	1.59	1.69	1.34	1.94	1.30	1.65	2.04
Male AVG	3.65	2.61	2.70	2.02	4.30	5.95	2.13	3.83
Male STDEV	1.81	1.72	1.82	1.65	1.95	1.27	1.34	2.09

Still, 119 (81%) participants did read news from paper newspapers, and 59 (40%) ordered them. Although the number of those who read online papers—141 participants (96%)—did clearly exceed the number of those reading paper newspapers (and only 6 participants (4%) did not read paper newspapers at all), the number of paper readers still underlines the importance of paper newspapers, too. And while 59 (40%) participants ordered and paid for paper newspapers, only 9 (6%) participants paid to access online news content.

On the positive side for the online news content providers—and Robert Murdock’s plans²⁰ to charge for online news content—the 9 people paying for the content were on the younger side, 4 being 20-29 and 5 being 30-39. The males were more likely to pay for online news content, as 7 out of the 9 were males, and 2 females. Thus, the future generations, especially males leading the way, might view paying for online content differently from the older

²⁰ <http://online.wsj.com/article/SB10001424052748704107104574570191223415268.html>

generation. In reverse, it is older people and females that order paper newspapers, as Table 8 shows. As our studies repeatedly show, males are spearheading the leap from paper-based tools to electronic tools, be it PIM or news paper reading.

8. *Participants who ordered paper newspapers by age and gender*

	Males	Females	No gender given	Altogether
< 20		2 (50%)		2 (50%)
20-29	9 (23%)	17 (46%)		26 (34%)
30-39	10 (31%)	7 (47%)	1	17 (35%)
40-49	3 (43%)	5 (100%)		8 (67%)
50-59	2 (100%)			2 (100%)
60 >	1 (100%)	1 (50%)		2 (67%)
No age given	1			1
	26 (31%)	32 (51%)	1	59 (40%)

The participants who read online news claimed to read them on average for slightly over one hour a day on computer (the average for those who paid for online content was slightly over 1 h 10 minutes), and slightly over 15 minutes on a mobile phone for those participants who both had an Internet connection on their phones and used it to read news (the 39 participants who fulfilled both conditions are shown in greater detail in Table9 - again, we see that the males are leading the adoption of online news reading with mobile devices). Out of the 39, 32 mentioned reading at least five minutes of news daily with their mobile phone.

9. *Participants who both had Internet connection on their mobile phones and used it for reading news online*

	Males	Females	Altogether
20-29	12 (30%)	4 (11%)	16 (21%)
30-39	13 (33%)	4 (27%)	17 (35%)
40-49	2 (29%)	2 (40%)	4 (33%)
50-59	1 (50%)		1 (50%)
60 >	1 (100%)		1 (33%)
	29	10	39

Altogether, 102 participants had internet connection (type—flat-fee etc.—was not specified) but out of them, 63 claimed not to read news with their mobile phone—although 2 mentioned in comment area (participants could comment their replies) that they did sometimes read news with their mobiles but not daily (one mentioned reading news with the mobile “rarely” and other “maybe once in two weeks”). (In turn, 3 people who claimed to use their mobiles to read news mentioned not doing so daily, and they had divided their estimated weekly news reading by seven to give us the daily reading time.)

Thus, it was 38% of the people with Internet connection in their mobiles that used them to read news.

Four people commented their news reading habits as far as location/situation is concerned, and all four mentioned using them when on the move, when sitting in a bus or train.

The users with mobile phone internet connectivity were more likely to recommend news in all categories but *Other* than the all the users in the sample, even if only slightly more so in case of *Tell a friend*. Furthermore, users who used their mobiles to read news were more likely to use each method than users with internet connectivity available who did not read news with their mobile, except for *Other* (half as likely) and *Tell a friend* (equal). Thus, internet connectivity in the mobile phone meant that those users used *Email, IRC, IM, and Social media* more than total of participants to recommend news. Users without internet connectivity in their mobiles, on the other hand, used clearly less other means than the users with internet connection, except when it came to *Tell a friend* (used slightly less) and *Other* (used a lot more).

5.3.2. Finding news

Table 10 summarizes how the participants reported that they typically found online news. On average, participants used 2.5 (out of the eight suggested ones) ways to find online news to read.

10. How participants typically found online news to read

	Method	No. who used
	News aggregation services	30
	Access online papers directly	123
	Access TV news's online services directly	68
	Recommendations from people	70
	Recommendation systems (e.g. dig.com or Slashdot.com)	15
	News letters from the news services	15
	RSS etc. feeds	32
	Other	8
		261

Although going directly to the source of the news, that is, a particular site one likes to read, dominated (accessing newspapers' sites and TV news's sites directly), recommendations from people did also have a clear role in finding and selecting what to read.

In Other, discussion forums were mentioned twice, and the same for Google and IRC. Using Twitter²¹ with keywords and iGoogle²² widgets were both mentioned once.

Of the news aggregation services, Ampparit²³ was the most popular (25 participants) and Google News²⁴ the second most popular (5 participants). iGoogle and Google Reader²⁵ were each mentioned once. On a scale of 1-7, the importance of aggregation services by those participants who used them was on average as 4.94. (One who claimed to use did not rate this.)

Among the RSS feeds mentioned by name (one user mentioned using over 400 feeds and another about 80—understandably neither named them), the most popular were HS (8 participants), YLE news (6 participants), and BBC (6 participants). For those who used them, the importance of RSS feeds was on average 5.28. (One who claimed to use did not rate this, and one who rated had claimed before not to use, and was thus not taken into consideration.)

5.3.3. Recommending news: Giving and getting

Roughly one third of the participants (46 participants, that is, 32%) recommended news at least several times a week (Table 11). In addition, roughly half (78, that is, 53%) got news recommendations from others at least several times a week. Recommending is thus an ongoing activity, and supporting it offers ways for online content providers to extend their reach and help their users find the most suitable content.

11. How often participants recommended and got recommended news

	How often do you get recommended news by others?	How often do you recommend news to others?	How often do you recommend news you have not read?
Several times a day	32	17	3
Several times a week	46	29	11
Several times a month	32	40	13
Once a month or less	27	43	40
Never	9	18 ²⁶	63
No reply	1	-	17
	147	147	147

²¹ <http://twitter.com/>

²² <http://www.google.fi/ig>

²³ <http://www.ampparit.com/>

²⁴ <http://news.google.com/>

²⁵ <http://www.google.com/intl/fi/googlereader/tour.html>

²⁶ One claimed never to recommend but then marked “Tell a friend” as a method of recommending.

As Table 11 also points out, participants did not tend to recommend news that they had not read, and thus the tools supporting recommendation can focus on recommending the open articles, as is done in NYTimes's TimesPeople²⁷.

When we look at how people recommend articles for others, we see that while there are no clear winners per se, social media, instant messaging, and IRC do have an important role and have already bypassed email in popularity. Such features as "Tell a friend," offered by many sites, on the other hand, do not play an important role in the equation (Table 12).

12. What channels participants used to recommend news items

Medium of recommending	No. of participants who use it
Email a link	44 (30%)
"Tell a friend" features	13 (9%)
IRC	62 (42%)
IM	51 (35%)
Social media	62 (42%)
Other	32 (22%)

When we look at what social media sites are popular here, we notice that there are basically only two players in the field, the clearly more popular Facebook²⁸ with 56 participants mentioning it and tenacious number two, Twitter, with 4 participants mentioning it.

Perhaps surprisingly, in *Other*, there are 21 mentions of recommending news face to face and 6 of recommending news over the phone (plus one mentioning of recommending them with SMS). The participants told, for instance, of showing a news item on their monitor to somebody in the room.

When the participants were asked what kinds of news they recommend, only 7 types of news are mentioned at least 10 times: Humor (49 times), Something the receiver is interested in (29 times), Odd/strange/surprising (19 times), Interesting (14), Work-related (13), Scientific (10), and All kinds (10).

The interestingness for the receiver factor was also evident in the comments by the participants:

"All kinds, anything at all, sometimes funny, sometimes shocking. It all depends on to whom do I recommend it."

"The news type depends on the receiver. If I know that somebody is not interested in politics, then I don't recommend that person any political news."

²⁷ <http://timespeople.nytimes.com/home/about/>

²⁸ <http://www.facebook.com>

"The kind that I think could interest them, have something to do with [their] hobbies or work."

Just like the participants did not recommend news items that they had not at least glanced, their recommendations were strongly influenced by their perception of the likes and dislikes of the receiver rather than using their own preferences as the starting point.

The kinds of news that the participants reported receiving (mentioned at least 10 times) were: Humor (54 times), Work related (23 times), All kinds (23 times), Odd/strange/surprising (20 times), What the receiver ("me" in this case) is interested in (14 times), Current affairs (12), and Politics (10).

When we contrasted the recommendation topic each participant recommended to others and the topics that s/he were recommended, we noticed that in case of 26, the topics were identical or very close, and in 74 cases different (the rest could not be determined from the answers). However, even in the case of 74, often the difference between the topics was not great.

When we look at the reasons why people recommend news items, we see that keeping in touch and maintaining relationships is a strong motivator behind the recommending activity. When asked why they recommended news items, 6 types of reasons were mentioned more often than 10 times: Humor (77 times), Informing (56 times), Conversation (esp. for starting one) (24 times), Taking part (15 times), Feeling that the item would interest the receiver (13 times), and Staying in touch (12 times). However, many reasons that were mentioned fewer times were also clearly related to maintaining social bonds, for instance, Topic of common interest (3 times), Sharing an emotion (3), Giving a good feeling for the receiver (2 times), Maintaining human relations (1 times) etc. Two participants even mentioned reciprocity by stating that they gave recommendations so that they would get recommendations.

However, reciprocity was not so evident in the ratings for the importance of recommending news to people who recommended news to us, as the participants valued the importance of recommending items back only at 3.75 on average.

On the other hand, having a possibility to comment on or otherwise acknowledge the recommendations was rated at 4.54, thus underlining the give and take involved in maintaining relations and staying in touch through recommending news.

Clearly over two-thirds of the participants (96 out of 129 participants who answered this question, that is, 74%) expected the receivers to read the news items that they recommended at least *Most of the time* if not *Always* (34 participants). Only 10 participants (8% of the 129 participants who answered

this question) expected the receivers to read *Less than half* (7 participants) or *None* (3 participants) of the news items that they recommended them.

Nevertheless, receivers reading the recommended items was not rated very important by the participants at 3.58. Neither was receiving feedback for the recommended news items seen as very important at 3.51.

