

Test Report

Contents

Test Report – Introduction	5
Test Plan – Template	6
Test Result – Template	11
Test Plan – Questionnaire	13
Test Result – Questionnaire	28
Test Plan – Observation	50
Test Result – Observation	57
Test Plan – Heuristic Evaluation I (version 0.1)	77
Test Result – Heuristic Evaluation I (version 0.1)	84
Test Plan – Assessment Test (version 0.2)	102
Test Result – Assessment Test (version 0.2)	110
Test plan – Heuristic Evaluation II (version 0.3)	121
Test Result – Heuristic Evaluation II (version 0.3)	128
Test Plan – Validation Test (version 1.0)	136
Test Result – Validation Test (version 1.0)	144

Test Report - Introduction

This Test report presents all the test plans and test results achieved from the different user tests and evaluations performed in the development of the inTelly system.

There has been used a different font and document template in the Test report. The main reason for using another layout is to illustrate that the Test report differs from the Main report. The tests and evaluations presented in the Test report should be seen as separate documents that have been made during the development of the system.

The first section of the Test report contains two templates for the Test plan document and the Test result document template developed in this project. This template has been used as the basis for the entire test plans and test result documents. The use of these templates has ensured a consistency in the user tests and evaluations. The templates are formulated in general terms and can be reused in other projects and/or system development.

The filled out test plans are each followed by a corresponding test result presenting the immediate results, the evaluated results, a discussion and a conclusion. The content of the Test report is:

- Test plan and Test result – template
- Test plan and test result for the:
 - Questionnaire
 - Observation test
 - Heuristic evaluation I (inTelly version 0.1)
 - Assessment test (inTelly version 0.2)
 - Heuristic evaluation II (inTelly version 0.3)
 - Validation test (inTelly version 1.0)

It should be noticed that the test plans for the two heuristic evaluations are almost identical, but they are both presented in the Test report in their full length for consistency.

The different types of tests and evaluations used are presented in detail in *Appendix C*.

Test Plan Template

Responsible:

kk, lkr, pmj

Test date:

4th of September 2000

State:

[Developing | Waiting for accept | Done]

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis | Initial | Iterative | Installation]

Test plan

This paper will describe the test plan. This will include an introduction, where it is stated why to use a test plan. The next section will describe the different parts that are considered to be necessary in a test plan. The parts are primary taken from [JR].

Introduction

The test plan should be made to have a document containing the most important aspects of the test to be performed. The document should contain subjects like the test objectives, the test methods etc. The main reason to make this document is that it is important for the value of the test, that the test has been well thought-out and described, because it gives the possibility to repeat the same test again if it becomes necessary. It also means that the design team is forced to think about the purpose of the test, which will ensure a goal-oriented test. The test plan is also a documentation of how the test is performed, which can help clarify uncertainties and problems at a later time.

The test plan contains the following items in this project [JR]:

- Purpose.
- Objective.
- User profile of test participants.
- Method.
- Accessories.
- Task list.
- Test monitor role.
- Evaluation measures.
- Report Contents and Presentation.

Furthermore our test plan will include:

- Possible outcomes of the test.
- Actions to be taken according to the possible test outcome.

The following sections will comment on the different items in the test plan.

Purpose

The test purpose should describe the purpose of the test in general terms (not very specific, but must serve as an overall purpose). The item must clarify the overall reason for performing the test.

Objective

The objective of the test is one of the most important items in the test plan. This item should hold a precise statement of the test objective. The complete test will get its focus point from this item. A test objective could be a statement like:

"Is prototype A faster to use than prototype B, when making task XX?"

It is important that the test objective is precise and it should be possible to carry out a measurement more or less directly from the test objective.

User Profile of Test Participants

In order to evaluate the product with test participants similar to the end users, it is naturally necessary to know who the end users are. If the test is performed with people who are not representative to the end users, then there will be a considerable risk that the test result will be misleading. If there is a significant difference between the end users and the test participants then it should be described and discussed in this section.

Method

This item contains the methods used in the test. There should be guidelines with all the details of the test. This includes how the test should be made in every detail. The reason to specify all test details is to ensure that the test will be the same even in situations with different test monitors.

It is important to define the goals and objectives before the test otherwise the test results will often tend to point in a direction given by the design team's thoughts.

Pilot tests are a well-suited tool in order to ensure that the test will perform well, when the test participants arrive.

By keeping the test as simple as possible it will be easy to keep the complete test consistent and easy to make equivalent tests, which is important for the quality of the test result.

The environment should be made as near the end environment as possible. This is especially important in cases where there are particular characteristics in the environment that needs to be taken into account (this could e.g. be very noisy places).

Accessories

In this section of the test plan all the necessary accessories needed for the test should be listed. This should include items like paper, pencil, computer etc. By making a somewhat complete list of test accessories minimizes the risk of forgetting something, which could decrease the yield of the test.

Task list

This section will present the subjects that should be covered by the task list.

As the name indicates the task list will include the tasks that the user has to make during the test. This will include a short presentation of each task.

It will also be a good idea to make benchmarks for the tasks in order to be able to evaluate the users during the test.

[JR] suggests putting primary focus on:

- Tasks that will be used frequently.
- Critical tasks (could be tasks where an error made during the task will have serious negative consequences).

Tasks that the design team think will be difficult to manage.

Test monitor role

The role of the test monitor should be completely described before the test. When the test monitor role is well defined the test will tend to be made in the same way every time it is performed even if the test monitor is a different person.

There should be made a description of what the test monitor should tell the test participant before, during and after the test.

Evaluation Measures

This item should describe the data collected during the test, used for the evaluation (error rates, time spend on different tasks, etc.). It is important to take the test objective into account when deciding which data to collect.

Report Contents and Presentation

In order to promote the test results it is necessary to determine how to communicate the results after the test.

Possible Outcomes of the Test

The design team should also consider what kind of result there could be expected from the test (both good and bad results). This will force the design team to take an extra look at the design before the test and this could result in adjustments in the design before the test is made. It can be discussed whether adjustments in the design just before the test is a good idea. But at first the expected results of a test can be defined early in the test preparations and second some design problems will make the test result invaluable and it will therefore be of great importance to correct them before the test.

The design team will also have to compare the possible outcomes given by this test plan item and the test objective. If the possible outcomes do not support the test objective then both items must be reconsidered.

Actions to be taken according to the possible test outcome

The considerations about the possible outcomes should include what to be done if the given result occurs. This means that the design team must have some ideas of both how the test will end-up and what to do accordingly. When the design team is prepared to handle some different outcomes, they will be better equipped to handle the resulting outcomes of the test.

Test Result Template

Responsible:

kk, lkr, pmj

Test date:

4th of September 2000

State:

[Developing | Waiting for accept | Done]

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative|Installation]

Test Result

This paper will describe the test result. The first section will contain the evaluated results followed by a discussion and a conclusion. The second section presents the immediate results, where it is possible to see how the participants handled the test.

Evaluated Results

In this section there will be an interpretation of the immediate results from the test. This could be calculations, or bringing the most important results of the test forward.

Discussion

In this section a discussion will sum up on the most important things about the test. The different results gathered from the test should be discussed, why the test ended up like it did. Is there anything that influences the test result?

Conclusion

What can be derived from the entire test. Bring forward a few points that have been the main points. These few points should be easy to remember.

Immediate Result

These immediate results are the raw data/comments, which came from the test participants during the test. This is the initial result from the test. This part often takes up all lot of pages.

Test Plan Questionnaire

Responsible:

kk, lkr, pmj

Test date:

30th of October 2000 to 6th of November 2000

State:

Done

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative|Installation]

Test Plan - Questionnaire

This questionnaire plan will take its starting point in the test plan template. The items in the test plan will not be followed strictly, but only the items relevant for the questionnaire will be considered. This means e.g. that there will not be made a task list the user should follow, because this will only contain a description of how the participant should fill out the questionnaire. It has nevertheless been chosen to use the test plan in the description of the questionnaire due to get a structure in planning process.

The content will be:

- Purpose.
- Objective.
- User profile of test participants.
- Method.
- Evaluation measure.
- Report Contents and Presentation.
- Possible outcomes of the questionnaire.
- Actions to be taken according to the possible questionnaire outcome.

There will also be a section presenting some general guidelines dealing with the creation of questionnaires. The questionnaire will be found in the last section of this test plan.

Purpose

To be able to determine what the user wants. The results will be used in the evaluation of the user opinions about using and reading news and TV-guides on the Internet. Finally the results should include some ideas and suggestions to the personalisation and integration of the news and TV-guide services.

Objective

The objective is to find out how the Internet users are using the Internet at present time concerning TV-guides, news reading and personalization on the Internet in general. The objective is also to get some user viewpoints of existing TV-guides and news services (with or without personalisation).

User Profile of Test Participants

The participants are Internet users who are familiar with the Internet (see also *Main report - 7. User definition* for more information about the definition of the user group).

The users are primary found by sending an email to fellow students at Aalborg University asking them to visit the project homepage (www.inTelly.dk) where they also can find the questionnaire. The link to the homepage is also send out to a relevant newsgroup on the Internet.

The emails are send to the 3rd, 5th, 7th and 9th semester at the Institute of Electronics and Data techniques by their group emails, e.g. 00gr301@kom.auc.dk, 00gr302@kom.auc.dk, etc.

In order to sort out possible non-advanced Internet users from the test result the following constraints have be made (if the participant does not fulfil all the items in the list he or she is not considered as an advanced Internet user).

In order to be categorised as an advanced user the participant must:

- Have an Internet connection at home.
- Use a rather new Web browser.
- Use the Internet more than twice a week.
- Have used the Internet for more than a year.
- Have used more than one of the different Internet information channels.

Method

The method used to collect the information from the users is chosen to be a questionnaire. The participants should be informed that they will be anonymous and that the data only will be used in the thesis made by the project group.

The questionnaire should contain some information about the user's personal data, like age group, level of experience on the Internet, etc. There should be a presentation of the overall idea of the project in order to put the participants in a situation where they are able to contribute with valuable ideas.

Next there will be some questions concerning TV-guides and news services in order to get an overview of which services that are popular and why this is the case. There should be room for the participants to enter critical comments and ideas in fields connected to each question. It is important that the participants are encouraged to contribute with their ideas and experience.

The questionnaire will be made as a Web page, where the participants are free to decide where and when they want to fill out the questionnaire. This is considered as an appropriate method when it is taken into account that the user group is advanced Internet users. The questionnaire will be tested before it is distributed to all the participants.

Evaluation Measures

The data collected will be of three types:

1. Comments entered in the comment fields.
2. Answers to yes/no questions.
3. Multiple-choice questions.

The different measures should be used to give the design team ideas of the needs of the user before the first draft of the TV-guide is made.

Report Contents and Presentation

The results should be presented by making statistical analysis on the answers and evaluations. The comments should be evaluated by making a discussion and conclusion of the results. It should be possible for the participants to view the final outcomes of the questionnaire at the project homepage.

Possible outcomes of the questionnaire

In the following list there will be presented two possible outcomes of the questionnaire:

1. The users will not give any personal information to make the personalisation.
2. There are many users that are not using any news services or TV-guides on the Internet.

Actions to be taken according to the possible questionnaire outcome

This section contains some suggestions to what to do if one of the outcomes listed above will be the result of the analysed questionnaire.

In the first situation the participants are sceptical about the generation of the user profile. This is a very alarming result and the design team must put focus on this subject during all development phases in order to make a user profile that will be generated without unnecessary disturb the users and also has a capability of filtering the information appropriate.

The second outcome means that the results would contain less information than expected. The idea was to collect information about the existing news services and TV-guides and some explanation about why the users are using them. If there are only a small number of users that has a favourite news service or TV-guide, then the design team has to use these, but must be careful not to put too much weight in these. The result could indicate that the existing news services and TV-guides do not fulfil the demands of a significant part of the Internet users.

Questionnaire guidelines

The guidelines given below are primary taken from [KL], but there are also added commonsense items made by the project group.

- The questions must be easy to understand.
- The formulations must be precise.
- Ask about one thing at a time.

- The questions must be unambiguous.
- The questions must be kept in general terms known by the participants.
- Guiding questions must be avoided.
- Hypothetical questions must be omitted.
- The questions must be formulated in order to give answers according to the objective of the questionnaire.
- The questionnaire must be tested in one or more pilot tests.
- The users must be anonymous and it should be stated clearly.

As already mentioned the final questionnaire will be found in the last section of this test plan document.

Questionnaire – evaluation

This section will present the evaluation of the questionnaire. The questions can be categorized in following four groups:

1. Questions concerning the use of the Internet in general.
2. Questions about the participant's use of Internet news services.
3. Questions about the participant's use of Internet TV-guides.
4. Questions concerning possible use of personalization of Internet information in general. This includes considerations about which personal data the user thinks he or she accept to send out on the Internet.

In the following section each question group will be presented in order to present how the different results will be evaluated. The reference to the questions will be Q1 for question number one and so on.

Internet in general

The questions concerning the general use of the Internet has been made in order to sort out possible non-advanced users who have answered the questionnaire. It has been chosen to eliminate possible non-advanced users because their experience with e.g. personalization could give wrong results (e.g. misunderstanding of the questions etc.) The general questions also have another purpose, which is to follow the guideline telling that a questionnaire should start with simple questions in order to make the participant feel comfortable (see also the questionnaire guideline presentation above). This means that there is included a starting question asking for the age of the test participant (Q1). And the result will therefore include an indication of the age of the participants, although they are primary expected to be in two age groups covering 15-35 years old. If there is an indication of some connection between the age of the user and some of the other questions then it will be investigated further.

The rules that determines if a participant is an advanced user or above is presented below:

1. The user must have an Internet connection at home (Q2 and Q3). If the user does not have an Internet connection at home it will be more difficult for him or her to use the TV-guide service when necessary.
2. The user must use a somewhat new Web browser (Q4). This will indicate that the user is using the Internet regularly because many homepages demands a new Web browser in order to display the pages correctly. The answers to this question could also be used to determine which Web browser the system should be optimised to.
3. The user must use the Internet more than twice a week (Q5). This means that the user must be in frequently contact with the Internet in order to be considered an advanced user.
4. The user must have used the Internet in more than a year (Q6). This could seem as a strict demand, but the relatively new users are in this project considered to be in an investigating phase concerning the Internet and is not using it as e.g. the news or TV-guide channel.
5. The user must have used more than one of the different Internet information channels (Q7). A user that is only using the Internet for one of the purposes is in this project considered to be a non-advanced user. The reason is that the user must have some idea of the many different information services available.