On the other hand, 9 participants had stopped giving somebody recommendations, and 2 of them mentioned the lack of response as the reason for stopping. Two also mentioned that since they were less and less keeping in touch with the person, they did not see a point in recommending news to them, thus underlining the importance of the relationship in general for recommending.

As far as recommending was concerned, being able to recommend articles as whole (average importance 5.40) or individual pictures in the news articles (average importance 3.39) were considered the most important. Being able to recommend pieces of news articles was not seen as very important, its average importance being only 3.27. Including a piece of the article in the recommendation was somewhat more important at 3.20, but being able to include one's own comments was clearly more important at 3.71.

"Comment could be used to emphasize what made the news article important in the first place of why it's sent to this person in particular"

"That's where I say what I think about the news article, I don't want somebody to think that I actually consider the item important, usually the opposite"

Thus, being able to comment the recommended article is important so that idea behind the recommendation becomes clear.

When we look at the media through which the participants received recommendations (Table 13) and contrast it with how they made recommendations (Table 12), we notice the same trends: Email, IRC, instant messaging and social media are the norms. In receiving the role of social media is slightly more emphasized than in making recommendations. In Other, face-to-face recommendations (19 mentions), phones (4 mentions plus one mentioning of SMS), and discussion forums (3 mentions) were the norms, just as in making recommendations.

13. Through which medium participants received recommendations

Medium of recommending	No. of participants who received recommendations through it
Email a link	61
"Tell a friend" features	10
IRC	63
Instant messaging	49
Social media	82
Other	27

Reading the articles that others sent them was considered rather important by the participants (4.79 on average) but there were differences caused by the medium used, as user comments showed:

*"It depends on the medium
Email=important
Facebook=not important"*

"I read practically always the ones recommended directly to me, the ones broadcasted in Facebook less often"

"If it's a personal recommendation, then important, otherwise not that important."

The more personal the recommendation—made to me and not to everybody factor—the more likely the item is read. The medium is important because some mediums, such as email, are more personal (at least in perception), and others are less personal and more broadcasting to the world, such as Facebook. Thus, the importance of social media in comparison to email can be somewhat illusory, and supporting direct, personal recommendation is important to getting the items actually read and not only recommended and ignored.

Twenty participants received news recommendations from online services but services proved to be very popular, and at least some participants commented not reading them that often:

"I don't remember. I ignore most of them."

"I get some but I consider them spam and I don't read them..."

"Much of the time I don't read them..."

The more people recommend, the more recommendations they get. Reciprocity is very evident in that sense and at that level. Also, participants felt that they received more recommendations than they gave.

5.3.4. Recommendations from online sites

Only 19 participants followed online recommendation sites, such as Slashdot²⁹ and Digg³⁰. The services used by them are shown in Table 14.

14. Recommender sites participants used

Site	No. of participants using
Slashdot	15
Digg	5
Reddit	3
Sektor.com	1
Reader.google.com	1
LWN.net	1
Fark	1

The participants were also asked to rate the importance of certain elements of recommendations independent of whether they used them or not. Following editors or editor-in-chiefs for recommendations was rated slightly less important (2.20 on average) than following friends or people the participants knew (2.29 on average). Following specialist in the given field was clearly more important but even it was rated only at 2.77. Following unknown people who shared the same interests was rated 2.48 on average. Interestingly, getting automatically generated recommendations that took one's interest into account was seen as the most important of these at 3.06.

Thus, the participants were ready to get recommendations based on specialists in the given field or to get generated recommendation based on shared interests or their own interests. Here we are most probably witnessing the difference between the needs to be social and stay in touch with friends and making sure that we get good recommendations. Say-do problem may also be raising its ugly head, as the participants were to large extent guessing their preferences.

5.3.5. The impact of recommendations

Table 15 summarizes how often the participants estimated that they read the news items recommended to them. The impact of the recommendations is evident in the numbers: 94 participants (out of 145 who answered the question), that is, 65%, read most, if not all, of the news items recommended to them and 120, that is, 83% read at least half.

²⁹ <http://slashdot.org/>

³⁰ <http://digg.com/>

15. How many recommended items participants read typically

Statement	No. of participants
I read the recommended news practically always	37
I read most of the recommended news	57
I read about half of the news recommendations	26
I read some of the news recommendations but less than half	19
I never read the news recommendations	6
No reply	2
	147

Again, the comments by the participants underline the importance of how personal the recommendation is and what is the relationship between the recommender and the receiver:

"I don't get that many news recommendations but when I do, I read them (if they're from somebody I know well)" (reads most)

"When they come an individual person" (reads most)

"If somebody sends a link by email, I check it out most of the time. Stuff shared in FB I check out only if they interest me." (reads most)

The importance of the person making the recommendation is underlined by the importance of 5.45 that the participants gave to the importance of the identity of the person making the recommendation in relation to reading the item:

"Well, if it's from some jerk-off, then I won't bother"

"I don't even regard as recommendations fro instance the things coming up in my old job's skype channel or things people put in their Facebook status; a recommendation is personal and comes preferably from somebody you really know"

On the other hand, the headline of the news item was slightly more important than the identity of the recommender: The participants rated the headline's importance at 5.63. Perhaps the headline is, as an earlier quote suggests, more important when the recommendation is not from somebody we know and less important when the identity of and our relationship to the person making the recommendation take precedence.

In spite of being more likely to read recommended articles when the recommendation is personal, the participants rated the importance of recommendations coming through social media at 4.22 while those received through email were rated slightly lower at 4.15 (Table 16). All in all, while

“Tell a friend” feature was not considered important, others were all quite important, even if email and social media did get the highest on average importance.

16. How important the recommendations coming from the different channels are to participants

Channel	Average importance (1-7)
Email	4.15
“Tell a friend”	2.77
IRC	3.56
Instant messaging	3.84
Social media	4.22
Other	4.09

If the social media is seen as slightly more important as a channel, it is clearly the most important when it comes to how the participants wanted to get recommendations (Table 17). IRC remains important to its users and email is not far behind it, but the dominance of the social media here is clear.

17. How participants would like to get recommendations

Channel	No. of participants (multiple choices possible)
Email	59
“Tell a friend”	14
IRC	62
Instant messaging	48
Social media	84
Other	30

All in all, we see an interesting dilemma here. On one hand, people act the easiest on personal, direct recommendations, but social media is seen as both the desired and the most important channel for recommendations. Perhaps here, too, we are seeing two things. One, people like to stay in touch with their friends and close ones, and recommending news and discussing them is one way to accomplish this. Such recommendations are very likely to get read because of the relationship between the recommender and the receiver. Two, social media allows finding interesting news better than other available means, and the decision to read is more based on the item in question (headline) than the person making the recommendation. Thus, we could have here two different needs and two different modi operandi.

5.4. Conclusions

Online papers have become a major source of news. However, the huge number of possible news outlets and news that the Internet offers means that we must develop efficient ways to find the news that we are interested to

read. The most popular method of reading news was to access some news site directly, that is, using the service one knows and likes. However, nearly half the participants (70 people or 48% of the 147 participants) typically also used recommendations from people to locate interesting news items. In fact, 22% received recommendations several times a day and 53% received them at least several times a week.

While recommender sites, such as Slashdot, Digg, and Reddit³¹, aggregation services, such as Ampparit and Google News, and RSS feeds, were all used to locate news by some participants, their share pales in comparison to the recommendations from people.

The popular channels for giving and receiving recommendations were social media, such as Facebook and Twitter, email, IRC, and instant messaging. Surprising many also mentioned face-to-face conversations and even phone conversations and SMS. In addition, discussion forums still do play a small part in the equation, even if their importance is likely to have decreased from their past glory.

Recommending news appears to be part of social behavior in modern times. People like to stay in touch and maintain relationships, and recommending news that we expect the receiver to be interested in is a way not only to give the receiver something interesting to read but also to maintain relationship and to indicate that I know what you like. The fact that the recommendations are, at least occasionally, continuations of conversations and occasionally invitations to conversations are all part of humans being social beings.

Part of this is that news items are read based on who the person making the recommendation is. When there is a close relationship between the recommender and the receiver, the recommended item is likely to be read. Although the receiver reading the item is not so important to the sender, the receiver appears to feel compelled to read the item when the relationship is close.

On the other hand, the rise of social media has moved the recommendations to the realm of such social media sites as Facebook and Twitter. While Facebook ruled supreme among our participants, Twitter also plays a role in how news finds their readers. However, a recommendation over a social media site is not as compelling as a personal, direct recommendation, and other factors, such as the title of the article, come into play in deciding which items to read.

Thus, in designing ways to recommend news through such services, giving people information about the distal item is important to help them make up their mind about reading it. Furthermore, being able to give such items

³¹ <http://www.reddit.com/>

personalized feeling could also help. A news recommendation on a Facebook wall is not personal but one shared personally with one through Facebook can be a personal recommendation.

In making recommendations, being able to comment on the news item one was recommending was appreciated, as it made it possible to make one's stance in relation to the item clear and help receiver to see the important part. Responses from the receivers were not seen as essential but there was evidence of some sort of give and take taking place. For instance, lack of response had caused two participants quit recommending news some people. It appear that occasional response is appreciated, and thus enabling it is important, but it does not need to be a constant barrage of responses, and being forced to respond every time did not appeal to the participants.

In receiving recommendations, especially when not involving the social give and take, there was a clear preference to recommendations that took one's preferences into account. Flooding people with news recommendations that do not take their preferences into account is seen as spam and is likely to lead to reduced reading and acting (that is, going to read the news). Fewer but good quality recommendations help dealing with the endless number of news while too many non-personalized recommendations are not appreciated.

The only group of people the participants expressed interest in following for recommendations was that of experts in a given fields. If recommendation channels are built around people, then the people around whom they are built need to establish their specialization in that field.

In general, people do not recommend news that they have not at least glanced, which is in keeping with the goal of keeping up relationships—bad recommendations are unlikely to improve our standing in the eyes of others. All in all, people would like to recommend whole articles or individual pictures in articles. Being able to recommend parts was not in the most participants' wish list.

In closing, this discussion only went through the replies of the participants as a group, and it is likely that there are distinct groups of users within the participants that have different needs and goals. One system rarely can serve everybody, and thus needs to be designed for a group of users with certain shared characteristics. For instance, the users of social media could be analyzed in separation of other participants for developing applications to be used in the context of the social media sites. The values and ideas of people who do most of their recommending by email, for example, would not greatly contribute to designing an application for social media, and consequently, should not be considered at all or at least with reduced weight in relation to those of the social media users.