If the user does not fulfil the above demands he or she is eliminated from the investigation. The reason is that the remaining answers will not be considered relevant if the user is not a part of the advanced user group.

News on the Internet

This part of the questionnaire is used to give an indication of what are the characteristics of a good news provider. The participants are asked which news providers they are using and whether the information is properly and well ordered. The evaluation of these questions should give the project group some ideas of how to design the user interface that presents the information to the user.

In the following list the evaluation of the different questions concerning Internet news are presented:

- Q8: This question is used to get an idea of the distribution of the different news channels among the advanced Internet users.
- Q9: If the users never have tried to read news on the Internet, then this question is used as a jump question (to Q15).
- Q10 and Q11: In order to get an idea of the looks of a good Internet news provider the participant is asked for his or hers favourite on the Internet. This will be taken into account when the considerations about the design are made. It should be noticed that is has been taken into account that some of the

participants will say that they are using the news service because of the contents. In this case the answers will still be used in the considerations of the user interface design, but the design team should keep in mind that the user seems to prefer the given Internet page because of the contents.

- Q12: This question will help the project group in deciding which participants to put weight on. The more often the participants are reading news on the Internet the more weight there will be put on their answers concerning the news.
- Q13 and Q14: These questions will determine why the participants are reading Internet news. From the answers it will also be possible to find out where the weaknesses in the Internet news services are (if some of the possible answers get significant lower score than the others then this could indicate that this subject needs extra attention in the design).

TV-guides on the Internet

This group of questions will have the same purpose as above (News on the Internet) apart from that these questions will be related to TV-guides on the Internet.

- Q15: From this question it will be determined how the participants are using existing TV-guides.
- Q16 and Q17: These questions will determine whether the participants do not check all possibilities when looking for TV-programs. If not, the TV-guide to be developed will have its justification.
- Q18: Will work as a jump question.
- Q19 and Q20: These questions will give ideas to why the participants are using a certain TV-guide on the Internet. This will be used in the design of the TV-guide.

Personalization of Internet information

In this part the participants states which personal data he or she is willing to inform a personalization system about. There will also be questions about the participant's experiences with personalization on the Internet. The result will be used to evaluate which data that can be considered available in order to personalise the information to the user.

- Q21: This question is used as a jump question concerning possible use or not of personalised Internet services.
- Q22 and Q23: From these questions it will be possible to get an overview of what the users think of currently existing personalization on the Internet. If e.g. the participants think that it is too difficult to make the necessary user profile then this must be taken into account in the making of the user profile in the project.
- Q24: This question will be used to determine which category of personal data that it can be expected to get from the users and taken into account in the development of the personalization. The question includes some data that is

considered very personal and that no one are expected to inform. These options are included in order to make sure that the participants are paying attention in filling out the relative long list of possibilities.

- Q25: This question will be used in order to get participants for future tests. It is considered important to get interested participants for the tests, because they will typically contribute with more valuable information.
- Q26: The final question is a field where the participants enter some general comments.

The Questionnaire

In this section the questionnaire is presented. This includes the introducing text as it was presented to the visitors on the Web page.

Questionnaire

This questionnaire will be used in the development of a personalised TV-guide. This means a TV-guide that automatically finds and presents the TV-programs that you are interested in. You therefore do not have to look through the TV-program listings for 30 channels.

The results of the questionnaire will be used in the development of a master thesis at 9th and 10th semester at the specialization of Intelligent MultiMedia at Aalborg University. The purpose of the questionnaire is made to get an idea of how the Internet can be used concerning news and TV-guides. The project will contain a suggestion for presentation, integration and personalization of information on the Internet.

We deeply appeal to you to contribute with your personal comments. The number of questions is for the same reason kept to a minimum.

Everybody that answers the questionnaire will be anonymous.

Dependent on the questions it is possible to select one or more answers (circular and square boxes respectively). In the comments fields it is possible to write a maximum of 255 characters. If you feel like writing more you are more than welcome to send us an e-mail with your comments.

Thank you,
Kenneth Kammersgaard
Lars Kromann
Peter Michael Jensen
support@inTelly.dk

See: <http://www.inTelly.dk> for details about the project.

At this homepage the first results of the questionnaire will be brought at the 6th of November.

1. To which group of age do you belong?

- 25
- 26-35
- 36-45
- 45-

2. Where do you use the Internet?

- Home
- Work
- Education
- Internet Café

Other

3. If you have an Internet connection at home; which type of connection do you have?

- Analogue modem
- ISDN
- Cable modem
- ADSL
- Local network
- WAP

Other

4. Which browser do you use?

- Internet Explorer 5.0 or newer

inTelly

- Internet Explorer older than 5.0
- Netscape 4.7 or newer
- Netscape older than 4.7

Other

5. How often do you use the Internet?

- Less than once a week
- Once or twice a week
- More than twice a week

6. For how long have you used the Internet?

- In one year or more
- In less than a year

7. Which Internet information channels are you using?

- News services
- Chat
- Categorized search (e.g. Jubii, Yahoo)
- Search (e.g. Altavista, Google)
- Newsgroups
- TV-guides
- Hit list (e.g. a top 10 of the most visited homepages)

8. How do you meet your news demands?

- Newspaper
- Radio
- TV
- Internet

Other

9. Have you tried to read news on the Internet? (If not then go to question 15)

- Yes
- No

Comments

10. What is your favourite Internet news service?

- DR Online
- TV 2
- Computerworld
- Politiken
- Jyllandsposten
- CNN

Another Internet news service

11. Why is the above your favourite Internet news service?



12. How often do you read news on the Internet?

- More than once a day
- Once a day
- Weekly
- Less than weekly

13. Do you read news on the Internet when you:

- Want the latest news
- Want more background knowledge about a certain subject

inTelly

- Want news about a special area of interest
- When you happen to run in to them

Other

14. What is the reason that you sometimes use the Internet for reading news?

- It is fast to find the different news
- It is easy to find the different news
- The presentation of the news are well-arranged
- You can decide yourself when to read the news
- You can decide yourself which news to read
- You can easily get more information about the different news

Other

15. How do you find out what to see in TV?

- From newspapers
- From Text-TV
- Internet services (DR online, TV-guides, etc.)
- TV-spots
- Radio-spots
- Friends

Other

16. How many TV-channels do you check for interesting programs?

- 1-2
- 3-8
- 9-30
- more than 30

17. If you have the possibility to watch more TV-channels than the number you check for interesting programs; is this because:

- There are never interesting programs on the TV-channels you do not check
- It is rare that there are interesting programs on the TV-channels you do not check
- You do not bother to look through all the TV-channels

Other

18. Do you use the Internet to find out what to see in TV? (If not then go to question 21)

- Yes
- No

Please write down the main reason for your answer

19. What is your favourite TV-guide on the Internet?

- DR Online
- TV 2
- Billed-bladet
- TV-guiden

Another TV-guide

20. Why is the above TV-guide your favourite?

21. Do you have experience in personalization on the Internet? (Personal adaptation of an information channel, so that it only presents the information that you are interested in). If not then go to question 24.

- Yes
- No

22. Where have you used personalization of information? (You are very welcome to write down the web address).

- In connection with news reading
- In connection with TV-guides
- In connection with search
- Receiving of newsletters

Other

23. What was your experience about the personalization that you have tried? (comment on your answer)

- Removes too much information
- Good at finding relevant information
- Removes the wrong information
- It is difficult to enter the personal data
- Feels insecure by giving your personal data in order to get the information personalised

Other/Comments

24. Which of the following personal data will you give out on the Internet, if they are not used for anything else but to personalise information to you and you will remain anonymous?

- Name
- Age
- ZIP code
- Address
- Civil registration number

inTelly

- E-mail
- Interests
- Income
- Job

25. If you are interested in participating in further testing concerning this project, then please write down your email address below.

- I would like to be a test participant
- I would like to receive an email when there is a new prototype available

Email-address

26. Comments in general



Test Result Questionnaire

Responsible:

kk, lkr, pmj

Test date:

30th of October 2000 to 6th of November 2000

State:

Done

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative|Installation]

Test Result

This document will present all the results of the questionnaire test. There will first be a presentation of the evaluated results, which contains an interpretation of the answers, comments and an overall discussion of the main question groups. This is followed by a conclusion of the test. The final sections contain a graphical representation of the results (one section for the advanced users and one for the non-advanced users). The estimated number of potential participants is 350, which means that 25% (86) has filled out the questionnaire. This is believed to be a relative high percentage considered that the questionnaire has been sent out uninvited.

Evaluated Results

All the numbers below refers to the answers in the questionnaire. Only the answers from the advanced users are used in this paper (see Test plan – Questionnaire – Test participants for details). The total number of advanced users is 62, which is 72% of the total number of users. All percentages are relatively to the number of advanced participants that have filled out each question described.

The first section will describe the results of each question primary concerning the frequencies of each answer. The second section primarily presents the comments given by the participants. The third section brings together the different questions and will give an evaluation of the four question groups (Internet in general, News on the Internet, TV-guides on the Internet and Personalization of Internet information). The contents of the described sections will be based on the advanced Internet users.

It should be noticed that 24 participants are considered as non-advanced users and therefore not included in the following description. The answers from the non-advanced users are also examined for comments and tendencies and they are found to be very similar to the advanced users. The reason is that the questionnaire is send out to potential advanced users (primary students at Aalborg University) and that the non-advanced users typically fulfils the demand for being advanced users just missing one of the five criteria's e.g. only not having an Internet connection at home.

Evaluation of questions

In the list below the results of each question is commented. This list will only comment on the counts of each question.

- 1) As expected the majority of the participants are in the age group below 25 years (80%), because the questionnaire was sent out to students at the university.
- 2) All participants use the Internet at home because this was one of the criteria's for being an advanced user. There is a lot of users that uses the internet from the

educational institutions (83%). 32% uses the Internet at work. This result is also expected because most participants are students and not all of those got a job.

- 3) The usual Internet connections at home are: Analogue modem (33%), Cable modem (29%) and local network (29%). The relative many having a connection via a local network could indicate that this connection is available on many student hostels. The speed of the Internet connections is spread evenly over all available speeds.
- 4) A very great majority of the participants are using Internet Explorer 5.0 or newer (90%). Only 30% are using Netscape 4.7 or newer. This indicates that if it is necessary to chose between the two browsers in the development of the system, then the system should be optimised to Internet Explorer 5.0 or newer.
- 5) The only participants used in the evaluation are advanced users, which use the Internet more than twice a week, which means that all the answers are more than twice a week. All the advanced users are using the Internet more than twice a week.
- 6) As in the previous question all the participants have answered that they have used the Internet for more than a year. (The reason is the same as in question number 5).
- 7) The most used Internet information channels are: search (87%), categorized search (74%), newsgroups (74%) and news service (72%). 38% are using the TV-guides available on the Internet. This is considered as a surprising low number of the advanced users.
- 8) In this question the most participants are meeting their news demands via TV (95%). But the Internet is surprisingly the second most used news provider (72%) for the advanced users.
- 9) This is a jump question where 80% of the participants have answered that they have tried to read news on the Internet.
- 10) The favourite Internet news providers are: Jyllandsposten (18%), TV-2 (12%) and DR online (11%), they are evenly spread between the different news providers. Further more the participants have mentioned other Internet news providers, which can be seen in the section evaluation of comments.
- 11) This question is about why the participants prefer the above news provider. This will be presented in the section evaluating the comments.
- 12) The frequency that the participants read news on the Internet varies through the whole spectrum (from more than once a day to less than weekly), but the most answered category is "once a day" (39%).
- 13) The primary reason for reading news on the Internet is to get the latest news (74%). The other possibilities get about 50% each. This means that it seems most important to provide news that are very up to date and secondly e.g. provide more background knowledge about a certain area. Another interpretation of the answers to this question could be that the there should be put extra focus on the issues that do not get a high score. The reason should be that by putting focus on these issues would eliminate some of the weaknesses of reading news on the Internet.

14) This question asks the participants to say why they are using the Internet for reading news. All questions were considered important and all possible answers are formulated positive in order to evaluate which of the possibilities that are not usually met at the news services. Three possibilities get the significant lowest scores and they are: "It is easy to find the different news", "The presentation of the news are well-arranged" and "You can easily get more information about the different news". The project group thinks that all possibilities are important and the system to be developed will therefore put extra attention on the above three possibilities.

15) This question asks how the participants are finding out what to see in TV. Text-TV is the primary source of what to see on TV (74%). All the other get about 40% (except "radio spots" with only 1%). This could indicate that the Internet is equally used to find out what to see in TV compared to e.g. newspapers. The Text-TV is typically the only source that is available when the TV viewer is sitting in front of the TV. This indicates that the best solution would be to integrate the TV-guide in the television. This is although not a subject for this project, except when the Internet will be available through the television or online TV gets common.

16) Most participants check 3-8 TV-channels for interesting programs (70%).

17) About 80% of the participants have the possibility to watch more TV-channels than they check for interesting programs. There are various reasons for this high number, but the primary reasons are that they do not bother to look through all the TV-channels and that it is rare that there are interesting programs on the rest of the channels. This means that a TV-guide that can easily monitor all TV-programs will be useful for most of the participants that have answered this question (55%).

18) Below half of the participants actually use the Internet as a TV-guide (44%). This although seems as a significant part of the advanced users.

19) The favourite Internet TV-guide is TV-2 and TV-guiden (56% and 28% respectively). The reason is to be found in the following section where the comments are described.

20) See the following section.

21) 41% of the participants have experience in personalization on the Internet. This is more than the project group has expected, but it should be noticed that the 41% is taken from the group of advanced Internet users.

22) Most have tried to use personalisation in connection with news and newsletters (44% and 48% respectively).

23) Most of the participants think that the personalization they have tried is good at finding relevant information (62%). It should be noticed that a significant part of the participants have said that the personalization removes too much information (29%) or that it is difficult to enter the personal data (25%). This indicates that these two topics to get extra attention in the development of the inTelly system. (See also comments in the following section).

- 24) Many of the participants that have tried out personalization are willing to give out important personal data like: Age, name, interests, e-mail and job. This should be taken into account when designing the system.
- 25) 20 of the advanced users have agreed to participate in further tests and this is a very useful fact for the project group, who hopefully can keep some dialogue with some of future users during the system development.

Evaluation of comments

In this section there will be an evaluation of the comments in the questions where the participants have contributed with valuable comments.