Appendix 2:

6. Haavi-Alma Toolbar usability problems and bugs reported in Redmine

This list (taken directly from Redmine) presents the volume and type of initial findings from Haavi-Alma Toolbar. No collective decision on how the reported problems and bugs should be named was made so some headings do not clarify the problem/feature in hand. Detailed descriptions and dialog between usability, design and development teams can be found from Redmine³².

#	Tracker	Subject
370	Feature	Details in new password request reply
369	Feature	Email invitation email content is misleading
368	Feature	Email invitation: the check of "@" missing
367	Bug	@Asking for new password: "Tilaa uusi" button pressed when text field left empty --> "ääkköset" do not work
366	Feature	@Registration "Jatka" button text needs to be changed to "Rekisteröidy"
365	Bug	@Registration "Keskeytä" button does not work
364	Feature	Email invitation: if new address rows can be added, those also need to be removable
363	Feature	In email invitation the ">" button should be removed
362	Feature	Removing own haavi or leaving from other user haavi: no indication to which haavi the actions is related to
361	Feature	@Registration "Rekisteröidy Haaviin" --> "Rekisteröidy Haavi-Almaan"
360	Feature	When leaving Haavi: "poista" "peruuta" --> "Kyllä" "Peruuta"
359	Feature	When accepting/declining invitation "Liity" and "Hylkää valitut" --> "Liity valittuihin" "Hylkää valitut"
358	Feature	Buttons need to look similar (activity)
357	Feature	Buttons need to look similar (coloring)
356	Feature	Cursor changes but there is no action on "+ Kerää ja jaa juttuja ystäväiesi kanssa!" and "+ Hiireile juttujen kohdalle ja lisää haaviin!"
354	Feature	Cursor does not change to finger on "Kutsu sähköpostitse"
353	Feature	Kaverilista needs button "Lähetä kutsu(t)" for Haavi-Alma invitations
352	Bug	Kaverilista "valitut(x)" remembers the last selection but does not show the content accordingly
351	Feature	Search in Kaverilista works unexpectedly

³² <https://redmine.media.tkk.fi/login>

#	Tracker	Subject
348	Bug	Kaverilista content is not clear - which users are shown as default?
347	Feature	Kaverilista search searches also other information than "nimi" like said in the field
346	Bug	Upper search field in Kaverilista does not work
345	Bug	Haavi members can be re-invited to certain haavi from the Kaverilista
343	Feature	The e-mail address of invited friend stays visible at Kaverilista
342	Feature	"Kutsu haaviin" text at email invitation email heading is not appealing
341	Feature	Email invitation receiver sees only the first name of the sender
340	Feature	"Rekisteröityminen onnistui" text is behind the "Kirjaudu" button
339	Feature	Registration could show mandatory fields with "*"
338	Bug	Email invitation directs user to the old pages
337	Feature	" Tähän tulee Käyttöehdot-linkki" at registration form
336	Feature	"Lue lisää..." at email invitation email is not a link
335	Feature	Dynamic search field remembers the previous input but does not limit the list accordingly
334	Feature	The contrast of the delete button ("x") and background is not enough
333	Feature	Is it possible that menus would close by clicking outside the the specific arrow
332	Feature	Cursor should not change to finger etc. if the object is not clickable
331	Feature	Tooltips are missing
327	Bug	Wrong kaverilista is shown when it is opened from haavi open list for haavi which is not currently selected)
326	Bug	Number of friends of certain haavi shows different numbers in different places
325	Feature	Notification of successful haavi item addition
324	Feature	Possibility to add "Oma viesti" while inviting users to certain haavi
323	Feature	Accepting/declining invitation to certain haavi pop-up should include sender details
322	Bug	Sharing work unexpectedly (order of my invitations versus invitations send by others)
321	Bug	Haavi item list shows duplicate of the each added item
320	Bug	Arrangement of haavi item list changes unexpectedly
319	Bug	"Tallenna haavi" button does not work in Kaverilista

Appendix 3:

7. Haavi-Alma Toolbar usability test report

Table of contents

Executive summary	50
7.1. Introduction	51
7.2. Usability test setup.....	51
7.2.1.1. Background	51
7.2.2. Participants	53
7.2.3. Test procedure	54
7.3. Results.....	54
7.3.1. General	54
7.3.2. Positive findings.....	55
7.3.3. Usability problems	55
7.3.3.1. Registration	56
7.3.3.2. Adding items and haavis.....	57
7.3.3.3. Haavi list properties	58
7.3.3.4. Haavi item list properties.....	58
7.3.3.5. Kaverilista (Buddy list) and Sharing	59
7.3.3.6. Concept and layout	61
7.3.3.7. Bugs.....	62

List of figures

6. LMP Toolbar interface on Mikko.fi site from design graphics by Petri Saarikko (Feb. 5 th , 2010).	52
7. Haavi-Alma Toolbar - before logging in (May. 31 st , 2010)	53
8. Pre-registration info screen.....	56
9. Left - Haavi list, middle - Haavi item list, and right - announcements and log out. 58	
10. Haavi-Alma Kaverilista (Buddy list).	59
11. Inconsistent buttons in Kaverilista: Kaikki, Poista valinnat, Lähetä kutsut, Peruuta.....	61

Executive summary

System Haavi-Alma Toolbar (<http://lucre.aamulehti.fi>)

Method Usability test (prior Expert walkthroughs)

Setup The Haavi-Alma Toolbar integrated on Aamulehti.fi was tested with 6+1 users in laboratory conditions. The tests lasted 1-2 hours including interview related to Haavi-Alma concept (Facebook/social media integration, Mobile application). The main aim of the testing was to debug and improve the application for the upcoming field study. The most urgent findings were reported via discussions and Redmine directly to design and development teams to guide the fast iteration. This reporting will focus on findings (from testing and walkthroughs) that were not possible to integrate for the field study but were left for possible next level.

Tested functionality of the Haavi-Alma Toolbar included: creating an account, logging in, adding items to haavis ("butterfly nets") from article ingress page and details page, adding haavis (default haavi created automatically), archiving and hiding items (but no way to see those items again), deleting items, deleting haavis, inviting friends to one's haavis (users of Haavi-Alma via system and other users with their email address), accepting/declining invitation to friend's haavis and reassigning from friend's haavis.

Conclusions 21 problems were found (direct reporting not included), six of them considered to be critical. The main problems were related to features that were had to be left unfinished (Kaverilista, Hiding/archiving, Adding comment) and insufficient sharing information.

Based on the use participant gave the To-do widget grade of 7,5 (at school grading scale 4-10) ranging from 6 to 8,5. SUS-score -presenting participants' subjective assessments of usability- was in average 65 (0-100) ranging from 38 to 85.

The positive findings from the test included that adding items to haavi was easy. Participants also appreciated that it could be done from the ingress page without first needing to read the article. The overall usability of the Haavi-Alma Toolbar was considered to be quite good as such. Some participant even preferred this system over bookmarks in private use but for using it actively in their every day lives would demand possibility to comment the items and the system to incorporate also other papers.

7.1. Introduction

This document outlines the results of the Haavi-Alma Toolbar usability evaluation. The Haavi-Alma Toolbar was evaluated by using two usability evaluation methods, expert walkthrough as a basis for usability testing. The goal of the evaluation was to confirm the usability before the proceeding field study and ideas for further development.

The test sessions, done in laboratory environment, included interviews on the Haavi-Alma concept (Facebook/ social media integration, Mobile application) but those are left for later. This document focuses on the findings strictly related to usability. Reported findings can be used to improve to usability of the Haavi-Alma Toolbar but those need to be incorporated with findings from other sources, e.g. the field study.

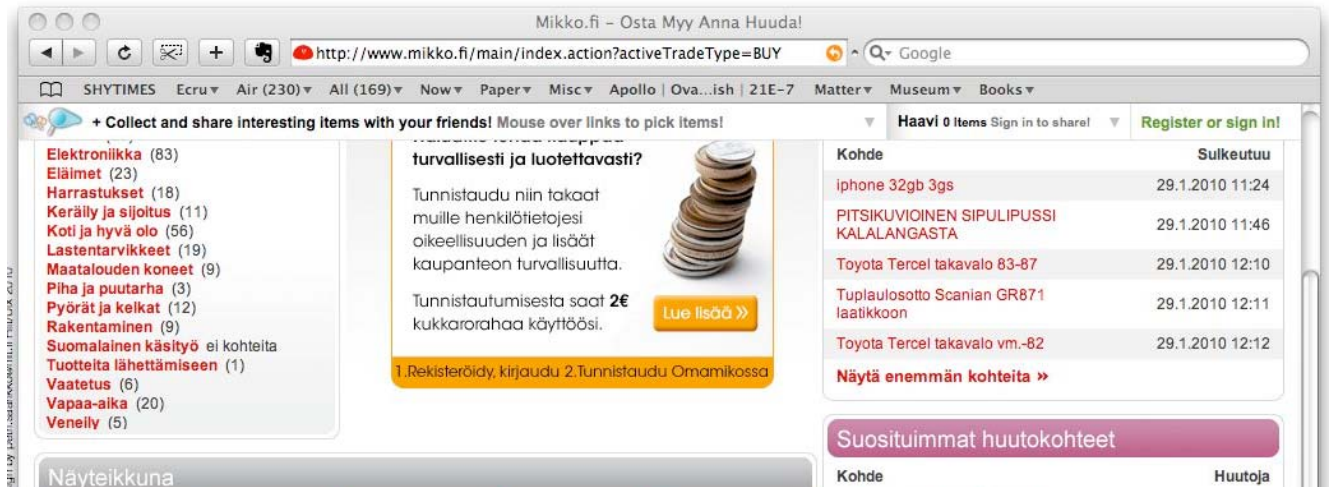
The document is constructed as following: Section 7.2. will explain the usability test setup and Section 7.3. will go through results.

7.2. Usability test setup

7.2.1.1. Background

In the beginning three different usability testing rounds were planned for Haavi-Alma Toolbar, social media (Facebook) integration and context aware Mobile application. Due the time constraints of the project in contrast to the development timetable, only one testing round was completed. Still we managed to combine two of the original themes: Browser Toolbar with shared lists and Facebook interaction on conceptual level.

Originally Haavi-Alma was to be integrated with Mikko.fi. It is a web service for buying and selling (Figure 6). The test plan with predefined tasks and semi-structured interview guidelines was done and 6+1 participants recruited based on a wide screener. The first criterion was that participants had used internet for buying and/or selling variety of things like in Mikko.fi, cars and/or apartments. Second criterion was that they had internet available in their mobile phones as a prerequisite also for testing the Mobile application.