10) Besides DR online, TV2 and Jyllandsposten, the participants primary use the following Internet news providers:

- Slashdot.org
- BT
- www.ing.dk
- Berlingske Tidende
- myCNN

11) The reasons that the participants are using their favourite Internet news provider are many. Below there is a list containing the most frequent categories:

- Fast (TV2, DR).
- Well arranged (TV2, CNN, Jyllandsposten, slashdot, Berlingske, DR).
- Have always used it (TV2, Jyllandsposten, Berlingske).
- Covering personal interests (slashdot, Jyllandsposten, BT, DR, ing, Computerworld).
- Contains the latest news (CNN, ing).
- Easy to remember the Web address (Jyllandsposten).
- Is considered objective (DR).
- Personalised news (MyCNN).
- News for free (Politiken).

14) The supplement reasons for reading news on the Internet are:

- To get local news from another country.
- Already on the Internet.
- To read about breaking news.
- The news are free.

15) Other comments about where to find out what to see on TV:

- Just zapping through the programs.
- TV-guide magazines.

17) Comments to why the participants are not checking all possible TV-channels for interesting programs:

- Prefers some given language.
- Too much to watch if all possibilities are used.
- Will not watch TV-channels where the programs are interrupted by commercials.
- All the TV-channels are not available where the participants are checking for programs.

18) Reasons for using or not using the Internet as a TV-guide:

- Easy to find out what to see (more channels available at a time).
- For free and right at hand (Internet).
- It is easy and fast.
- It is difficult and slow!

19) Other favourite TV-guides:

- www.rmas.dk (Ritzau)
- www.worldonline.dk

20) Reasons for the favourite Internet TV-guide:

- Well arranged (TV-guiden, rmas, TV2).
- Covering all TV-channels (TV-guiden, TV2).
- Always used it (TV2, TV-guiden).
- Categorized (TV-guiden).
- Fast (rmas, TV-guiden).
- Combined TV-guide and other information (TV2).
- Personalization (TV-guiden).
- Updated (TV2).
- No commercials (worldonline).
- Contain supplements to the programs (TV2).

22) Specific Web sites where personalization have been tried out:

- www.borsen.dk/trader
- www.in.gr
- www.pan.gr
- www.berlingske.dk
- www.tv-guiden.dk
- www.bt.dk/MitBT.pl
- www.pavei.com
- www.mycnn.com
- www.politiken.dk
- www.ing.dk
- www.computerworld.dk (newsletter)

23) Supplement comments from those who have tried personalization:

- Removes too little information.
- Very difficult to adapt to the user's needs.
- Often the design of the personalization changes.

Discussion

As mentioned in the section evaluated results is the questions grouped in the following subjects:

- Internet in general
- News on the Internet
- TV-guides on the Internet
- Personalization of Internet information

In the following sections these groups will be commented according to the result of the questionnaire.

Internet in general

This part of the questionnaire (question 2 to question 7 except question 3) was used to evaluate whether the participants were advanced users or not. In the final result there was 62 advanced users and 24 non-advanced users. This means that about 28% of the participants was not categorized as advanced users. This number is considered so high that it was decided to study the answers from the non-advanced users in order to see why they were not advanced users. The result was that they typically failed on only one of the parameters that determine whether they should be categorized as advanced users or not. The result was that the answers from the nearly advanced users were also considered in order to ensure that the project group did not miss any important information from the questionnaire. But it proved that the answers and comments from the non-advanced users were very close to the advanced users so the high percentage of non-advanced users is not considered a problem.

News on the Internet

Question number 8 to number 14 deals with news on the Internet. A significant high number of the advanced users are using the Internet for news reading (72% - only news on TV gets a higher score). In question number 10 and 11 the participant were asked to write down their favourite news provider on the Internet and why this is the case. The reasons were many and the project group will check up on the different news providers and find out how and why they live up to the expectations given by the participants. The primary reason for reading news on the Internet was to get the latest news. As already mentioned in question 14 the Internet news providers seems to fail to present the news well arranged and to make it clear how to get more information about certain news etc. This is a result that will be used in the system design.

TV-guides on the Internet

The most important result in the questions concerning the TV-guides is the reasons why the participants are not checking all TV-channels for interesting programs. The reason is typically that they do not bother to check more TV-channels for interesting programs. A personalised TV-guide should make it possible to look through all relevant channels when presenting the TV-programs to the user and this is a task that will be considered important in this project. Again the participants suggest some Internet TV-guides that will be investigated further in order to get inspiration to the design.

Another problem is the low percentage of advanced users that uses Internet TV-guides. A personalised TV-guide that makes it easy for users to find out what they want to see could attract new users. The TV-guide might also be accessible via a handheld or portable device making it possible to use it when watching TV, which might result in a number of new users.

Personalization of Internet information

Personalization on the Internet is only tried out by about 40% of the advanced users. This could indicate that this opportunity is not yet a common feature on the Internet. The participants that have tried personalization seems more or less satisfied but there is a tendency indicating that the personalization removes too much information and that it is difficult to enter the personal data. These two topics will therefore get extra attention in the development of the system.

Conclusion

The outcome of the questionnaire is considered useful. The project group has received a number of pros and cons of existing Internet news providers and TV-guides. The results will as already mentioned be used in the system development in order to fulfil the demands of the advanced Internet users. There has been put focus on certain subjects that seems not to be highly prioritised by the current existing news providers and TV-guides. Further more the project group has received some ideas of what is good and bad about personalization on Internet in general.

Graphical Results

In the following section there will be a graphical representation of the results from the questionnaire. The section is split into two sections, advanced and non-advanced users respectively. The graphical data presented here are meant to give a quick overview of the results.

Advanced users

The graphical representation of the results concerning the advanced users is presented below. The presentation first shows the question (the number of answers to the question is in the parenthesis). Below the question there is a table that presents the possible choices, the number of hits, a percentage value and percentage graph. All the percentage values are relatively to the number of participants that have filled out each question (seen in the parenthesis after each question).

1) To which group of age do you belong? (62)

Option	Hits	Percentage	Graf
-25	50	80%	
26-35	6	9%	
36-45	2	3%	
46-	4	6%	

2) Where do you use the Internet? (62)

Option	Hits	Percentage	Graf
Home	62	100%	
Work	20	32%	
Education	52	83%	
Internet café	1	1%	
Other	1	1%	

3) If you have an Internet connection at home; which type of connection do you have? (62)

Option	Hits	Percentage	Graf
Analogue modem	21	33%	
ISDN	3	4%	
Cable modem	18	29%	
ADSL	6	9%	
Local network	18	29%	
WAP	0	0%	
Other	1	2%	

4) Which browser do you use? (62)

Option	Hits	Percentage	Graf
Internet Explorer 5 or newer	56	90%	
Internet Explorer older than 5	1	1%	
Netscape 4.7 or newer	19	30%	
Netscape older than 4.7	0	0%	
Other	2	3%	

5) How often do you use the Internet? (62)

Option	Hits	Percentage	Graf
Less than once a week	0	0%	
Once or twice a week	0	0%	
More than twice a week	62	100%	

6) For how long have you used the Internet? (62)

Option	Hits	Percentage	Graf
In one year or more	62	100%	
In less than one year	0	0%	

7) Which Internet information channels are you using? (62)

Option	Hits	Percentage	Graf
News services	45	72%	
Chat	15	24%	
Categorized search	46	74%	
Search	54	87%	
Newsgroups	46	74%	
TV-guides	24	38%	
Hit list	6	9%	

8) How do you meet your news demands? (62)

Option	Hits	Percentage	Graf
Newspaper	34	54%	

Radio	31	50%	
TV	59	95%	
Internet	45	72%	
Other	3	5%	

9) Have you tried to read news on the Internet? (62)

Option	Hits	Percentage	Graf
Yes	50	80%	
No	12	19%	
Comments	11	18%	

10) What is your favourite Internet news service? (54)

Option	Hits	Percentage	Graf
DR Online	6	11%	
TV 2	7	12%	
Computerworld	4	7%	
Politiken	6	11%	
Jyllandsposten	10	18%	
CNN	4	7%	
Another Internet news service	17	31%	

11) Why is the news service above your favourite? (43)

Option	Hits	Percentage	Graf
Comments	43	100%	

12) How often do you read news on the Internet? (51)

Option	Hits	Percentage	Graf
More than once a day	9	17%	
Once a day	20	39%	
Weekly	12	23%	
Less than weekly	10	19%	

13) Do you read news on the Internet when you: (51)

Option	Hits	Percentage	Graf
Want the latest news	38	74%	
Want more background knowledge about a certain subject	25	49%	
Want news about a special area of interest	26	50%	
When you happen to run in to them	22	43%	
Other	1	2%	

14) What is the reason that you sometimes use the Internet for reading news? (51)

Option	Hits	Percentage	Graf
It is fast to find the different news	24	47%	
It is easy to find the different news	17	33%	
The presentation of the news are well-arranged	12	23%	
You can decide yourself when to read the news	39	76%	
You can decide yourself which news to read	33	64%	
You can easily get more information about the different news	17	33%	
Other	9	18%	

15) How do you find out what to see in TV? (62)

Option	Hits	Percentage	Graf
From newspapers	23	37%	
Text TV	46	74%	
Internet services	27	43%	
TV-spots	25	40%	
Radio-spots	1	1%	
Friends	21	33%	

Other	7	11%	
-------	---	-----	---

16) How many TV-channels do you check for interesting programs? (62)

Option	Hits	Percentage	Graf
1-2	8	12%	
3-8	44	70%	
9-30	10	16%	
more than 30	0	0%	

17) If you have the possibility to watch more TV-channels than the number you check for interesting programs; is this because: (56)

Option	Hits	Percentage	Graf
There are newer interesting programs on the TV-channels you do not check	11	19%	
It is rare that there are interesting programs on the TV-channels you do not check	28	50%	
You do not bother to look through all the TV-channels	31	55%	
Other	7	13%	

18) Do you use the Internet to find out what to see in TV? (61)

Option	Hits	Percentage	Graf
Yes	27	44%	
No	34	55%	
Please write down the main reason for your answer	11	18%	

19) What is your favourite TV-guide on the Internet? (32)

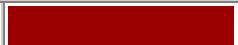
Option	Hits	Percentage	Graf
DR Online	1	3%	
TV 2	18	56%	
Billed-bladet	0	0%	

TV-guiden	9	28%	
Another TV-guide	4	13%	

20) Why is the TV-guide above your favourite? (23)

Option	Hits	Percentage	Graf
Comments	23	100%	

21) Do you have experience in personalization on the Internet? (60)

Option	Hits	Percentage	Graf
Yes	25	41%	
No	35	58%	

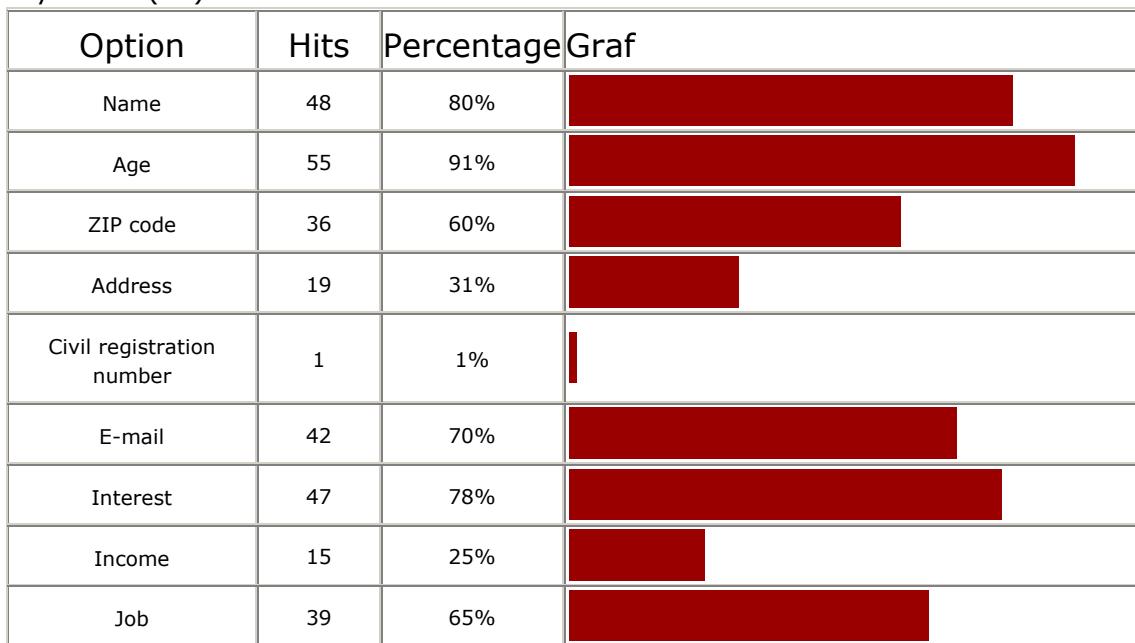
22) Where have you used personalization of information? (25)

Option	Hits	Percentage	Graf
In connection with news reading	11	44%	
In connection with TV-guides	7	28%	
In connection with search	6	24%	
Receiving of newsletters	12	48%	
Other	9	36%	

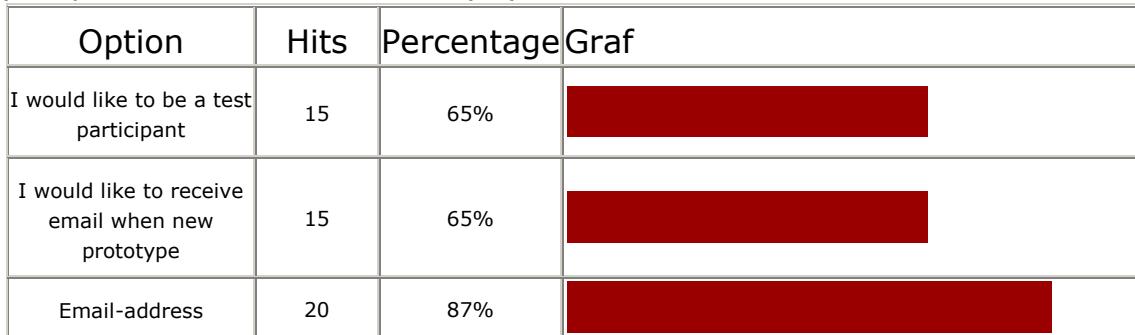
23) What was your experience about the personalization that you have tried? (24)

Option	Hits	Percentage	Graf
Removes too much information	7	29%	
Good at finding relevant information	15	62%	
Removes the wrong information	3	12%	
It is difficult to enter the personal data	6	25%	
Feels insecure by giving your data in order to get the information personalised	5	20%	
Other/Comments	13	54%	

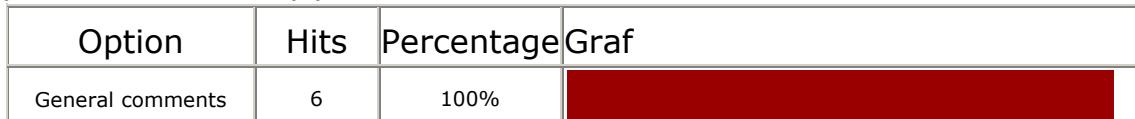
24) Which of the following personal data will you give out on the Internet, if they are not used for anything else but to personalise information to you and you will remain anonymous? (60)



25) Do you wish to receive an email (23)



26) General comments (6)



Non-advanced users

The presentation of the results concerning the non-advanced users is presented below. The ways the graphics are structured are the same as in the foregoing section.