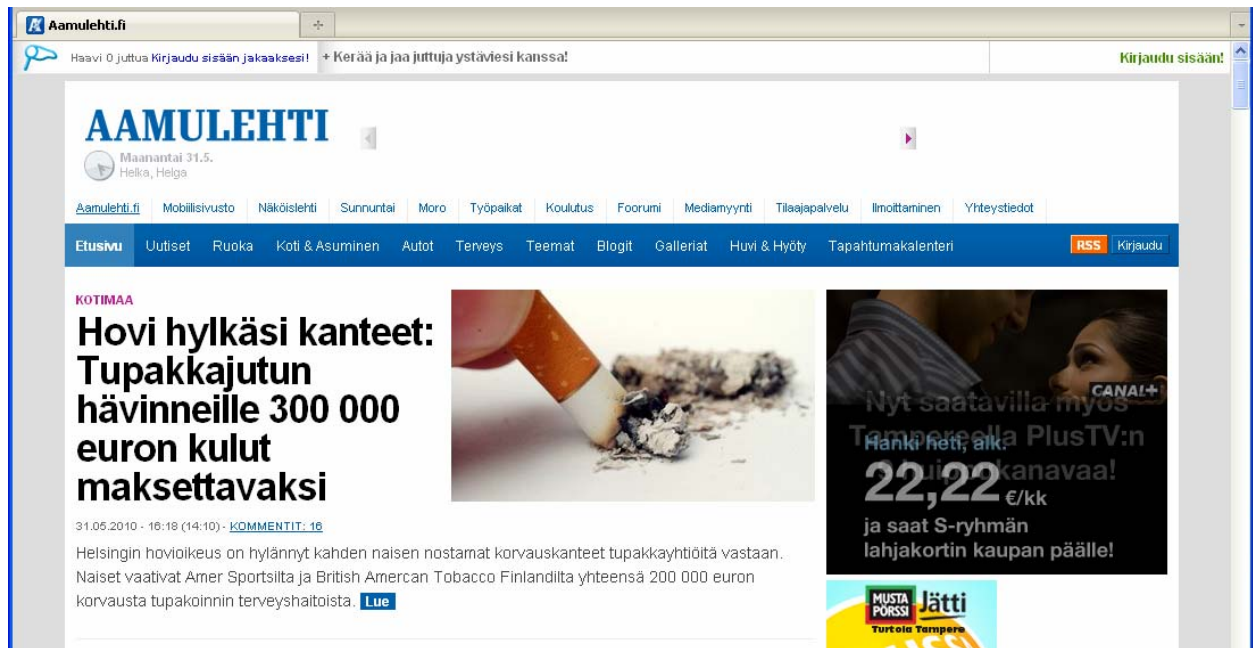
6. LMP Toolbar interface on Mikko.fi site from design graphics by Petri Saarikko (Feb. 5th, 2010).

The integration partner changes in the middle of development and the test were carried out with Aamulehti.fi service. This meant that new test task and interviews were planned and new set of participants recruited.

Design process included cognitive walkthroughs to guide the development prior to the tests. Also expert evaluations were made. Technical problems made it mandatory to use Mac computer and the newest Firefox browser for the testing. A traditional windows mouse and keyboard were attached to minimize the possible confusion while using the evaluated system.

The tested version of the Haavi-Alma Toolbar (Figure 7) supported following features:

- Creating an account
- Logging in
- Adding items to haavis ("butterfly nets")
 - From article ingress page and article details page
- Adding haavis (Default haavi created automatically after registration)
- Achieving items (no achieve available)
- Hiding items (no way to see hidden item)
- Deleting items
- Deleting haavis
- Inviting friends to one's haavi
 - Users of Haavi-Alma (via system)
 - Others with email address
- Accepting/declining invitation to friend's haavis
- Reassigning from friend's haavis



7. Haavi-Alma Toolbar - before logging in (May. 31st, 2010)

7.2.2. Participants

Pilot and six other participants were recruited for the usability test based on the wide screener (i.e. no direct user profile is to be used). The main recruiting criterion was that participants read either traditional newspaper version of Aamulehti or used Aamulehti.fi service. Among them were different types of readers from reading both forms of the paper daily, to reading mainly only the web version or the paper version to one who only read the paper less than once a month but was chosen because he lived on the Aamulehti coverage area.

The secondary aim was to select people who either recommend news themselves (some in Facebook) —and are so familiar with the concept of sharing— or only receive recommendation (some in Facebook)—or both—and, because not enough such participants were not quickly found from the existing resources project timetable limiting the time available for recruiting, one person who did not do either. On the contrary there was one who did not use any social media but was active on discussion boards.

The tertiary aim was to select both female and male participants of somewhat different ages because the actual sex and age divisions in the current user population of Aamulehti.fi were not known. Four of our participants were females and three males, ages 21-40 average of 28 years.

Four of the participants also had an Internet connectivity on their mobile phones which was kept in mind regarding the Mobile application

7.2.3. Test procedure

Each test session lasted 1-2 hours, consisting of different parts:

The test tasks covered the main functionality of the widget and tried to simulate the common use cases as much as possible at this state of widget development (one actual user, others imaginary -- private list, shared list with imaginary user). Some tasks were altered and verbal interventions were used to overcome technical immaturity.

User satisfaction questionnaire was based on System Usability Scale questionnaire³³. This ten-item five point Likert scale gives a global view of subjective assessments of usability within participants.

Post-test interview using the test task as bases for ideation of the concept, focusing on sharing between users, interaction across different sites, and mobile use.

Test procedure or test tasks did not change after the pilot test so also the pilot test data is included.

7.3. Results

7.3.1. General

21 usability problems in together with some bugs are reported here to be taken into consideration in further development of the To-do widget. 6 of them were considered to be critical. Most urgent findings were reported directly to Redmine, which is a flexible project management web application³⁴. There were only one week between the usability tests and the field study so most critical and obvious problems were reported already when usability tests were ongoing. There were also unified reporting with the designers based on discussions on the features of the next version, e.g. moving items from one haavi to another related to achieving and hiding items. These are not discussed here. These findings also date prior to the field study so some parts can be in contradiction to later findings.

SUS-score was in average 65, ranging from 38 to 85 and participants gave the widget 7,5 in the scale of 4-10 ranging from 6-8,5.

³³ Brooke, J. (1996) SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & A. L. McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis. <http://www.usabilitynet.org/trump/documents/Suschapt.doc>

³⁴ <https://redmine.media.tkk.fi/login>

Even though six participants had only tried Mac computers earlier and one never even seen Mac computers the strange operating system did not seem to affect based on participants actions and responses.

7.3.2. Positive findings

Participants mentioned that major strength of the service was that adding items was easy and could be done also before reading the article to be used later. Some found the application also useful as a substitute for bookmarks. All but one participants considered the scope of the widget to be somewhat useful but to be used on a large scale commenting possibility and some other papers need to be added so news items could be collected from different sources and users stand on those can be clearer.

7.3.3. Usability problems

Problems are presented within views and functions to help discussion on collective changes. Some design suggestions are given to individual problems but the big picture also needs to be revisited with the findings from the field study for cohesion if the Haavi-Alma development continues.

The following usability problem severity rating is used:

[1] Critical usability problem - Prevents the users from using the product in a feasible manner and therefore should be repaired immediately.

[2] Major usability problem - Complicates the use significantly and should be given high priority.

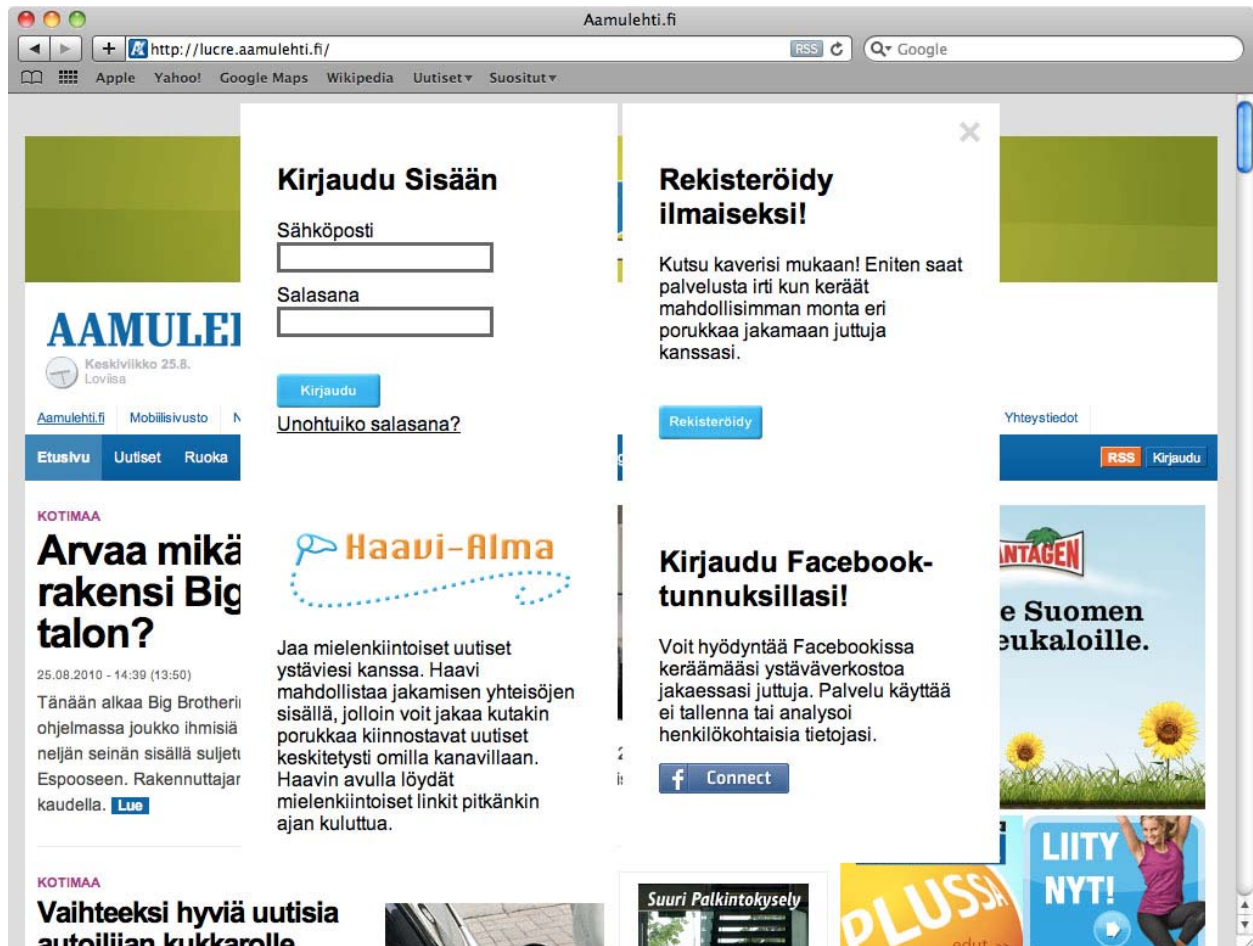
[3] Minor usability problem - Complicates the use of the product and should be given lower priority.

[4] Cosmetic usability problem - Should be repaired for the use to be as pleasant as possible.

[T] Technical problem - Technical problems are most likely due to technical problems with the product.

[C] Comment - Comments and observations by the test team or participants that are not directly related to usability problems but possible new features etc.

7.3.3.1. Registration



8. Pre-registration info screen.

Pre-registration screen does not fit to the screen with all resolutions and font sizes (Figure 8). [1] Important information about the service can be missed. During tests only two participants remembered the name of the service after registering and logging in. Others assumed that the name is Haavi which is visible also when logged in to toolbar.

→ If the instructions are to be kept on four parts “About the service” could be on top-right, registration on bottom-left (or rather unified with Log-in on top-left and Facebook login). The four parts could also be unified, bordered from the site content to add visibility and separated from each other with visible frames if needed.

Unclear what membership is valid for the system. [2] Text “Already a member sign in here” does not clarify whether the membership refers to Aamulehti or Haavi-Alma. It is also unclear whether Aamulehti subscription information is needed.

→ Information on pre-registration.

Shared personal information is not explained. [1] System does not clarify which parts of the information will be made public to the other users in the system, e.g. first name, last name, and email address.

→ Information on registration.

Shared information with Facebook is not explained. [2] System does not clarify what information is sent to either direction between Haavi-Alma and Facebook. This was a huge concern among test participants who also use Facebook for sharing actively. Because the needed information was not available they would have continued to use these two systems separately.

→ Information on pre-registration.

Registration process does not give needed details. [3] At registration there is no information whether some info will be sent to users email (convention that is familiar from other systems).

→ Information on registration.

Password is not changeable. [2] If user forgets the password and asks a new one he/she will receive a long row of letters and numbers. This password can not be changed so the safety is questionable because the password is sent via email and hard to remember.

→ To be made changeable or some information on registration.

Indication on successful registration is hard to notice. [2] After registration the user is not automatically logged in and the only indication is between the name and password fields and the log in button.

→ Information could be placed on the top of the screen where user's visual attention is at.

7.3.3.2. Adding items and haavis

Same news items can be added to same haavi many times. [1] Problematic in shared condition with extensive usage.

→ Important to fix so that one news item can only be added once to a haavi.

There is n possibility to leave a comment on added item. [1] This was a feature that did not get implemented for the field study due the time constraints.

→ Most important feature to be added.

Participant: "Need to be able to add comment so other people know how user reads the world. This is enough and other commenting and discussion with other forums."

[C] Some participants mentioned that the Haavi icon moving the heading of the article is problematic and should be moved to the end of the heading. On the other hand this is a good indication that something is happening.

[C] Some participants mentioned that adding items to haavi functionality needs to be also where other share tools are at the bottom of the article. This was especially because Share in Facebook icon is located there (though different and separate from Haavi-Alma Facebook integration).

[C] Some participants wished that items could be added to toolbar by clicking the toolbar.

[C] New Feature: Possibility to add same item to many haavi's at the same time. Relevant in situation where there are private and shared lists.

7.3.3.3. Haavi list properties

'+' is not clear enough when adding new haavi. [3] Some participants considered '+' to be unclear when adding haavi. They would rather have text 'Lisää'.
→ At least a tooltip could be added.

7.3.3.4. Haavi item list properties

Difference between "Arkistoi" and "Piilota" is unclear. [1] This was the main thing lowering the users grade. First of all buttons "Näytä piilotetut", "Avaa arkisto" are needed if these two functions are kept separately.

Users had very different ideas on what separates archiving and hiding, especially shared condition needs.

→ The features need to be carefully designed especially for shared conditions.



9. Left - Haavi list, middle - Haavi item list, and right - announcements and log out.

The connection between Haavi list and Haavi item list is not clear (Figure 9). [3] Haavi list is visually separate from Haavi item list.

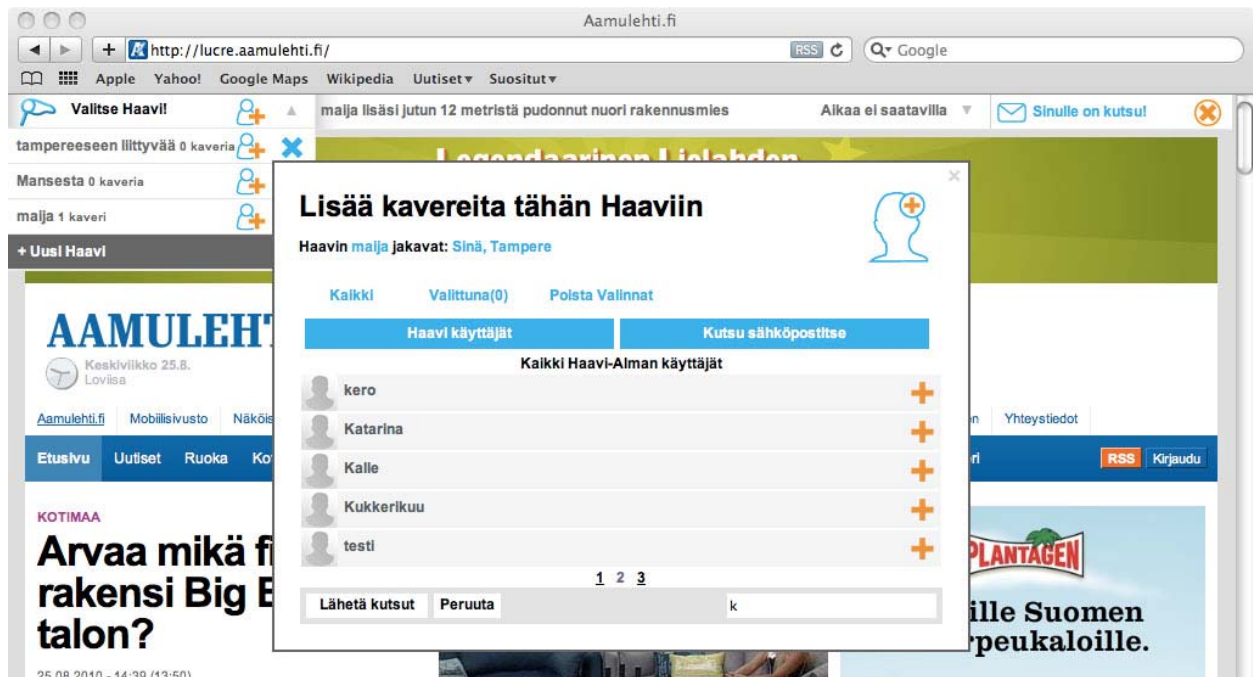
→ Indication in which haavi user is would be added to Haavi item list. e.g.

Name of certain haavi: Nimimerkki lisäsi jutun *Name of the article ...*

Another question is whether to word "juttu" is needed?

[C] How much content can be added to one haavi? What happens when haavi is "full"
→ Archiving?

7.3.3.5. Kaverilista (Buddy list) and Sharing



10. Haavi-Alma Kaverilista (Buddy list).

Kaverilista functionality was unfinished during the usability testing and some features were to be experimented during the field study. The current version (Figure 10) can benefit from the finding listed here but concluding decisions are needed to finalize the design.

Haavi-Alma invitation process is not clear. [3] After user has invited Haavi-Alma members or email friends to certain haavi, there is no confirmation on whether they were already added or if some confirmation on their behalf is needed. Participants were confused on what actions are needed on their behalf to complete the process.

→ E.g. following text could be added

“Kutsuit [Nimimerkki, nimimerkki], näet kun hyväksyvät kutsusi”

Kaverilista does not show the friends invited to haavi. [3]

→ Kaverilista could show who (=Haavi-Alma members) have been invited to certain haavi. This would prevent accidental double invitations.

No possibility to add personal note to the Haavi-Alma/Email invitations. [3] This was a feature that did not get implemented for the field study due to the time constraints.

→ Like in individual item posting also here the personal note could be used for additional information.

Received Haavi-Alma invitation does not include enough details. [2] Screen name and/or email are not enough. User need to know exactly who has sent the invitation.
→ First name, Last name are to be added to the invitation.

Kaverilista does not show which haavis have been shared certain user. [3]
→ This information could be added to the space between screen name and the icon of adding/removing this user from selected haavi.

Not clear whether user can invite third party to a haavi he/she has been invited to. [3]
→ In current design this is possible but there could also be different kinds of haavi memberships and indication for those.

[C] Default haavi is private so users need information on how to find sharing icon and the concept at the first place.

[C] Some participant wished that there can be certain haavis from which all the added items are also shared to Facebook.

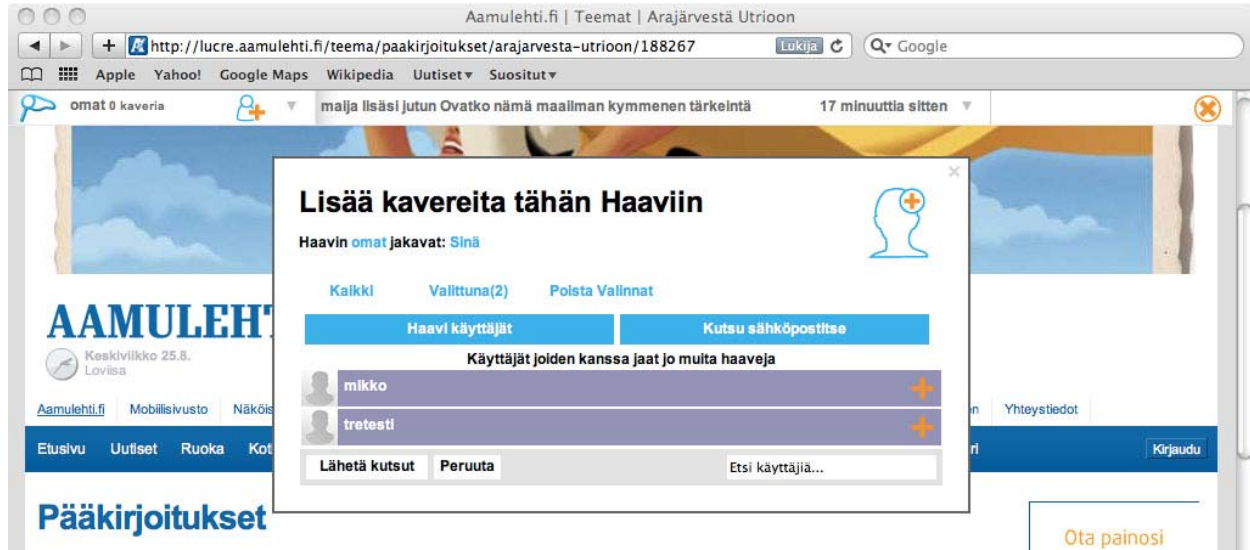
[C] Some participants wished that they would be able to control of items posted to Facebook, e.g. only for those Facebook friends who also are members of current haavi.