1) To which group of age do you belong? (24)

Option	Hits	Percentage	Graf
-25	12	50%	
26-35	8	33%	
36-45	0	0%	
46-	4	16%	

2) Where do you use the Internet? (24)

Option	Hits	Percentage	Graf
Home	13	54%	
Work	9	37%	
Education	15	62%	
Internet café	0	0%	
Other	1	4%	

3) If you have an Internet connection at home; which type of connection do you have? (17)

Option	Hits	Percentage	Graf
Analogue modem	8	33%	
ISDN	2	8%	
Cable modem	3	12%	
ADSL	0	0%	
Local network	3	12%	
WAP	0	0%	
Other	1	6%	

4) Which browser do you use? (22)

Option	Hits	Percentage	Graf
Internet Explorer 5 or newer	13	59%	
Internet Explorer older than 5	0	0%	
Netscape 4.7 or newer	7	31%	
Netscape older than 4.7	4	18%	

Other	2	9%	
-------	---	----	---

5) How often do you use the Internet? (23)

Option	Hits	Percentage	Graf
Less than once a week	1	4%	
Once or twice a week	0	0%	
More than twice a week	22	95%	

6) For how long have you used the Internet? (21)

Option	Hits	Percentage	Graf
In one year or more	17	80%	
In less than one year	4	19%	

7) Which Internet information channels are you using? (23)

Option	Hits	Percentage	Graf
News services	9	39%	
Chat	4	17%	
Categorized search	21	91%	
Search	16	69%	
Newsgroups	11	47%	
TV-guides	5	21%	
Hit list	2	8%	

8) How do you meet your news demands? (22)

Option	Hits	Percentage	Graf
Newspaper	11	50%	
Radio	15	68%	
TV	18	81%	
Internet	9	40%	
Other	4	18%	

9) Have you tried to read news on the Internet? (23)

Option	Hits	Percentage	Graf
Yes	14	60%	
No	9	39%	
Comments	1	4%	

10) What is your favourite Internet news service? (14)

Option	Hits	Percentage	Graf
DR Online	2	14%	
TV 2	1	7%	
Computerworld	0	0%	
Politiken	2	14%	
Jyllandsposten	2	14%	
CNN	0	0%	
Another Internet news service	7	50%	

11) Why is the news service above your favourite? (13)

Option	Hits	Percentage	Graf
Comments	13	100%	

12) How often do you read news on the Internet? (14)

Option	Hits	Percentage	Graf
More than once a day	1	7%	
Once a day	5	35%	
Weekly	5	35%	
Less than weekly	3	21%	

13) Do you read news on the Internet when you: (14)

Option	Hits	Percentage	Graf
Want the latest news	12	85%	
Want more background knowledge about a certain subject	2	14%	

Want news about a special area of interest	6	42%	
When you happen to run in to them	6	42%	
Other	1	7%	

14) What is the reason that you sometimes use the Internet for reading news? (13)

Option	Hits	Percentage	Graf
It is fast to find the different news	3	23%	
It is easy to find the different news	4	30%	
The presentation of the news are well-arranged	4	30%	
You can decide yourself when to read the news	9	69%	
You can decide yourself which news to read	10	76%	
You can easily get more information about the different news	5	38%	
Other	4	31%	

15) How do you find out what to see in TV? (22)

Option	Hits	Percentage	Graf
From newspapers	11	50%	
Text TV	9	40%	
Internet services	3	13%	
TV-spots	11	50%	
Radio-spots	0	0%	
Friends	6	27%	
Other	6	27%	

16) How many TV-channels do you check for interesting programs? (23)

Option	Hits	Percentage	Graf
1-2	5	21%	
3-8	16	69%	

9-30	2	8%	
More than 30	0	0%	

17) If you have the possibility to watch more TV-channels than the number you check for interesting programs; is this because: (21)

Option	Hits	Percentage	Graf
There are newer interesting programs on the TV-channels you do not check	2	9%	
It is rare that there are interesting programs on the TV-channels you do not check	8	38%	
You do not bother to look through all the TV-channels	15	71%	
Other	2	10%	

18) Do you use the Internet to find out what to see in TV? (23)

Option	Hits	Percentage	Graf
Yes	5	21%	
No	18	78%	
Please write down the main reason for your answer	6	26%	

19) What is your favourite TV-guide on the Internet? (6)

Option	Hits	Percentage	Graf
DR Online	0	0%	
TV 2	4	66%	
Billed-bladet	1	16%	
TV-guiden	1	16%	
Another TV-guide	0	0%	

20) Why is the TV-guide above your favourite? (5)

Option	Hits	Percentage	Graf
Comments	5	100%	

21) Do you have experience in personalization on the Internet? (21)

Option	Hits	Percentage	Graf
Yes	8	38%	
No	13	61%	

22) Where have you used personalization of information? (8)

Option	Hits	Percentage	Graf
In connection with news reading	0	0%	
In connection with TV-guides	1	12%	
In connection with search	4	50%	
Receiving of newsletters	4	50%	
Other	4	50%	

23) What was your experience about the personalization that you have tried? (8)

Option	Hits	Percentage	Graf
Removes too much information	2	25%	
Good at finding relevant information	5	62%	
Removes the wrong information	2	25%	
It is difficult to enter the personal data	4	50%	
Feels insecure by giving your data in order to get the information personalised	1	12%	
Other/Comments	3	38%	

24) Which of the following personal data will you give out on the Internet, if they are not used for anything else but to personalise information to you and you will remain anonymous? (22)

Option	Hits	Percentage	Graf
Name	15	68%	

Age	18	81%	
ZIP code	14	63%	
Address	6	27%	
Civil registration number	0	0%	
E-mail	15	68%	
Interest	14	63%	
Income	4	18%	
Job	13	59%	

25) Do you wish to receive an email (10)

Option	Hits	Percentage	Graf
I would like to be a test participant	9	90%	
I would like to receive email when new prototype	8	80%	
Email-address	9	90%	

26) General comments (8)

Option	Hits	Percentage	Graf
General comments	8	100%	

Test Plan

Observation Test

Responsible:

kk, lkr, pmj

Test Date:

27th to 29th of November 2000

State:

Done

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative|Installation]

Test plan

In this paper there will be a presentation of the test plan concerning the observation of the users performing some given tasks, including a dialogue concerning what is good or bad about the different web services used.

The questionnaire and the observation test in combination should form the tests made in the initial design phase. The observation test will concentrate on **why** issues in order to get overall ideas and opinions from the users concerning the task of gathering information about TV-programs and news, furthermore it will be examined what the user thinks about personalisation on the Internet.

Introduction

The observation test plan contains the following items:

- Purpose
- Objective
- User profile of test participants
- Method
- Accessories
- Task list
- Test monitor role
- Evaluation measures
- Report Contents and Presentation

Furthermore our test plan will include:

- Possible outcomes of the test.
- Actions to be taken according to the possible test outcome.

The following sections will comment on the different items in the test plan.

Purpose

The purpose of the observation test is to find out which tasks and subtasks the Internet users perform when they are to find news and TV-programs on the Internet. It also has to give an idea about how people think about personalization and how they use this possibility.

Objective

The objective of the observation test is to perceive how the users find some predefined news and TV-programs given by the design team. The tasks to be solved can be seen in the section below named "Task list". The observation of the users working will make it

possible to extract which tasks and subtasks there typically are in e.g. finding news and TV-programs on the Internet. These extracted tasks and subtasks will be used when the functional requirements and testable goals will be made (see *Main report – 13. Functional requirements* and *14. Testable goals*). During the observation it will be allowed to keep a dialogue with the user in order to clarify what he or she is thinking and get some general or more specific ideas from the user. This interaction with the user should give a good idea about where and why the user gets frustrated and which things the user likes or dislikes about the news or TV-guide sites.

User Profile of Test Participants

The users performing the test will be selected from the final user group. This means that the participants will be advanced Internet users. The participants will be 5 persons from the project group and MINDPASS A/S.

Method

As already mentioned the test method will be an observation of the participants performing some given exercise on the Internet.

The following list will describe the outline of the observation:

1. The first thing to happen is to ensure that the participant recruited is an advanced user. This means that the participant is asked some questions in order to determine whether he or she is an advanced user.
2. The method will be that the test monitor asks the participant the exercises given in the Task list one at a time.
3. When the test monitor has ensured that the participant has understood the exercise, the participant should start finding the news or TV-program he or she was asked to.
4. The participant will be asked to inform the test monitor about what he or she is doing during the exercise. This will include good and/or bad things about the visited web pages etc.
5. The test monitor will write down relevant actions performed by the participant and also the utterances of the participant. This task will be rather demanding, but it will be allowed to interrupt the participant in order to get observations clarified and to get a further description of the comments uttered by the participant.
6. When the participant has completed the exercise by finding the given news or TV-program, the next exercise is given to the participant until the final exercise has been performed.
7. The test participant should be debriefed, about his experience with using the different sites.

The test monitor must make it clear that the exercises are made in order to evaluate the performance of the Internet services and not the participant. This means that the participant should not be frustrated or stressed when the news or TV-programs are hard to find. The time used to solve the exercises has shown to be approximately 30 minutes. This duration is considered appropriate, because the time is not too long so that the participant becomes impatient. It should be noticed that pilot tests of the exercises have shown that the amount of time is appropriate to solve all the exercises.

The order of the exercises will be the same for all participants.

Accessories

The accessories to be used are:

- A computer with an Internet connection and a browser installed.
- A list of tasks to be performed.
- A list of possible news and TV-guide services.
- The test environment will be an office-like room.

Task list

The tasks the participant should go through are given by the different exercises listed below. The reason for asking the questions is placed after each question.

1. Find a news story that is interesting for you. This is a simple start question and should make the participant feel comfortable due to the relative simple task.
2. Find out what you are going to see in the television tonight. This task is a bit more complicated, but it is not considered difficult for an advanced user. The question is asked in order to get inspiration from advanced users performing this type of exercise, which is considered highly relevant to this project.
3. Find out who was thrown out of the Robinson expedition last time. This is a typical exercise where the information sought is a combination of TV-guide information (Robinson is a TV-program) and news in general. It is typically not possible to find the answer in a TV-guide, so the participant is forced to look for the information elsewhere. The reason for selecting Robinson is that this program is very popular so there is a high probability that the participants know what the question is about.
4. Find out which movies that are shown on DR1, TV3, and NKR1 tonight via. www.tv-guiden.dk.

This is considered a typical task for a TV-guide user. The participant is placed in a situation where he or she "wants" to see a movie and will have to find possible subjects among the TV-channels available.

5. Imagine that you want to record the coming Sunday's movie on TV3. Your video recorder has to be programmed via a showview code, so try now to find the show

view code via. <http://www.billed-bladet.dk/tv.php3>. This type of exercise is only relevant for TV-guide users that possess a video recorder, but the number is assumed so high that the exercise is considered relevant.

6. Create a personalised account for yourself at <http://www.billed-bladet.dk/tv.php3> and adjust it to your needs/demands. This and the next exercise are asked to get the participants to create a personalised account on two of the TV-guides that provides the best personalization. The creation of a user profile is somewhat different on the two TV-guides and it is therefore interesting to see how well they are performing.
7. Create a personalised account for yourself at www.tv-guiden.dk and adjust it to your needs/demands. See the preceding question.

In the list below there will be a presentation of which questions the project group expects to get an answer to according to the list of tasks presented above (each number refers to the corresponding question above).

1. Is there used a search mechanism on any of the sites? Are the front-page used or the user going to a specific category?
2. Is a search mechanism used to scan some/all channels? Is the way the information is presented acceptable?
3. Is TV3 homepage used to find this information and is there here used a search mechanism or is there used a search engine like Altavista? Is there used a general news provider or a specific newspaper?
4. Is a general TV-guide used to find the information for all three channels or are the homepages for each specific channel used? Is there used a movie finder or is the daily program page used. Is the way the information is presented acceptable?
5. Is TV3's TV-guide used or is there used a general TV-guide to get the code. Is there used any kind of search mechanism.
6. Is the personalised account easily created? Is the TV-guide registration site easily found or is there used a search mechanism. Is there any problem in understanding what is asked one the site. Are people frightened about the amount of checkboxes presented? Is the way the information is presented when using the personalised account acceptable? Is the information filter working properly?
7. See the previous question.

Test monitor role

The role of the test monitor will be the following:

- Presentation of the superior tasks to be performed by the participant (remember to state that it is the different web sites that are evaluated and NOT the participant).
- The test monitor should ask the participant to perform the different exercises.

- When the test participant is having problems with the exercises, the test monitor will have to help the participant.
- During the test the test monitor writes down the actions made by the participant.
- The test monitor can ask the participant to comment on specific and common actions.
- The test monitor can stop a task if the participant is having serious problems.

Evaluation Measures

The evaluation measures will in this case be the comments given by the participants during the test and the tasks (or actions) they perform in order to solve the exercises.

Report Contents and Presentation

The outcome of the observation test will be a discussion of the comments and tasks performed by the participants. In the end of the discussion there will be a conclusion, where a common task analysis will be presented.

There will also be a list of good and bad things about different news, TV-guide and personalisation services.

Possible Outcomes of the Test

In this section there will be a presentation of two possible outcomes of the observation test:

1. There is no general pattern in the tasks performed by the participants during the exercises.
2. The participants have the opposite opinion about the web sites visited (e.g. some likes the facilities that some dislikes).

Actions to be taken according to the possible test outcome

In the previous section two different outcomes of the observation test was listed. In this section the actions to be taken according to the outcomes are presented. This includes both actions to be taken before and after the test in order to avoid the situations described.

1. If there is an obvious risk that the participants will solve the exercises differently there should be put extra focus on the exercises before the test. A possible solution could be to formulate the exercises more specifically (e.g. the participant should use a particular web page to solve the exercise etc.). The risk could be minimized by the use of pilot tests. If there are no patterns in the final result it

should be tried to sort out the most atypical results and see whether there are patterns in the remaining results. In the worst case it will impossible to make out any patterns from the results. In this case the concluding task list should be made from a supplement section where it is discussed why the patterns are completely different.

2. If the participants have the opposite opinions about some of the web sites it is necessary to evaluate why this is the situation. It could be necessary to determine whether the participants are disagreeing in general or is the situation unique. In order to determine which participant that has the "right" opinion is could be necessary with a discussion in the project group. A solution to a situation like the one described could be to avoid the facility that has diverging opinions in the implementation of the system. Primary because the result could indicate that the pros and cons are strongly individual and therefore should be completely avoided.

Test Result

Observation Test

Responsible: kk, lkr, pmj
Test date: 27th to 28th of November 2000
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|Iterative|Installation]

Test Result

This document will describe the test results of the observation test. The first section will contain the evaluated results followed by a discussion and a conclusion. The second section presents the immediate results, where it is possible to see how the participants solved the exercises.

Evaluated Results

In this section there will be an interpretation of the immediate results from the observation test. There will be a task analysis of the exercises that do not have any predetermined sequence and an evaluation of the given comments given during the test.

Task Analysis – Scenarios

It has been chosen just to look at exercise number 2, which is: Find out what you are going to see in the television tonight. This task is considered a typical task for users of a TV-guide on the Internet. Another reason to put focus on this exercise is because all the other exercises, except number one, have a specified URL. This is an exercise where the participant is free to choose a web site to find the information.

The approach for solving the task has been almost identical. It has mostly just been the web site which has changed. The three different web sites, which have been used to fulfil the exercise, are www.billed-bladet.dk, www.tv-guiden.dk and www.tv2.dk.

The way of solving the task itself can be split into two groups, where one of them uses a user profile and the other “filtering” facilities available for non-registered users. The two scenarios are as followed:

With user profile:

1. Figure out what web site to use to find the relevant information
2. Finding site
3. Choosing user profile
4. Scrolling through data and reading the shows presented
5. Read deepening information about shows when in doubt
6. Finding interesting program

Without user profile:

1. Figure out what web site to use to find the relevant information
2. Finding site
3. Provide information (date, channels, category) to narrow down search
4. Scrolling through data and reading the shows presented
5. Read deepening information about shows when in doubt
6. Finding interesting program

The two scenarios are very similar. The main difference is that the users with a user profile get a filtered result immediately after the login, while the users without a user profile have to make some filtering before the relevant TV-programs are presented.

Evaluation of comments

Exercise 1

All participants have solved this first exercise easily and without any problems. The only thing that really changes from participant to participant is the favourite news site that is used. The most important outcome of this exercise is that the articles that he or she has read by the user should be highlighted or removed from the list presented articles because it is annoying for the user to check/read the same news twice.

Exercise 2

This exercise was also solved without problems and this time the sites used for gathering information about the evenings TV-program was narrowed down to three different. The first we will look at is the negative sides about the used sites which are: The main point here is that there is presented too much data and information which then makes it difficult to remember what you want to see. Another thing that seems to be a problem is the channels and the way of choosing them. First there is too many channels and it is not always easy to see how many channels that are chosen. Another thing is that right now it is not possible to choose just three or four channels and the get program presented for these.

If we are looking on the most important of the good features about the used TV-guides there are things like highlighting of primetime programs and the feature "Right now" which gives you the programs that are running on the different right now. And a last nice feature is the automatically determination of the present date, so that you do not have to choose this.

Exercise 3

There is not much to say about this exercise more than it went well. The way of dealing with this problem was done in two ways. The first was just going to the TV3's homepage and lookup under "Robinson" and there you go – an answer. The second way of doing it is by first using a search engine and then just try some of the links.

There have been no comments of any kind so the task must have been easy and the web site is therefore found to be well structured.

Exercise 4

This exercise is a bit more specific than the earlier, because it is here specified which TV-guide to used for solving the problem. There are again both positive and negative comments about this site. Good things about this site is that it is easy to use and that it has the feature "Tidspunkt" (time), which divides the whole day into 24 time slots. The

more negative things about this site is that some participants mean that there are presented to many channels if just using default settings. Another thing about this list is that channels are missing which makes the user think there is no program of that kind on the channels today or gets a feeling that the system is outdated and/or unstable. The same goes if the user is checking just one channel and the user might have to select another category to see if the system is working, which is not very appropriate. It would also have been nice with a small popup text describing the links on the left-hand side. A last feature that has been missed is a sort of notepad or "shopping basket" into which you can put interesting programs and the view them when finished with the program search. The worst is that the above-mentioned feature is already build into this site, but nobody was aware of it and there were drawn no attention to it.

Exercise 5

The exercise number five was a bit more demanding than the others because of the show view code, which might not have been known by all participants, but even that did not cause any problems. The results/comments of this exercise are unlike the others because there are some opposite comments. The superior opinion about the site is that it is both simple and logic to use or difficult and not very easy figure out. Looking a bit deeper into the results gives the following positive opinions about www.billed-bladet.dk. As already mention some of the participants think that it is a very simple and logic TV-guide, that gives good and fast results and that it is better than the one they normally use. Their way of structuring/grouping the channels after its language is also seen as a good feature.

Now going to the negative features about the site we first have the selection possibilities like category and channel and so on, which are meant to be difficult to find by some of the participants and this also goes for the indication of the chosen channel(s). There also seems to be problems with the channels and their categorisation, which seems to be misguiding. The change from language based channel grouping to category-based grouping is also seen as abrupt. Also the category "film" is confusing user because it is represented both at the top and the bottom of the page. A last annoying thing is that the length of the program and the show view code is presented without any indication of what it is.

Exercise 6

All the participants solved this exercise without major difficulties but the opinion about the page with its personalisation is widespread. The overall opinion goes from OK and might use the site again to never ever using this site again. First we will look on the negative outcome/results of the test. The first that the participant's notices is that there is no indication of the possibilities of a personalised user account and they also have difficulties in finding where to signup for a personalised user account. The next problem is when creating an account the user is presented to things like "kode-husker" (code reminder) and email which purpose is not explained. There are also problems when making a mistake or using already used passwords, because this results in clearing all data fields, which is not very user friendly. The problems with the next page are the

missing description of "Rediger brugerprofil" (adjust user profile) and "Ny kanalindstiller" (new channel preferences) because the participants do not know how to proceed. Also the pre-programmed "kanal indstiller" (channel preferences) are confusing and annoying to people.

Setting up of "kanalindstilling" is generally a very negative experience because of the keyword sorting possibility is not working and the way the user has to handpick all channels from an alphabetic list, instead of using some smart grouping like Danish, Nordic and so on. Also the missing of channels and the showing all-possible and many unknown channels is irritating the participants. The way of selecting interesting categories is just irritating and frightening as the channel selection and especially the "other" and "serier" which are not described are annoying the participants. There are missing some categories or they might have been classified as other this is difficult to know.

The positive about the site is that it is not difficult to create a user account when you found out where to do it. Another good feature is that you do not have to login, but just go to the site it recognises you.

There have also been some features, which have been missed by the users, and these will be listed in the following. The users are missing the possibility for keyword sorting where keywords are highlighted in the text. It would have been nice if the channels selected in earlier "kanalindstiller" were reused when creating a new "kanalindstiller". It would also have been nice if it had been possible to select the series that you want to see, so that you do not get every program, which is considered as a series presented. At last there is the missing notepad for making notes about the programs the user wants to see.

Exercise 7

The last exercise here is like the foregoing, just based on another web site. Here it seems that people have had a better experience with generation of the user profile. In the following there will be a presentation of the outcome of this exercise. The first that seems to be an annoying factor is the unnecessary data like date of birth, sex and zip code, which is required by the system to create a user account. Another thing is that the birth of date is using an unnatural format (14041975 for 14th April 1975). Next is the selection of the channels, which here is much better than at www.billed-bladet.dk because of the grouping (Danish, Nordic, English, etc.), but still there to many channels that have to handpicked and there are also problems with missing channels. Now continuing with selecting the categories and time, where it seems that people are missing the possibility to manually specifying the timeslot, but the separation of weekdays from the weekend is found good. The problem with the categories, are that they are not very specific and therefore cause a lot of noise, but they are easily overviewed. The next is the page with the "interessegrupper" (groups of interest) and keywords. First of all a lot of the participants did not understand what was meant by "interessegrupper", because it

seemed to be the same as the category specification. There is also the choice "women" which can be understood in at least two ways and is therefore a confusing factor. The keyword feature is seen as a good feature, but there is missed the possibility of having keywords attached to every category and then also more than only three keywords.

When the participants are finished with all the selections and are trying to save the profile they are presented to a "Are you sure" dialog box, which is making the participants insecure. This box is not necessary because nothing important is happening they are just saving their user profile.

On the next page is both the user name and the password is presented to the user, in a way that is visible for everybody who might be near the user. After this page the participants proceed directly to their personalised TV-guide. Here they are presented to a page, which has a strange set-up and they do not understand. The TV-programs located in the right and left frames seems to be randomly chosen and of no interest to the user. This page also contains some structural errors like the placement of "genre" which is the categories for the user that is logged in but this facility is not together with the other profile facilities. There are also things that do not work, like the back button when the user has been checking program details and when the "√" is changed to "X" then it is not possible to change it back again. Generally the page layout is not very impressive. The participants are also missing things like a sorted and compressed overview over all interesting TV-programs and also the notepad facility for making program notes.

Discussion

In this section a discussion will sum up on the most important things about the news search, TV-guides and the creation of personalised user accounts, which are the three categories into which the observation can be split.

News search

The exercise one and three which were based on news search gave the following things to think about when presenting news. When searching for news it seems that people either use a search engine to find a news site containing the interesting news or that they just go directly to a site where they know it is possible to find interesting news.

The things that are important for the user when reading news is mainly to get the interesting news and that these are presented in a well-arranged way and also the indication of which news that have been read is an important factor.

TV-guide

The exercises two, four and five is meant to give some ideas about the things that are important to TV-guide users and to tell the project group what is good and bad about the existing TV-guides.

One of the most important things about the existing TV-guides seems to be the amount of information, which is presented to the user. This should be kept at a minimum and in a well-arranged way. This also means not to display major program descriptions and things like that, but just the most necessary things like program name, channel and time.

Many of the participants where also looking for a sort of notepad, where they could put the programs they want to see, so that they do not have to remember them when browsing for further interesting programs.

The participants also requested the possibility of selecting a few channels, instead of one or all channels, and this without having to create a personalised account. They would also like to have a more flexible time selection, instead of the one where they only have the possibility selecting between four partitions into which the day have been split.

A thing that is important is, when using the category search and there is found no programs, this should be stated clearly to the user.

Creation of personalised user accounts

Exercise number six and seven should give some ideas of how to create better user interfaces and make it easier for the user to create their user account.

First of all when you want people to create a personalised user account it should be visible for them where to do this. All the way through the creation process there should also be clear exits and clearly stated where to go and what to do next. As well as it at any time should be possible to stop with the set-up of the user profile and then use current settings. When the user is making a mistake it should also not be necessary to start all over again.

The data from the user should be kept at a minimum. This means only ask them about necessary data like username and password. When asking for additional information it should be stated what this information is used for.

The selection of the channels should be easy and fast and the channels grouped logic way, like the language or type of programs they show. People would like both the possibility of selecting whole channel groups and of handpicking channels. There are all also a demand for the possibility of adding missing channels in some way. The selection of the categories should as well as the channels be easy and fast overviewed and set-up

and also grouped in a logic way. There is also a request for keywords that are attached to each category, to make the sorting as good as possible.

If there is a possibility for several set-up files it should be possible to reuse things like the channel configuration from the earlier files.

Conclusion

This observation test can be considered as both useful and successful. The project group has received a number of good and bad things about the existing services (primary concerning the two best Danish Internet TV-guides).

The result of this test will be used later on in the system development and will help to fulfil the demands of the advanced users.

Immediate Result

These immediate results are the raw data/comments, which came from the test participants during the test. The numbers of test participants for this test was five and they are all listed below.

Test person: #1

Test monitor: Peter Michael Jensen

Date: 27th of November 2000

Exercise 1

- 1) Finds www.computerworld.dk
- 2) Reads a news story at the front page

Exercise 2

- 1) Finding www.billed-bladet.dk/tv.php3
- 2) Hitting already created profile
- 3) Scrolling down in the list
- 4) Long list of 80 programs (view all selected)
- 5) Looking at a program details
- 6) Going back
- 7) Scrolling again
- 8) Another program details
- 9) Back
- 10) Scroll
- 11) Found a film.

12) Have forgotten what to see and when (there were different programs). Did they overlap, what channel?, too much information (do not need showview) etc. (not sure if all program possibilities are present, did I miss something?)

Exercise 3

- 1) Go to www.tv3.dk
- 2) Click's Robinson short cut
- 3) Reads the page
- 4) Short cut to "Seneste nyt" (latest news)
- 5) Reads a lot of text
- 6) The person is found

Exercise 4

- 1) Goes to www.billed-bladet.dk (wrong).
- 2) Goes to www.tv-guiden.dk
- 3) Selects Film icon
- 4) All details about films on the different channels are available
- 5) Problems about detecting NRK-logo
- 6) Very easy task on the web page, but too many channels shown

Exercise 5

- 1) Goes to www.billed-bladet.dk/tv.php3
- 2) Selects TV3
- 3) Selects Sunday
- 4) No programs under TV3!!!!
- 5) Difficult to find out where to select film
- 6) Hits "genre"
- 7) Hits "film"
- 8) Still no programs!!!
- 9) To much to select between
- 10) Mission impossible webpage not updated so it is impossible to find the TV-programs for Sunday.