[C] In public haavi "Kaveri" is not a good term. One participant suggested "Followers" (seuraaja) instead.

[C] New feature: Possibility to change haavi from Kaverilista. That way the user would not need to first navigate to the right haavi from the Haavi list.

[C] New feature: Possibility to follow others actions throughout his/her haavi's. This would able some kind of following possibility as reported in current practices.

7.3.3.6. Concept and layout



11. Inconsistent buttons in Kaverilista: Kaikki, Poista valinnat, Lähetä kutsut, Peruuta.

Buttons do not look the same (Figure 11). [1] This is also affected by the fact that buttons don't look like buttons and are different by colors, e.g. *Lähetä kutsut*, *Peruuta* and *Kaikki*, *Poista valinnat*, and *Haavi käyttäjät*, *Kutsu sähköpostitse*.

Tooltips are missing. [2] E.g. '+' for adding and 'x' for removing good and consistent but when the user is not totally familiar with the concept it is hard to tell whether x means that friend is added or not.

→ Tooltips to all buttons.

Unclear what is clickable. [2] The cursor does not change to finger every time there is something to be clicked and sometimes there is a finger even when nothing clickable.

→ Revisiting the consistency.

[C] Many participants tried to open and close the toolbar outside the arrow by clicking the toolbar itself or outside the application.

[C] What haavi and which item is opened when logging in?

[C] Some participants wished that Haavi-Alma could be always open in browser and get notification like in Facebook. This form also need to be able to mark one self as not active if shown.

[C] Some participants wished that they could see new items and users in certain haavi when logging in. Selected important haavi events could also be sent directly to email or phone (need fast reading while dealing with apartments or cars etc.)

[C] Some participants wished that they could get desired information from Haavi-Alma to phone or email from selected haavi(s). One participant also mentioned a case where some person among friends group is looking for apartment. The haavi action notifications are important only to him/her. Different roles were suggested: informants and recipients.

7.3.3.7. Bugs

Haavi icon appears slowly the headings on the page. [T]

Whole page content has to be loaded before the haavi icon appears. This can sometimes take a while.

Haavi-Alma toolbar and pop-ups are behind some advertisement banners (@Windows/Safari). [T]

→ The required Firefox browser needs to be informed to the users if this is not possible to fix.

Can PC's firewall in some cases block the sending of referer in email invitation and the invitation is not delivered? [T]

Appendix 4:

8. To-do widget V1 usability test report

Table of contents

List of figures	64
Executive summary.....	65
8.1. Introduction	66
8.2. Usability test setup.....	66
8.2.1. Background.....	66
8.2.2. Test procedure	69
8.2.3. Participants	68
8.3. Results.....	69
8.3.1. General	69
8.3.2. Positive findings.....	70
8.3.3. Usability problems	70
8.3.3.1. Log in, log out	71
8.3.3.2. Creating items	72
8.3.3.3. Editing items	73
8.3.3.4. Reserving items.....	74
8.3.3.5. Moving items	76
8.3.3.6. Starring items.....	77
8.3.3.7. Sorting items.....	77
8.3.3.8. Searching items.....	78
8.3.3.9. General problems of different views.....	78
8.3.3.10. Bugs.....	81

List of figures

12.	<i>Original graphical design of the To-do widget interface graphics by Ayush Bajaj (version 1.9).</i>	67
13.	<i>Tested To-do widget interface.</i>	68
14.	<i>The log in view.</i>	71
15.	<i>The default view of the To-do widget toolbar.</i>	71
16.	<i>View for adding new item.</i>	72
17.	<i>The Save button is not visible editing view 1 (left) but only after entering editing view 2 (right).</i>	73
18.	<i>Three different colors of the reservation lock.</i>	74
19.	<i>Reservation feature within the editing view.</i>	75
20.	<i>Sorting view.</i>	77
21.	<i>Cluttered icon on the left.</i>	78
22.	<i>Folder reference is visible only at the general level not by the folder name.</i>	79
23.	<i>Home button and the state of the widget are not visible enough.</i>	79
24.	<i>Item list does not show priorities or deadlines.</i>	80
25.	<i>Three different pie icons on the right.</i>	81
26.	<i>The third row of the heading is only partly visible.</i>	81
27.	<i>Different size of red dots in items and folders.</i>	82
28.	<i>Closed To-do widget toolbar.</i>	82

Executive summary

System	To-do widget V1 (http://flexi.tml.hut.fi/todo/index.xformsdb)
Method	Usability test (prior Expert walkthroughs)
Setup	The To-do widget prototype was tested with 10+1 participants in laboratory conditions. The tests lasted 60-120 minutes including interview related to diary study of the use of mobile phones in PIM (reported in D1.4). The reporting here focuses on guiding the possible next development level of the To-do widget (mobile use requirements and widget environment integration left for later).

Tested functionality of the To-do widget V1 included: adding tasks and folders, moving tasks and folders on the list, tasks also into folders, changing task details (priority, status, deadline, start time, reservation), making tasks starred, viewing starred tasks, sorting tasks (priority, status, deadline, start time, reservation), and searching tasks.

Conclusions 39 usability problems were found, nine of them being critical. The main problems of the widget were related to drag-and-drop feature, unfamiliar conventions (e.g. American style date marking), and missing details (e.g. folder reference). Also some icons were found problematic but it was already a known fact because there was no time to finish them before the testing.

Based on the use participant gave the To-do widget grade of 2,9 (at university grading scale 0-5) ranging from 2 to 4. SUS-score -presenting participants' subjective assessments of usability- was in average 56 (0-100) ranging from 43 to 83.

Participants mentioned that one of the advances of the widget was that adding items is clear and simple and that completed items are strike through in the item list. Dynamic search was also appreciated. The overall usability of the To-do widget was considered to be quite good and the scope also useful but revisiting some design issues would make it even more usable.

Simplified interface could solve many problems but on the other hand versatility was one of the widgets advances. One design principle was to make the widget as suitable as possible for both desktop and mobile use but on the other hand provide desktop users advancements like drag-and-drop. The vacuum from desktop to mobile still needs its' own design decisions.

8.1. Introduction

This document outlines the results of the To-do widget V1 usability evaluating. The To-do widget was designed to be a part of larger widget ecosystem but at this first state it was evaluated as an individual widget by using two usability evaluation methods, expert walkthrough as a basis for following usability testing. The goal of the evaluation was to confirm the current usability and to get ideas for further development. More testing rounds were planned but never conducted due LUCRE's focus change from EVENT to LMP.

The test sessions, done in laboratory environment, included interviews on participants' current use of PIM. The interview was partly based on diary on the use of mobile phones in electronic PIM all participants filled-in before the test session. Interview results are reported in D1.4 *Current practices and resulting services*. This document focuses on the findings directly related to the To-do widget usability to help possible later revisions.

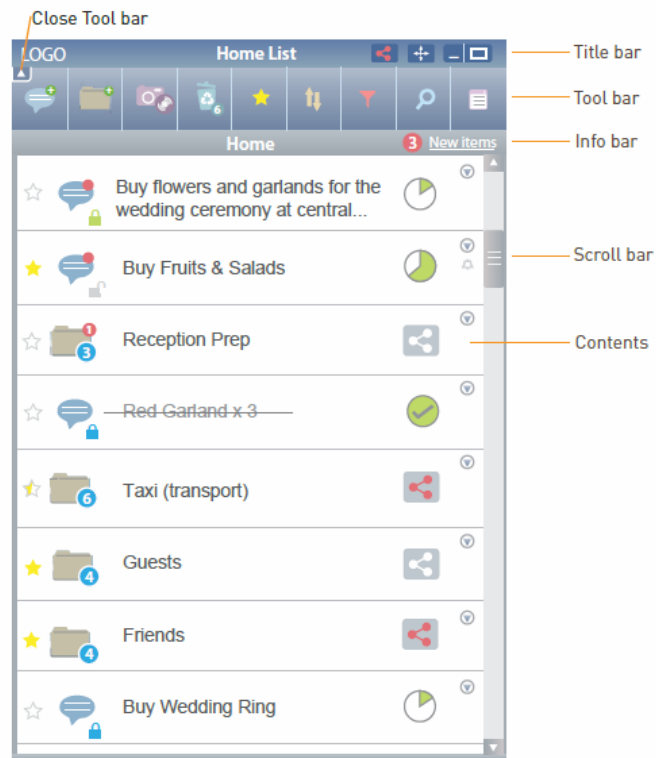
Because the project's focus had already sifted to the next pilot during the to-do widget usability testing the resources needed to be prioritized. That is why the findings here mainly present the individual problems and give some design suggestions but do not form a complete design for V2.

The document is constructed as following: Section 8.2. will explain the usability test setup and Section 8.3. will go through results.

8.2. Usability test setup

8.2.1. Background

Design process of the To-do widget was diversified. Figure 12 shows the original design. The basis for this design was to be able to help users with various complex tasks, wedding organizing process used as a scenario. At that point there were still some design issues undecided and the project focus started to shift away from Wedding Organizer Service (EVENT) to LMP.



LUCRE TO-DO Widget Interface (Home Screen)
310px * 470px

12. Original graphical design of the To-do widget interface graphics by Ayush Bajaj (version 1.9).

The time limits, technical constraints, and begun focus shift lead the user testing to be conducted with simplified version, without final graphics for icons (Figure 13).