Exercise 6

- 1) Finds www.billed-bladet.dk/tv.php3
- 2) Click on "indstillinger" (preferences)
- 3) Click on "ok"
- 4) Enters personal data (do not understand "kode-husker" (code reminder))
- 5) Layout disappears because name was taken and no exit! He feels trapped
- 6) Hits back

- 7) All data erased....
- 8) Enters the data once again...
- 9) Same problems again...
- 10) Test99 was not taken...
- 11) Mysterious icons ("taxa" and "fodbold" (football))
- 12) Click on "taxa"
- 13) Big mystery about "taxa" (show under "kanalindstillingsnavn" (channel preferences name))
- 14) "Taxa" is now key word and can not be altered unless expert option selected
- 15) Only name of profile is altered
- 16) Looking for "ekspert" option
- 17) Selects "ret I kanalindstilling"
- 18) Gets only back to previous page (no "ekspert")
- 19) Shortcut "tilbage" (back)
- 20) Clicks on "Personlige indstillinger" (personal preferences)
- 21) Hits "Rediger din brugerprofil" (adjust your user profile)
- 22) Selects "ekspert"
- 23) Hits "Ret brugerprofil" (adjust user profile)
- 24) Hits "Ny kanal indstilling" (new channel configuration)
- 25) Enters new name for shortcut profile
- 26) Difficult to select relevant channels
- 27) Easy to select the different categories
- 28) Clicks on "fjernsyn" (TV) to see the result
- 29) The profile is finished and the result is ok

Exercise 7

- 1) Goes to www.tv-quidens.dk
- 2) Clicks on very small icon named "opsætning" (preferences)
- 3) Enters data
- 4) Enters more data (necessary data) data of birth is strange format
- 5) Clicks "ok"
- 6) Easy to select channels (good categorised channels, dk, uk, etc)
- 7) Good shortcuts to unselect/select each category.
- 8) Moves on
- 9) Interests selects (good overview)
- 10) Enters more info
- 11) Popup "are you sure?" (Unnecessary)
- 12) Info about the making of the profile
- 13) Good to enter "login"

Test person: #2

Test monitor: Peter Michael Jensen

Exercise 1

1. Finding www.sonofon.dk/MIPortal
2. Finding "Nyheder" (News)
3. World Radar is also good (www.worldradar.dk)
4. Deselects "Ekstrabladet" (sport) and "Børsen".
5. Does not remember which articles that have been read. Maybe not present these articles again or at least mark these.
6. Chooses "I dag" (today).
7. Chooses an IT-news

Exercise 2

1. Goes to www.tv-guiden.dk
2. Selects "TV-guiden"
3. Selects box containing the current channel.
4. Uses arrow down to change channel.
5. When selecting all channels nothing is visible/shown.
6. Does not believe that programs are sorted and put into the right categories.
7. Likes the feature "Lige nu" (right now) which shows what is shown on the different channels right now.
8. The highlighting of the primetime programs is a good feature.

Exercise 3

1. Goes to www.bt.dk
2. Finds "nyheder" (news)
3. Searches for "Robinson + 27." (27. is the date of the latest Robinson show)
4. Gives no result
5. Searches for "Robinson" only
6. Gets the answers from an article

Exercise 4

1. Goes to www.tv-guiden.dk
2. Do not use username
3. Selects type "film"
4. Selects the different channels one by one.
5. There is no film at DR1 and the participant selects type = "all" to be sure about the result

Exercise 5

1. Goes to www.billed-bladet.dk
2. Selects Sunday
3. Selects channel (TV3)
4. Thinks that the channel indication is hidden (difficult to find)
5. Scrolls down to find the Sunday movie and finds it.
6. Selects details about the film
7. Prefers to keep the old page when selecting the details (gets a complete new page instead).
8. Thinks that the categorisation of the channels are misguiding.
9. The (tiny) mark that indicates the selected channel disappears after the above navigation!?!?

Exercise 6

1. Goes to www.billed-bladet.dk
2. Selects "indstillinger" (preferences)
3. Makes a new account.
4. Is confused about the default profiles named "taxa" and "football" (mentioned earlier).
5. Is confused about the term "resten af dagen" (the rest of the day). What does this refer to when he enters the personalised TV-guide?
6. Bad grouping of the channels (it is unnecessary to display all possible channels).
7. The lines separating the channels do not seem to indicate any categorization.
8. Many unknown channels are shown.
9. TV-Danmark 1 is completely missing
10. What are the categories "andre" (others) and "øvrige" (something else)
11. Save the account data
12. Are not sure what to do now
13. Finds the account
14. And finds finally out to delete the default accounts (taxa and football)
15. Select a channel number (is very hard to see – too small font)
16. Thinks that it was easy enough to make a user account (has experience with www.tv-guiden.dk)
17. Insecure about the works of the personalization (might miss something)
18. Usually it will not be necessary to log in (too time consuming!)

Exercise 7

1. Goes to www.tv-guiden.dk
2. Selects "opsætning" (preferences) and "ny bruger" (new user).
3. Thinks that they ask for sensitive and non-relevant information (birth, etc.) and not explaining what they are used for!
4. Better overview of the channels (compared to www.billed-bladet.dk)

5. Misses a description of each channel.
6. Will not de-select children programs because there is a risk that there is some programs that could be missed.
7. No freedom in the selecting of time of day.
8. "Interest groups" are very similar to "genre".
9. "Women" can be understood in at minimum two ways.
10. Do not bother to enter some keywords
11. "Are you sure" – dialogue box makes the participant insecure. Is there some very important happening here? No just saving the user account data!!
12. Shows the password!!!!
13. Login-possibility is misleading because it is not a traditional login
14. What is shown after the login???? To tables of TV-programs but no description at all!!!!!!
15. "Genre" is misplaced on the user interface
16. Back button does not work when the details are selected

Test person: #3

Test monitor: Peter Michael Jensen

Date: 28th of November 2000

Exercise 1

1. Finds www.ekstrabladet.dk
2. Uses Ekstrabladet because they have fast server, but might have found more interesting news at other sites
3. There is nearly always a somewhat interesting story there
4. Reads headlines

Exercise 2

1. Finds www.billed-bladet.dk and clicks on "TV-guiden".
2. Annoying that you first have to choose "indstillinger" (preferences) and then "aften" (evening).
3. Scrolling through 169 programs, which is too many programs
4. It is a bit weird that the evening starts at 16.30
5. Chooses a program for the evening.

Exercise 3

1. Goes to www.google.com
2. Searching for "robinson and dk"
3. Choosing the link www.robinson.tv3.dk
4. Choosing "Seneste nyt" (latest news)

5. Finds the answer.

Exercise 4

1. Goes to www.tv-guiden.dk
2. Choosing "Alle film" (all films), "Station" (channel), "19.00-00.00"
3. The time section is weird.
4. Overview over all movies on the different channels
5. Insecure about why TV-3 is not shown on the channels list (is there no movies or?)
6. Missing pop-up for icons
7. "Tidspunkt" (time of day) is a good facility, like that the day is split into hours

Exercise 5

1. Finds www.billed-bladet.dk / "TV-guiden"
2. Choosing TV3
3. Choosing "Søndag" (Sunday)
4. Choosing "Aften" (evening)
5. Choosing "Film" (movie)
6. Very fast – only one movie left to choose - show view code is directly shown, but without no indication that it is actually the show view code

Exercise 6

1. Finds www.billed-bladet.dk / "TV-guiden"
2. Chooses "Indstillinger" (preferences)
3. It is not very clear that there are any possibilities for personalisation and where to find it.
4. Choosing "OK" for creating new user
5. Keying in profile data (easy for him)
6. Choosing "Rediger profil" (adjust user profile)
7. Says no to both expert questions
8. Choosing channels
9. There are too many channels
10. Choosing categories
11. Wondering about the subcategories under "Andre" (other) and "Serier" (serial)
12. Saves changes

Exercise 7

1. Goes to www.tv-guiden.dk
2. Choosing "Min TV-guide" (My TV-guide)
3. Keying in data for the profile

4. Wrong birth date – keys in 140475 instead of 14041975
5. Choosing channels
6. This feature is a bit better sorted than at www.billed-bladet.dk but still to many channels to choose. It is difficult but has to be done.
7. Choosing interest.
8. Wondering what "andre" (others) is covering over
9. The time selection is OK here. It is good that the week is split into workdays and weekend
10. Saves profile
11. Choosing "Login"
12. No description of what is seen on right and left side
13. If the checkmark change to "X" it is not possible to change it back. The goes for the other way around
14. Poorly selection of programs. As if it just picks some programs

Test person: #4

Test monitor: Lars Kromann

Date: 28th of November 2000

Exercise 1

1. Starts at www.dr.dk
2. Wants to read about major bank robbery but can not find anything
3. Disappointed because he thinks it should be on the first page
4. Goes to Index and chooses "Nyheder"
5. Checks "Alle nyheder" and "Indland" but nothing about robbery
6. He gives up and reads something else

Exercise 2

1. Goes to www.billed-bladet.dk / "TV-Guiden"
2. Chooses "Aften" (evening), "I dag" (today) and "Alle genre" (all genres)
3. Insecure about how many channels that are chosen.
4. Missing that the channels within the different groups are offered as an option, which gives the possibility to choose among them.
5. Scrolling through the pages.
6. When he changes program group has to make a new time selection or select a specific channel to update the screen/site. This is seen as an annoying factor.

Exercise 3

1. Goes directly to www.robinson2000.dk
2. Reads the name directly from the first page.

Exercise 4

1. Goes to www.tv-guiden.dk
2. Because someone has used this machine earlier and has been using www.tv-guiden.dk he is still logged on. Annoying and weird.
3. Logs off.
4. Chooses "TV-guiden"
5. Chooses channel, category, date and time (DR1 (channel), film, Tuesday the 28th of November, 19.00-00.00)
6. This seems as a good sorting of the programs.
7. Chooses TV3, where nothing then is shown and he gets suspicious and chooses "Alle genrer" to see if the system works.
8. Exactly the same happens with NRK1.
9. He is now finished with checking the programs, but now he cannot remember what he wanted to see. He is missing a "shopping basket"/notepad

Exercise 5

1. Goes to www.billed-bladet.dk / "TV-Giden"
2. Chooses "Søndag" (Sunday), "Aften" (Evening), "Film" and TV-3.
3. Gets immediately the movies for that evening presented.
4. It is easy, simple and logic.

Exercise 6

1. Goes to www.billed-bladet.dk / "TV-Giden"
2. Chooses "Indstillinger" (preferences) but it does not seem logic that the user profile creation starts here.
3. Says OK to new profile.
4. Keying in user profile data – he makes a mistakes which result in an clearing in all fields for user data – not very good.
5. Wondering why they want email address and writes a dummy address so that he does not get spamed.
6. Does not understand "Rediger burgerprofil" (adjust user profile) and "Ny kanal indstillinger" (new preferences for channels)
7. Guessing that he has to choose "Ny kanal indstillinger" (new channel preferences) to get on with the exercise.
8. Does not understand why the possibility to key in the keywords is not there and why he has to go back to expert settings in "Rediger brugerprofil" (adjust user profile).
9. The keywords do not work even if all expert possibilities are chosen.
10. Scrolling through channels and wondering about how they are sorted/grouped
11. Thinks that the grouping of the channel is so poor.

12. Missing channels like TV2 Zulu.
13. Does only choose a few channels to test personalisation.
14. The category groupings are not logic and tiring.
15. Does not know what "andre" (other) is covering over
16. Would never ever use the personalisation at www.billed-bladet.dk again because it is irritating and very poor

Exercise 7

1. Goes to www.tv-guiden.dk
2. Chooses "Min TV-guide" (My TV-guide).
3. Annoying that they want to know date of birth, sex and zip code and that they do not tell what they are using it for as they do where they ask for the email address
4. The way of typing in date of birth is weird
5. The grouping of the channels is OK, but still annoying that they have to be hand picked
6. Missing channels (VH1)
7. Does not understand what is meant with ""interessegrupper" (groups of interest).
8. OK with keywords
9. Done with setting up profile.
10. It is a problem that the interest is giving a lot of noise data. He would like just to see when the Simpsons are on but he gets all series.
11. Would like keywords on every category.
12. The timeslots into which they have split the day are bad. Manual specification would be better
13. Does not understand what is presented to him on the right and left side when using the personal TV-guide.
14. Test participant gets angry and breaks off the test.

Test person: #5

Test monitor: Lars Kromann

Date: 28th of November 2000

Exercise 1

- 1) Goes to www.jp.dk
- 2) Checks "Indland", "Udland" (at home and abroad), "Kultur" (culture), "Erhverv" (industry) and "Børs og finans" (finance) but does not find anything interesting. Goes back to front page and chooses a robbery article.

Exercise 2

- 1) Uses www.tv2.dk

- 2) Is always using this site, has it bookmarked.
- 3) Not only for tv2 but also for many channels.
- 4) Chooses "film" in the menu TV.
- 5) Chooses "TV-guide" and then "alle kanaler" (all channels), "18.00-24.00" and "film".
- 6) Site will automatically determine the day.
- 7) Shows nothing – seems that server has gone down.

Exercise 3

- 1) Tries first www.tv3.dk, but server is not responding
- 2) Decides then to use MINDPASS A/S own search engine <http://search.mindpass.com>
- 3) Searches on "Ørød" (island council) and gets some links.
- 4) Chooses www.robinson.tv3.dk
- 5) Amused over the good result of the search engine

Exercise 4

- 1) Uses www.tv-guiden.dk
- 2) Would have used www.tv2.dk if he had the possibility. Did not know www.tv-guiden.dk
- 3) Chooses genre - film and then "dato" (date), "station" (channel), "alle film" (all films), "hele dagen" (all day) is then preset.
- 4) Scrolls through the films.
- 5) Could be nice if he had had a calendar or notepad to save choices.

Exercise 5

- 1) Goes to www.billed-bladet.dk / "TV-guiden"
- 2) Chooses "Film", "Aften" (evening) and "dato" (date)
- 3) Is insecure about the meaning of the numbers listed below all movies but thinks that it is the length of the movie and the show view code
- 4) Thinks it is better than www.tv2.dk
- 5) Likes the grouping of the channels "danske" (Danish), "nordiske" (Nordic), "tyske" (German)....
- 6) When changing from language based grouping to category based grouping of channels this is done very abrupt
- 7) Confusing with the category "film" in both the top and bottom of the page. What is the difference

Exercise 6

- 1) Goes to www.billed-bladet.dk / "TV-guiden"

- 2) Can not immediately find where to create a new user account.
- 3) The possibility of the personalisation not visible.
- 4) Says OK to create account.
- 5) Keying in profile data.
- 6) Does not understand page with "Rediger burgerprofil" (Adjust user profile) and "Ny kanal indstillinger" (new channel preferences)
- 7) Irritated about the two pre-programmed channel settings ("taxa", "fodbold").
- 8) The way of choosing channels is annoying. To many channels that has to be handpicked.
- 9) Would be nice with a grouping of Danish channels, Nordic channels and so on instead of an alphabetic listening.
- 10) The category choosing is frightening. He gives up the further setting up and saves the profile.
- 11) Very hard to use but could be used.
- 12) Category "profilgenre" (genre of profile) is not intuitive
- 13) Adds new "kanal indstilling" (channel preference).
- 14) Missing the possibility to choose time interval.
- 15) Would have been nice if the channels from the earlier "kanal instillinger" (channel preferences) could have been reused so that you do not have to choose them again. Annoying that this is not possible.
- 16) When selecting series it would have been nice if it had been possible to choose the few once that he wanted to see.
- 17) Missing sailing under the category sport – so he chooses "andet" (other)
- 18) Saves "kanal indstillinger" (channel preferences)
- 19) Testing the new profile and finds that sport gives way to much noise data that he do not want so he deselects sport again.
- 20) Too much text when presenting the program.
- 21) Would like to have the type of sport presented in the headline for the program when it is presented.
- 22) Might maybe use this TV-guide again but does still think that it is very hard to use and that it has its irritating moments during set-up.
- 23) Is missing some sort of keyword sorting, where these then where highlighted in the program text.

Exercise 7

- 1) Goes to www.tv-guiden.dk
- 2) Chooses "Min tv-guide" (My TV-guide).
- 3) Filling out user data but is annoyed about the extra data they want to know.
- 4) Likes the way the programs are grouped.
- 5) Much faster, better and easier than at www.billed-bladet.dk.
- 6) Category set-up OK.
- 7) Would have like more keywords or better keywords for each category.
- 8) Saves profile.

- 9) Likes the small rotating visualisations of the category.
- 10) Missing a sorted and compressed view of the possible programs.
- 11) Generally that page set-up is awful.
- 12) Does not know what is presented in the right and left side of the page.
- 13) Thinks that www.billed-bladet.dk is more attractive/inviting than www.tv-guiden.dk but the way of setting up/choosing the channels is better at www.tv-guiden.dk.

Test Plan

Heuristic Evaluation I

(version 0.1)

Responsible: kk, lkr, pmj
Date: 27th to 28th of November 2000
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|Iterative|Installation]

Heuristic evaluation

This paper presents the test plan for the heuristic evaluation. The heuristic evaluation is based on [JN1, p. 115-163], [JN2] and [JN3].

Although there is a difference between a user evaluation and a user test [JN1, p. 157] it has been decided to use the same test plan template for both heuristic evaluations and tests. The main difference is that in a test situation the test participant does not evaluate e.g. the user interface directly but the observer or test monitor interprets the participants' responses/actions. In an evaluating situation the participant is responsible for the analysis of e.g. the user interface.

The above differences do not result in any problems in the making of the evaluation plan because the template is usable in both situations due to the specification of every detail in the test/evaluation (this includes the role of the test monitor, if present).

Introduction

This evaluation plan contains all the aspects of the heuristic evaluation. The usability heuristics used in this evaluation are presented in detail in *Appendix B*.

The evaluation plan contains the following items (no changes from the test plan template):

- Purpose
- Objective
- User profile of test participants
- Method
- Accessories
- Task list
- Test monitor role
- Evaluation measures
- Report Contents and Presentation

Furthermore our test plan will include:

- Possible outcomes of the test
- Actions to be taken according to the possible test outcome

The following sections will comment on the different items in the evaluation plan.

Purpose

The purpose of the heuristic evaluation is to evaluate a number of user interfaces due to some usability heuristics. The reason to perform the evaluation is to clarify how some user interface mock-ups are performing seen from a usability point of view.

Objective

The objective of the first evaluation is to determine how the user interface mock-ups are performing according to the usability heuristics given by [JN1, p. 115-155]. The user interface mock-ups are made on the basis of:

- The results of the exploratory tests performed in the initial design phase (questionnaire and observation test).
- Project group ideas.
- Inspiration from existing products (see also *Appendix J*).
- Ideas given by other people (e.g. MINDPASS A/S employees, supervisor etc.)

User Profile of Test Participants

The test participants used in the heuristic evaluation could in principle come from any of the following categories (used by [JN1, p. 161]):

- Novices with computer knowledge in general, but no special knowledge about usability aspects.
- Single experts who are usability experts, but not specialized in the domain of the interface.
- Double experts who both are usability experts and experts in the kind of interface being evaluated.

The results differ dependent upon which type of the groups that is used for the evaluation. According to [JN1, p. 161] will double experts be best at finding usability problems, while novices find the smallest number of usability problems. This indicates that double experts are to be preferred in heuristic evaluations. In this heuristic evaluation the participants are taken from the "single experts"-group and the "double experts"-group.

The number of participant should be between 3 and 5 due to some cost benefit calculations made in: [JN3]. In Figure 1 the number of usability problems as a function of number of evaluators is presented. It can be seen that the curve flattens out when there is more evaluators than 5 and a number below 3 results in a significant lower number of usability problems found.

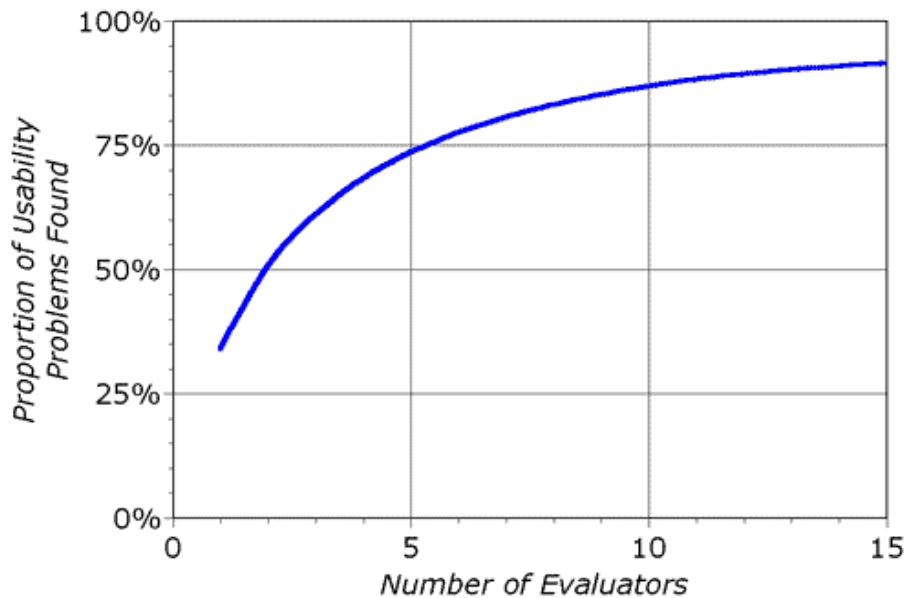


Figure 1: The number of usability problems found as a function of the number of evaluators [JN3].

In this heuristic evaluation it has been decided to use the group members as evaluators (3 persons). This means that the evaluation can be performed efficiently (no recruiting of external evaluators). The solution is considered to be acceptable due to the fact that there will be used usability experts from MINDPASS A/S in the second heuristic evaluation. The validity of the use of the group members implies that the members are considered to be "single" or "double experts", which the project group has taken the liberty to estimate to be true.

Method

Jacob Nielsen originally presents the evaluation method used. The method is used to determine usability problems in a user interface so that they can be attended to as part of an iterative design process [JN1, p. 155].

The heuristic evaluation is performed by letting a number of participants evaluate a number of user interfaces according to the usability heuristics given below (described in detail in *Appendix B*):

- Simple and natural dialogue
- Speak the users' language
- Minimize the users' memory load
- Consistency
- Feedback
- Clearly marked exits
- Shortcuts

- Good error messages
- Prevent errors
- Help and documentation

The participants are evaluating the user interfaces individually. The results can be collected in several ways (written by the evaluator, made by an observer etc.) In this evaluation it has been chosen to let the participants make a written document as their contribution to the evaluation. This gives the participants the freedom to decide when they want to perform the evaluation (in a given period). To ensure that the documents are somewhat easy to evaluate and to ensure that the participants remembers all the usability heuristics it has been decided to make an evaluation scheme to be filled out by the participant (one scheme for each user interface). An extract of a filled-out scheme is given below as an example:

Usability heuristic	Comments
Simple and natural dialogue	#1 The sequence of tasks to perform when opening a file is too difficult. #2 There are too many colours on the user interface (makes it confusing). #3 ... #4
Speak the users' language	#1 I don not understand the word "inTelly". #2 ...

When the schemes are filled-out the design team makes a new scheme containing all the usability problems found by the participants. The next thing to be done is to make a prioritising among all the usability problems. The reason is that it might not be possible to correct all the problems listed therefore it is necessary to have an idea of which problems are considered as major problems and which are minor problems. The method used to determine the rating between the usability problems is based on a "Severity Rating" given by [JN3]. The rating is done by sending out all the usability problems to not less than three and not more than five of the original participants, asking them to evaluate each usability problem according to the scale below:

- 0 = I do not agree that this is a usability problem at all.
- 1 = Cosmetic problem only: need not to be fixed unless extra time is available on the project.
- 2 = Minor usability problem: fixing this should be given low priority.
- 3 = Major usability problem: important to fix, so should be given high priority.
- 4 = Usability catastrophe: imperative to fix this before product can be released.

The severity of a usability problem is a combination of three factors [JN3], which should be taken into account by the evaluation participants:

- The **frequency** with which the problem occurs: Is it common or rare?

- The **impact** of the problem if it occurs: Will it be easy or difficult for the users to overcome?
- The **persistence** of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

When the ratings are returned from all the participants the mean is found for each usability problem and they are sorted accordingly. It is now possible to determine which problems that are considered most important to fix etc.

The user interfaces to be evaluated are a combination of WebPages, GUI shells etc. All the user interfaces considered in this evaluation are presented in *Supplement M*.

Accessories

The accessories to be used in this evaluation are all the user interfaces to be considered. Each participant should be given the user interfaces to be evaluated and the same number of heuristic schemes to be filled out. Every user interface should be given a unique identification, which is also used to mark each scheme in order to simplify the collection of the usability problems. Further more the description of the usability heuristics is given to each participant (see also *Appendix B*).

Task list

The tasks to be performed are not given by this evaluation plan, because the tasks to be considered only consist of filling out some heuristic schemes.

Test monitor role

The role of the test monitor is to handout the user interfaces and the evaluation schemes to the participants. This will include explaining that there belongs a specific scheme to each user interface.

Evaluation Measures

The evaluation measures will as already described be the severity ratings given by the participants. As a starting point all usability problems with a rating of 3.0 or higher should be corrected.

Report Contents and Presentation

The presentation of the evaluation result will be a list of the means of the usability problems ordered with the major problems at the top of the list.

Possible Outcomes of the Test

The following outcomes are considered in this and the following section:

- The final list of usability problems is very long and indicates several severe problems.
- The list contains usability problems that, if corrected, could result in new usability problems.

Actions to be taken according to the possible test outcome

The actions to be taken according to the above listed results are respectively:

- If the list of usability problems is very extensive it might be necessary to increase the rating indicating that a problem should be corrected (default set to a mean of 3.0). This will be acceptable in the first heuristic evaluation due to the fact that it will be possible to catch the minor usability problems later in the design phase. Furthermore it is possible that the correction of the major problems will result in a more or less automatic correction of some of the minor problems. If the second heuristic evaluation shows severe problems with the design there must be made some decisions about whether it is possible to redesign the user interfaces giving the problems and/or the validation test should be carried out with the problems indicated by the second evaluation. This will strongly depend on the severity of the problems and time available to correct the problems.
- If there are potential new usability problems in the correction of some of the problems in the list it should first of all be considered how to avoid this problem. If the potential problem is considered to be critical, a total change of the UI is a possible solution (acceptable in the first heuristic evaluation, but if present in the second evaluation it might not be possible to redesign the user interface due to time constraints). If it is impossible to avoid some potential new usability problem it should be evaluated whether this problem will be less significant compared to the problem corrected. And if this is the situation it should be corrected anyway.

Test Result

Heuristic evaluation I

(version 0.1)

Responsible:

kk, lkr, pmj

Test date:

27th to 28th of November 2000

State:

Done

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative**|Installation]**

Test Result

In this document there will be a description of the results of the first heuristic evaluation. The first section will contain an evaluation of the results followed by a discussion of some the most severe usability problems and at last there is a conclusion on the whole evaluation and its results.

Evaluated Results

The evaluated results will be an interpretation of the immediate results of the rated usability problems. There will be presented a description of the problems that are assumed to be the most important and annoying to the user.

The whole list of problems will be used in the next iteration of the design, but this section will concentrate on the major problems, to draw out the most severe problems in the user interface, to focus the attention on these problems. The major problems are the ones rated with an average of 3 or higher (see also the Test plan for the Heuristic Evaluation). It should be noticed that the test participants agree quite much on the rating of the usability problems, this indicates that the evaluation result is quite thrust worthy.

Each answer in the evaluation is marked with which user interface component it belongs to (see table 1). This shows clearly that the user interface with most usability problems is the Evaluation of programs component (it got the top four rankings). Each user interface component is listed with the evaluated comments. The components are listed by the rating of the comments.

Evaluation of programs Component:

- The evaluation type used, should be the same all places (consistent).
- Each degree of evaluation should be indicated in some way, e.g. icons, colours.
- Pop-up description of degrees (icons).
- There should be immediate feedback when a program is evaluated.
- There should be a discussion of how many degrees of evaluation the system should be able to handle.

List Component:

- It should be indicated which programs that are in the notepad.

Notepad Component:

- It should be possible on the front page to see which programs are in the notepad.
- It should be discussed how the different degrees of evaluation in the notepad should be presented.

Help User Interface:

- The user do not know why some programs are hidden.

Description of a program Component:

- The use of the word "Omtale" is not consistent.
- It is not clear how to select which description is shown in the frame.
- It should be possible to evaluate the program in the detailed description.

User profile User Interface:

- There should be a warning if the user moves to the next page without saving the data entered.
- It should be discussed if the keywords should be divided into the categories.
- It should be possible to specify the categories by entering some associated keywords.
- The number of Repeaters presented to user should be kept at an appropriate level.
- There is missing a preview function on the page where the columns in the List component are specified.

General:

- When the system is working for more than a second, this should be indicated in some way.
- Maybe there should be some help/indication of the functions that are not visible to the user.

Easy corrections:

- "Fuzzy %" should be replaced with another term.
- There should be an indication of which programs has relevant news present.
- There should be pop-up descriptions of the different GUI elements.
- "Nulstil" should be replaced with another term.
- "Gem" comment on UI#10 should be removed.
- There should be a Cancel button at the create user interface.

The comments where evaluated and categorised by the components to which they belong, and listed by the overall rating of each component. This was only done with the comments with a rating of 3 and higher.

The three major usability problems in the inTelly user interface, are considered to be:

- The presentation and interaction in connection with evaluating programs.
- The interactions with the notepad.
- The help concerning why programs are hidden and the evaluation of programs.

These three problems will now be discussed.