The tested To-do widget V1 supported the following features:

- Adding tasks and folders
- Moving tasks and folders, tasks also into folders
- Changing task details
 - Priority, Status, Deadline, Start time, Reservation
- Making task Starred and viewing Starred tasks
- Sorting tasks
 - Priority, Status, Deadline, Start time, Reservation
- Searching tasks
 - Name and description



13. Tested To-do widget interface.

Five expert walkthroughs were made with the tested prototype. The walkthroughs commented the bugs, available functionality and the logic behind the widget to aid the developers in finalizing the to-do widget V1. Time constraints postponed some larger iteration to time after this first testing cycle. The following operating systems, Internet browsers, screen resolutions and text sizes were used for walkthroughs:

Mac OS X 10.5.8, Firefox 3.5.2, Resolution on projected image approximately 1280x960. Text size: Larger (*1st walkthrough*)

XP SP3, Firefox 3.5.2/3, Resolution 1440X900 / 1280x800. Text size: Normal (*2nd-5th walkthrough*)

8.2.2. Participants

Pilot and ten other participants were recruited for the usability test based on the wide screener (i.e. no direct user profile is to be used). The main recruiting criterion is prior experience with smart phones, phones with an Internet connection because these people would most probably be the users of the developed system. There participants were six females and five males, age 20-39 average age being 28 years. More detailed background information can be found in D1.4 Section 3.

8.2.3. Test procedure

Each test session lasted a maximum of 120 minutes, consisting of different parts.

Pre-test interview was about current use paper to-do's and electronic memory aids partly based on the background questionnaire online and diary study in which the participants were asked to report their use of mobile devices as external memory aid for five days. The diary consisted of date, time, used application, purpose of use, and whether it helped the participant to remember that thing. Findings are reported in D1.4 Section 3.

Test tasks covered the main functionality of the widget and tried to simulate the common use cases as much as was possible at this state of widget development. There were many rather small tasks due to the fact that using a to-do list typically requires frequent administrative tasks. First set of tasks was done with an empty list and the second with an already used and shared list. Scenarios were related to studying to the participants' background.

User satisfaction questionnaire was based on System Usability Scale questionnaire³⁵. This ten-item five point Likert scale gives a global view of subjective assessments of usability within participants.

Post-test interview used the test task as bases for ideation of to-do list future functions, mobile use and the widget ecosystem.

Test procedure or test tasks did not change after the pilot test so also the pilot test data is included.

8.3. Results

8.3.1. General

39 usability problems in together with some bugs were found. Ten of the problems were considered to be critical. Since LUCRE's focus has shifted from EVENT to LMP this document only lists the findings, including some ideas from participants and expert suggestions on how the individual problems could be addressed. Complete design for V2 would also require the big picture to be revisited for cohesion and the use context to be taken into account. It is worth remembering that this study focused the To-do as stand alone widget and separately from any widget environments or mobile use.

³⁵ Brooke, J. (1996) SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & A. L. McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis.
<http://www.usabilitynet.org/trump/documents/Suschapt.doc>

SUS-score was in average 56, ranging from 43 to 83 and participants gave the widget University grading of 2,9 (at the scale of 0-5) ranging from 2-4.

8.3.2. Positive findings

Some participants did not realize the meaning of the widget ecosystem/environment and the features and functionalities coming from that concept. Still the scope of the widget was considered somewhat useful as the widget offered flexibility e.g. in following kind of situation: Participants were asked to make a note of reading a library book before the return date they had various ways to do it. One participant made one item for reading and other for returning, one calculated the time needed for reading and marked that as the start date, one participant moved the return deadline earlier (based on the time needed for reading) and marked the actual deadline in description field.

Participants mentioned that one of the major advances of the widget was the fact the view for adding items is clear and simple (i.e. no need to fill in all the details). The fact that completed items are strike through in the item list was considered to be useful and the dynamic search was also appreciated even though it did not include the usernames.

The idea of improving usability with simplex and complex views got confirmation from the test. Simplex view should be the default but allow the user easily to increase and modify the features needed in his/her to-do towards the complex view.

8.3.3. Usability problems

Problems are presented within views and functions to help discussion on collective changes. The following usability problem severity rating is used:

[1] **Critical usability problem** - Prevents the users from using the product in a feasible manner and therefore should be repaired immediately.

[2] **Major usability problem** - Complicates the use significantly and should be given high priority.

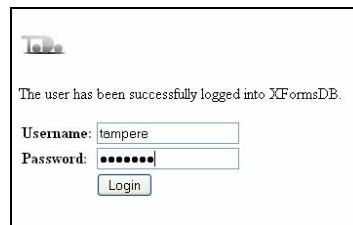
[3] **Minor usability problem** - Complicates the use of the product and should be given lower priority.

[4] **Cosmetic usability problem** - Should be repaired for the use to be as pleasant as possible.

[T] **Technical problem** - Technical problems are most likely due to technical problems with the product.

[C] Comment - Comments and observations by the test team or participants that are not directly related to usability problems but possible new features etc.

8.3.3.1. Log in, log out



14. The log in view.

Log in view shows unnecessary details (Figure 14). [4]

→ No need to show the "XFormsDB" text because it does not tell anything to average user and may even confuse them.



15. The default view of the To-do widget toolbar.

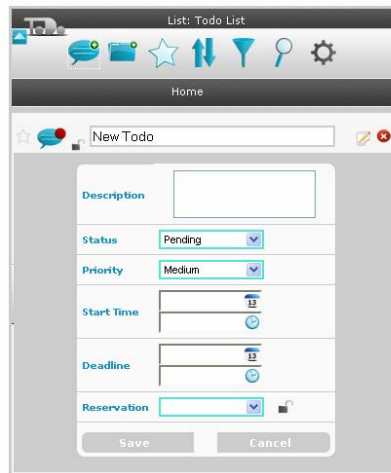
Log out button can not be found from the toolbar (Figure 15). [1] Log out can be found only inside the Settings menu. Some participants tried the Close toolbar button for this because there was nothing else; others just closed the browser window.

→ Log out button needs to be on the high level on the toolbar, if it can be used as a standalone widget.

User does not see to which profile he/she is logged in (Figure 15). [3]

→ Problem is irrelevant if to-do is used in widget ecosystem, still some identification of the user and the current to-do list is necessary if different profiles and/or to-do lists are possible.

8.3.3.2. Creating items



16. View for adding new item.

All views of the widget allow creating an item but those are visible and editable only in Home view. [1] User can create items while in different views (Home, Searching, and Sorting) of the widget. While not in the Home view participant gets no feedback about the item created and can not either add information to it. The created items appear only when returning to the Home view.

→ It remains to be decided whether the user should be able to add items in every view of the widget. The widget development actions need to be planned accordingly.

There is no text focus after opening (Figure 16). [2]

When the text field is clicked the focus goes in the end of default text. [2]

→ Focus need to be at the heading field by default and need to paint the whole default text

Date and time features look the same but function differently (Figure 16).

[2] The date and time selections look the same but only the date can be selected from a calendar view in addition to writing it to the text field.

→ Participants wished that the time would be able to select the same way as the date by clicking the clock icon, but suitable time interval a) easy to select and b) suitable for different kinds of items remains as a question. Participants also wished that the current date and time would be used as the default.

American style date marking confuses Finnish users. [1] Some participants noticed it already while making the item, some just later when asked to sort the items so that they can see three next deadlines.

→ Local conventions needed

The Save and Cancel buttons do not look like buttons (Figure 16). [3] The gray color makes the buttons look inactive and the pink when mouse over is very different from the blue icons used elsewhere.
→ Same color schema needed for the whole widget.

The defaults can not be adjusted. [3]

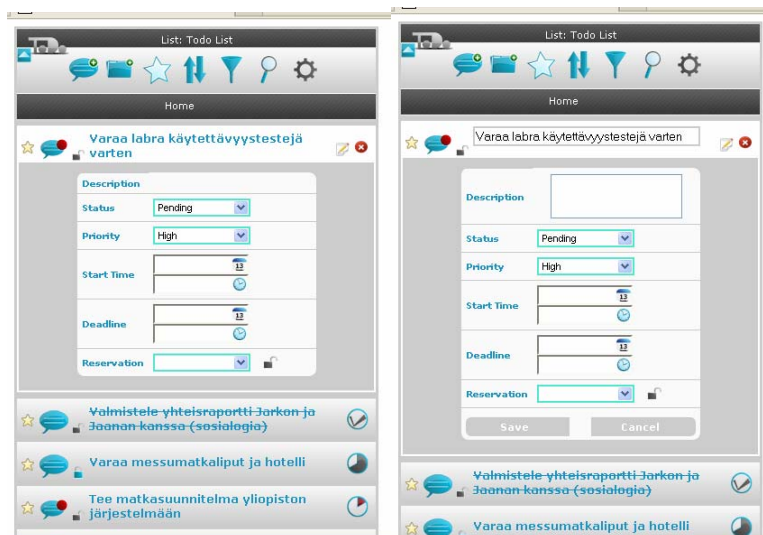
→ Participants wished that they could choose defaults themselves, depending on the use verified use cases.

[C] Some participant mentioned that the icon for adding new item resembles more like IM than adding a to-do item, but the function was easily understood from the widget context.

[C] Participants often marked importance with capital letters or explanation marks instead of changing the Priority. Still when asked or explored widget freely, they felt it was understandable and useful.

[C] Participants wished to get information to email or SMS e.g. when item deadline is approaching especially if they have not visited the widget for a while.

8.3.3.3. Editing items



17. The Save button not visible editing view 1 (left) but only after entering editing view 2 (right).

The items can not be opened from outside the heading text. [3] Participants tried to open the items from the bubble and the pie.

→ Technical reason why the whole row can not be used to open the item?

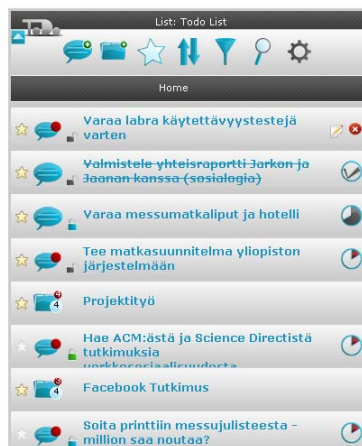
The Save button is not always visible when changing item details (Figure 17). [2] The Save button is not visible when item is opened from its heading to editing view 1(left) but only after the editing button on top right of the heading is pressed to enter editing view 2 including possibility to edit heading and description (right). The changes at view 1 in item details are saved without separately saving. The lack of Save button does not communicate this procedure to the user. There is no feedback on the changes, so while changing details some participants closed and opened the item again to check whether the changes had been saved.

→ Close arrow could be used here to communicate the next step.

[C] Against prior assumption the participants quite naturally learned to close the individual items from the heading text.

[C] One participant wished that the item could be marked done without opening it.

8.3.3.4. Reserving items



18. Three different colors of the reservation lock.

The coloring of the reservation locks is not clear (Figure 18). [2] The idea of locks was widely accepted among the participants, but the lock itself is quite small and the color differenced are too (e.g. compared to the red dot). The closed lock icon helps the understanding. Assumptions about the reservation colors (*black* not yet reserved, *green* for me, *blue* for someone else) were typically different (green available/done). To one participant the lock resembled user rights and he suggested the reservation to be face, avatar etc. → Tooltips (and Help)



19. Reservation feature within the editing view.

Not clear whether reservation selection is needed (Figure 19). [2] The empty field does not indicate what to do.

→ The reservation list should have text “Choose” as default (if no other default is used, e.g. the user him/herself).

It is unclear whether the lock icon is clickable. Many participants tried to click the lock icon because there are also other similar kinds of items functioning as buttons.

The place of reservation feature is confusing (Figure 19). [3]

→ Reservation could be placed upper in the item details because it is so vital part of it in shared list.

[C] “Reservation” was mentally combined to choosing location among participants. This is partly because the username in tests were “tampere”.

→ One participant suggested “Responsible” in stead.

[C] Depending on use cases there might be need to be able to give same item to many users e.g. by creating some user groups.

[C] Participants wished that some confirmation would be sent the person assigned either via email or widget itself.

[C] Participants had some concerns about the fact that anyone can change the settings or the reservation.

8.3.3.5. Moving items

Drag-and-drop does not function in the same pace as user does the actions.

[1] Participants were familiar with the concept of drag-and-drop from other than web contexts. Many preferred it to be the best way to move the item if it can be made to function better.

→ Need to move smoother, maybe without the transitions, at least start moving right away.

Indication where the item is to be dragged is unclear. [1] The point of the cursor does not indicate the spot where the item is to be dragged but the middle point of the item bar. So depending on where the user has taken the item the place changes. Also open items can be moved and then the target spot activates the same way in the middle.

→ The yellow line is a good indication color for the target but it needs to be a wider for the users to notice it.

The widget does not provide any means for undoing actions. [2] The fact that there is no undo feature is especially problematic when drag-and-drop feature is used.

→ There should be either a way to undo the actions or to have same kind of confirmation in drag-and-drop feature like when deleting items. If this is considered to be technically too difficult, solutions of the other problems need to be addressed more thoroughly.

“Item + item = folder” is unclear. [2] Some participants noticed that a folder was made when one item was moved on top of another but most did not even notice it. When asked what happens when one item is dragged on top of another item, most participants assumed that the target makes room for the new item on top itself. According to one participant this unexpected feature developed like this brings more harm than use.

→ While dragging item on top of another item before the folder is created the user should be asked confirmation about the event, possibly also name for the folder. The default can be to target items name like in current design.

Moving can not be done with arrow keys as an alternative to drag-and-drop.

[3]

→ One participant wished that moving would be possible also with the arrow keys, especially if the drag-and-drop works only as it works now. This is important to take into consideration also when designing the mobile widget.

[C] Participants had troubles while dragging and dragged outside the widget. They mentioned that this should not be possible without trying the concept of widget used in widget ecosystem.

8.3.3.6. Starring items

Starred items list shows also others than starred. [1] Starred items list shows also the items which deadline is near. This confused participants.

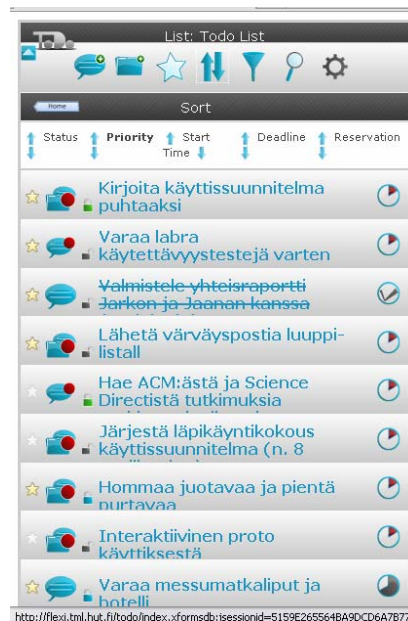
→ If the functionality is kept as it is now this needs to be communicated also to use.

The highlighting of starred items is not visible enough. [2] Participants mentioned e.g. that the red dots showing pending items is much more visible but only half as important as important to stand out as this one.

→ Different background color or bolding were suggested by participants but they self also considered the situation where everything is marked very brightly unbearable.

[C] Some participants assumed that starring is directly related to high priority.

8.3.3.7. Sorting items



20. Sorting view.

The direction of sorting is not clear (Figure 20). [1] Participants were not able to tell whether the sorting is done ascending or descending. Also the text itself is clickable.

→ Clearer visual feedback is needed to tell the user how the sorting is done. Tooltips on arrows would partly help.

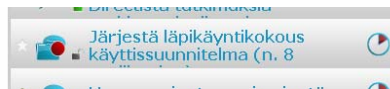
It is unclear how the sorting is done: passes deadlines, completed items etc. [2] E.g. what happens with the items that have already reached their deadline and/or have been completed?
→ Revisiting the design.

It is unclear whether multiple sorting is possible (Figure 20). [3] No visual feedback on what has been selected.
→ Adding correct visual feedback related to previous problem will solve also this one.

8.3.3.8. Searching items

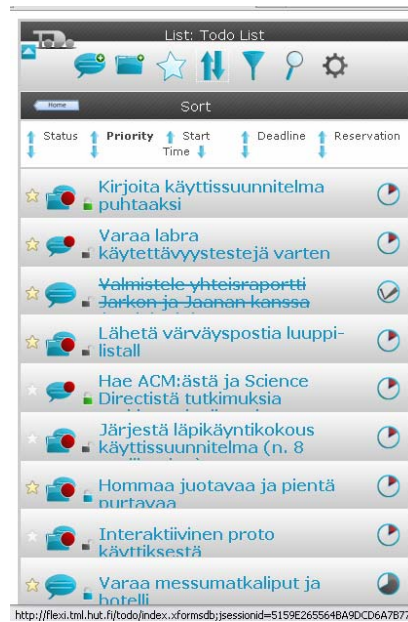
Search function does not work as assumed. [2]
→ Participants assumed that the search would include also the usernames.

8.3.3.9. General problems of different views



21. Cluttered icon on the left.

Icons are cluttered (Figure 22). [2] When the item list is opened from folders the item belonging to some folder is cluttered because so many individual images are combined in it on top of each other.
→ Finalizing of the iconography.



22. Folder reference is visible only at the general level not by the folder name.

Folder name reference is not always visible (Figure 21). [1] When the item list is opened from folders e.g. while sorting the folder reference is lost. →Item needs to display the information in which folder the item is in.

All items can not be seen in the same list at the Home view. [2] In the home view the items in the folders can only be seen opening folder by folder. → Participants wished that the items inside folders could be unified to a collective list like in Sorting view (Figure 21).



23. Home button and the state of the widget are not visible enough.

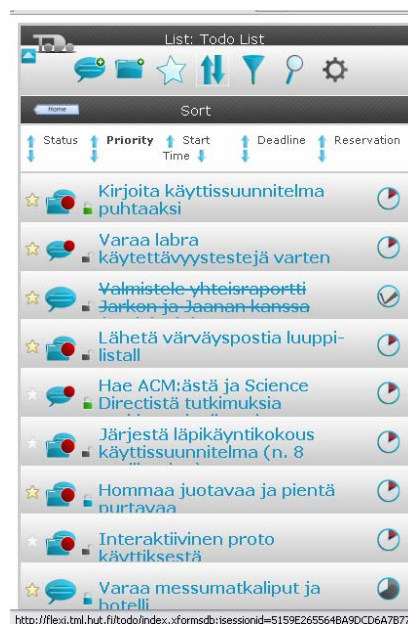
The widget state is not clear (Figure 24). [2] The widget does not communicate its state (Home, Sorting, Searching) clearly enough. Sometimes participants were unaware in which view they were in and what actions they were able to do there, e.g. moving items. →Widget state could be presented e.g. with a bigger font.

Returning to home view is not noticeable enough (Figure 24). [2] Some participants tried to return to the home view by clicking the buttons on the toolbar again.

→ Bigger Home button

Moving between views is unclear. [2] Some participants were not sure how to return the Home view. In Search view one participant deleted the search term and the dynamic search gave him view with all the items. Participant thought that he was in Home view. In Starred items view other participant clicked the star icon on the toolbar again trying to get to Home view.

→ Design suggestion to previous problem might help also with this one.



24. Item list does not show priorities or deadlines.

Priority is not visible in the item at the item list (Figure 25). [1]

→ Priority need to be included to the item.

Deadline is not visible in the item at the item list (Figure 25). [1]

→ Showing the item deadline can be done many ways, e.g. showing it on the item list or enabling timeline view with calendar integration.

Some tooltips are missing. [2]

→ Tooltips needed for all icons, even though they were considered to be quite clear.

No Help (file) available. [2]

→ Some participants asked for the help already during the test session.

Deleting folder but not its content is unclear. [3] Participants assumed different outcomes, e.g. all the items would go on their original places.
→ Rethinking of the concept. Is it even possible to remove the folder if there is something inside?

Red dots do not indicate what they stand for (Figure22). [3] The red dots in current version mark items with Pending Status. Participants on the other hand assumed that they display the new or important items.
→ Tooltips (and Help), maybe reiteration of the concept.



25. Three different pie icons on the right.

Pie icon is not clear (Figure 25). [4] Participants had problems interpreting the pie icon (from the top: Completed, Pending, In progress). Some thought that the second pie icon in Figure 25 is a clock telling about the approaching deadline. Participants had different kinds of suggestions how they would like the presentation best but they varied a lot.
→ Re-thinking the concept: maybe space for priority and/or deadline instead.

[C] Participants wished that private items could be combined with shared ones so that one can get the overall image of the things to be done. This was also mentioned in the to-do widget requirement concerning scenario where one person is having multiple to-do lists.

8.3.3.10. Bugs

The item arrangement changes unexpectedly when one item is moved from one place to another. [T] This happens at least when item is tried to move to the top (but left at second due the indication problems?).

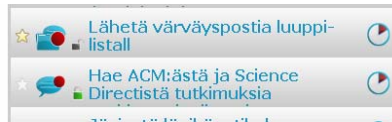


26. The third row of the heading is only partly visible.

The third row of the item heading shows only partly (Figure 26). [T]
→ If the item heading is more than two lines the rows the third row could be marked with "..." in the end of the second row.

Inconsistent terminology, e.g. tooltips "delete item" vs. "create todo note". [T]

→ Preferable terminology needs to be decided and used consistently throughout the widget.



27. Different size of red dots in items and folders.

The red dots are different sizes in folders and items (Figure 27). [T]

→ Finalizing of the iconography.



28. Closed To-do widget toolbar.

Close toolbar arrow does not change its direction when pressed (Figure 28). [T]

→ Arrow needs to change its direction when the toolbar is closed.