Discussion

In this section there will be a discussion of the three main usability problems.

Evaluating programs

The major usability problem in the system is for the user having to evaluate the programs in an efficient and intuitive way. There is proposed several different degrees of evaluation for both the user and system. The number of degrees has to be discussed and tested with users, when some of the functionality in the system is implemented. The indication of already evaluated programs and the feedback when the user has just evaluated a program also has to be tested.

Interaction with Notepad

One thing that is found out for certain in this test, is that there can not only be a separate page with a notepad, it has to be indicated at the front page which programs is added to the notepad. The notepad can be implemented in different ways:

- A combination of a frame with a notepad besides the List component and a separate page containing a notepad.
- A combination of a notepad that is integrated in the List component and a separate page containing a notepad.
- A notepad that is integrated in the List component.

Help

The system should provide help for the user concerning why some programs are hidden, because it is not intuitive that the system hides some programs from the user. This help should be provided for the first time users, because when the user knows why some programs are hidden, it seems logically. The evaluation of the programs should also be assisted by the system, because this kind of evaluation is also not known by the user, and may be difficult to understand. This should be tested with new users when the help and the functionality are implemented, to see if it is comprehensible.

Conclusion

This evaluation of the first user interface drafts gave some good ideas of what to be aware of when designing these further. The test has drawn the project group's attention to the usability problems of the user interfaces and given an idea of what parts that has to be redesigned.

The three major usability problems in the inTelly user interface, are considered to be:

- The presentation and interaction in connection with evaluating programs.
- The placement and appearance of the notepad.

- The help concerning why programs are hidden and the evaluation of programs.

This test also gave the impression that there is the need for an addition to the heuristics from Jacob Nielsen. This addition should concern the new or missing functionalities, which the user might find during the evaluation.

Immediate Result

Below the immediate results from the heuristic evaluation is presented. It contains all the comments from the three test participants together with their rating of the problems and the average of their rating.

UI : User Interface number
 He : Heuristic
 #1 : Test person #1
 #2 : Test person #2
 #3 : Test person #3
 Av : Average

Components

Li : List Component
 Fi : Filter Component
 Mi : Miscellaneous Component
 No : Notepad Component
 Ev : Evaluation of programs Component
 De : Description of a program Component
 Ne : News Component
 Ch : Channel description Component

User interfaces

Lp : List of programs
 Di : Dialog agent
 Lo : Login
 Us : User profile
 He : Help
 Mb : Menu bar and logo
 Me : Messages
 Po : Pop-up help

	UI	He	#1	#2	#3	Av	Comments
Ev	1	4	4	4	4	4	The priorities should not contain more boxes for the hidden programs.
Ev	1	1	4	3	4	3,7	The task of selecting which programs to see and not

							to see (to be removed) is not immediately clear. What to do when given this GUI?
Ev	2	1	4	3	4	3,7	Have to guess what the "flueben" and the "X" is for.
No	2	3	4	3	4	3,7	How does the user know if a program is already in the notebook?
Ev	11	1	4	3	4	3,7	The evaluation should be the same as on the front page.
Ev	1	5	4	3	3	3,3	When the programs are moved from e.g. watch to "not watch", there should be some feedback if this takes more than one second
Li	1	5	4	3	3	3,3	If the TV-guide presents some programs that the user typically not wants to see (e.g. programs seen by many users within the same group), they should be clearly marked.
Li	1	8	4	3	3	3,3	The TV Danmark, message is not very informative. There could be a lot of other problems with our system or the connection so there should be several options.
Ev/He	1	10	3	3	4	3,3	Missing description of green, yellow and red
Li/He	2	3	4	3	3	3,3	The user don't know why some programs are hidden.
De	2	4	3	3	4	3,3	"Omtale" is used twice and they are not directly connected.
Ev	2	5	4	2	4	3,3	Can't see what has happened when the positive sign is clicked.
Ev/De	4	3	4	3	3	3,3	It should be possible to select to watch or not to watch the programs in this detailed description (decisions will often be taken when the user is looking at the detailed description).
Po/He	5	1	3	3	4	3,3	No description of any of the icons
Po/He	5	1	3	3	4	3,3	The "?" could be understood as a help symbol.
Us	8	9	3	3	4	3,3	There should be a warning if the user moves to the next page without saving the data entered!
Ev	11	1	4	2	4	3,3	There is not a directly connection between the list of possible answers and the coloured radiobuttons and the radiobuttons in the table.
Li	13	2	3	3	4	3,3	"Fuzzy %" is a difficult term.
	13	6	3	3	4	3,3	The user feels trapped when finished with this UI. There is no exit (e.g. to the program listing).
Ev	17	4	4	3	3	3,3	Should be the same good/bad indication as all over the system (prioritering)
Li	1	1	2	3	4	3	There is no indication of which programs that has news links.

Li	1	8	4	2	3	3	Important to present if there is missing some channel data in the program listing.
Li	1	10	3	3	3	3	Missing the indication of the sort possibility
Li	2	1	4	2	3	3	Missing an indication of how many programs "skjulte" contains. Want to know how many programs there are before I want to decide if I want to check them.
De	2	2	3	3	3	3	It is not clear what changes the description.
No	2	2	4	1	4	3	The user has to look at the notebook to see what is selected to watch.
	4	6	3	3	3	3	There should be a "back"-link.
No	5	1	3	3	3	3	Not clear what the arrows do
No	5	2	3	3	3	3	The terms "Optag", "Se" and "Måske" are perhaps not the best available.
	5	6	3	3	3	3	There should be a "back"-button (e.g. back to the program listing).
	6	6	3	3	3	3	No indication of how to get back to the front page
Us	9	9	2	3	4	3	It is not clear what "Nulstil" does (see also UI#8).
Us	10	1	2	3	4	3	Missing keywords at all categories
Us	10	1	2	3	4	3	"Gem" comment should have been mentioned before so not necessary now
Us	10	1	2	3	4	3	It should be possible to specify the categories by entering some associated keywords.
Us	11	1	2	3	4	3	To much data presented to the user (if all possible series are shown)
Us	13	1	3	3	3	3	There is missing a preview function.
Us	13	9	2	3	4	3	There is no idea to make the notepad the start page, because this side is empty until the user moves some programs from the program listing.
	1	1	2	3	3	2,7	Maybe you don't understand where "Mine kategorier" is defined.
	1	1	3	1	4	2,7	What can the search box be used to? Search within the selected programs or?
	1	2	3	2	3	2,7	You can be in doubt about what "Vis skjulte" means. And why they are hidden.
	1	2	3	2	3	2,7	Is "prioritet" the best word?
	1	2	3	2	3	2,7	What is "skjulte programmer"?
	1	4	2	3	3	2,7	Missing advanced in dato, tidsrum and kategori
	1	4	4	1	3	2,7	The fonts used seem to be different (filter text and program listing).
	1	4	2	2	4	2,7	Search is placed in two different places are they different?
	1	5	3	2	3	2,7	It is not clear that the system has selected the

							programs presented to the user, when he or she logs in.
	1	10	3	2	3	2,7	There should be pop-up messages on the different GUI items.
	2	2	2	3	3	2,7	The program selected for the detailed description is not clearly visible. (It is not possible to see which program the description is connected to).
	2	5	3	2	3	2,7	There should be an indication when the system is working on getting, filtering or sorting data.
	3	1	2	3	3	2,7	"Notesblok" does not contain enough information. Missing data and time for program
	3	3	2	3	3	2,7	The user can't see data about programs in the notepad.
	3	7	3	2	3	2,7	Shortcut to details of notepad.
	4	1	3	3	2	2,7	Sequence: 1. Read details, 2. Decide if to see program, 3. Select/deselect program, missing, 4. Back-button, missing
	4	5	2	3	3	2,7	What about missing data, try indicate this way - Showview code: - (if none)
	5	1	3	2	3	2,7	The state each program is in, should be stated by the icons.
	5	1	2	3	3	2,7	It is unclear if a click on an arrow, moves the program one row down, or down to next section.
	5	1	2	3	3	2,7	How do I get more data about the programs?
	5	1	3	2	3	2,7	The different categories "Optag", "Se" and "Måske" are hardly visible.
	5	4	2	3	3	2,7	It is not possible to sort the table by clicking on the title bar like in the program listing.
	6	1	2	3	3	2,7	Which characters should be used, is there any demands for the username and password? (Present legal characters, etc.)
	6	5	2	3	3	2,7	If the user name is taken, it should be stated which username it was.
	6	6	3	2	3	2,7	I would like to Cancel.
	6	8	3	3	2	2,7	What is presented to the user if e.g. he/she does not enter data in a field or enters a wrong password in the last field etc.
	6	9	2	3	3	2,7	There should be a presentation of legal characters/examples etc. for each input box.
	6	9	3	2	3	2,7	It is not clear what the "...under Brugerprofil..." refers to.
	7	1	3	2	3	2,7	It would be nice to know, whose user profile it is. And

							when it was last updated.
7	2	2	3	3	2,7		The term "Side XX" is not a usual Web term.
8	1	2	2	4	2,7		Does next button save?
8	2	2	2	4	2,7		Does "Gem" save all data or only data on this page. The user could be in doubt whether it is necessary to hit "Gem" for each page or?
8	5	3	2	3	2,7		How do I know if it has been saved. There could be a date for when it was last saved.
8	9	2	2	4	2,7		What does "Nulstil" do? (Clears the whole user profile or only the data on this page?)
8	9	2	3	3	2,7		Missing examples of how to write the different data
8	10	1	3	4	2,7		General for all user profile UI's: There is no help available!
9	9	2	2	4	2,7		It is not what will happen is the user hits "Tilbage" when he or she comes directly from e.g. the user profile start page (UI#7).
11	1	3	2	3	2,7		Should they be divided into categories?
11	1	3	2	3	2,7		The user should be able to suggest new Repeaters.
12	1	2	2	4	2,7		It seems dangerous that the "Nulstil" removes all keywords, but a very visible undo facility could be a solution.
12	1	3	2	3	2,7		If the previous UI makes it possible to enter keywords associated with each category, then the user could be confused about the keywords on this page.
13	1	2	3	3	2,7		Description of what to do on this page (it's purpose)
13	1	2	3	3	2,7		Have not been presented to "notesblok" or "nyheder" before
13	1	2	3	3	2,7		Description of the different options on the page
13	1	3	2	3	2,7		It is not clear what the start page means (is it the start page after logon?)
14	6	2	3	3	2,7		There is no back or "home" buttons.
15	1	2	3	3	2,7		I agree that the primary thing to explain is the priority, but the other issues in the program listing should also be described.
18	1	3	3	2	2,7		Why only for registered users
18	3	2	3	3	2,7		It is important that the user can see the original web page when navigating in the dialogue (he might forget what he or she was originally looking for after a while...) It should therefore also be possible to move the dialogue window.
1	1	2	1	4	2,3		The remarks to the user are placed in the bottom. This could be a problem if the user is not aware of the

							comments when the program list is long.
	1	2	3	1	3	2,3	Should it not be "Korte beskrivelser" instead of "Omtale"?
	2	1	1	3	3	2,3	Seems a bit scary when you get the description presented that way (to much text).
	2	2	1	3	3	2,3	Why is there a scroll bar, when it is not used?
	2	2	1	2	4	2,3	There is no indication of what the filter select-boxes contain.
	2	2	1	3	3	2,3	Not clear how the user can select more details about a program.
	2	4	3	1	3	2,3	The checkbox "Omtale" should be named "Kort omtale" or "Kort beskrivelse".
	3	1	1	2	4	2,3	If two programs have the same title (e.g. "Nyhederne") then this difference is invisible in the notepad.
	3	1	1	3	3	2,3	The term "Notesblok" appears twice, but do they have the same meaning?
	4	1	2	2	3	2,3	There is no link to the complete news articles.
	4	3	3	3	1	2,3	When going back to the list of programs the user should end up at the program just visited. Maybe the program should be marked in some way.
	4	8	2	2	3	2,3	If one of the boxes is empty, there should be an explanation.
	5	1	1	3	3	2,3	No indication of what the first column contains
	5	1	2	2	3	2,3	Change "record" symbol to "R" or "VCR" or something else that gives clear indication of what is meant instead of the red button
	5	1	3	1	3	2,3	More space between the different sections.
	5	1	2	3	2	2,3	How do I remove data?
	5	1	1	3	3	2,3	It is not immediately obvious that "Optag" is higher prioritised than "Se".
	6	3	2	2	3	2,3	The "code reminder" function should be explained.
	6	9	2	2	3	2,3	Missing description of what the user shall do to create a profile.
	7	1	2	2	3	2,3	The page number does not say anything.
	7	6	2	2	3	2,3	No way "back".
	8	1	3	1	3	2,3	"Gem" and "Nulstil" should be disabled if nothing is changed.
	8	1	2	2	3	2,3	Is it necessary to save every page? Maybe there should be a button for save all.
	8	2	2	2	3	2,3	Will "Gem" fade when the data is saved? (Should be that way).

	8	5	2	2	3	2,3	There is no presentation of why we need the different data.
	8	7	1	2	4	2,3	It should be possible to get to main menu of help.
	8	9	1	2	4	2,3	The day and month are not clearly stated (the order is different in e.g. English web sites).
	8	9	2	2	3	2,3	The "Nulstil" command sets up a need for an undo command.
	8	9	2	2	3	2,3	Missing describing text about buttons function
	9	1	1	3	3	2,3	The small "package" icon seems to have no function.
	9	1	1	3	3	2,3	Does not understand "fjernsyns" icon
	10	2	2	2	3	2,3	The category "Øvrige" is not understandable.
	10	5	3	1	3	2,3	The user should be informed if e.g. some categories could result in many or few hits, and what to do to about it (e.g. enter some additional keywords)
	11	1	2	2	3	2,3	It is not clear what the user can do if a series is not present in the list.
	11	1	2	2	3	2,3	The colours should be present in the whole page or completely removed.
	11	1	2	2	3	2,3	The "+/-" signs are not explained.
	12	1	2	2	3	2,3	Does the system add keywords?
	12	1	2	2	3	2,3	It seems to be impossible to enter phrases.
	12	9	2	2	3	2,3	Missing description how to write more than one keyword at the time.
	13	1	2	2	3	2,3	I would like some examples of column data.
	13	1	2	2	3	2,3	It is not possible to understand what the time selection means for the layout.
	13	1	3	2	2	2,3	It is not possible to decide the order of the information in the program listing.
	13	1	2	2	3	2,3	There is no information about what is happening when data is not available for the program listing (will the columns still be visible or?)
	13	2	2	2	3	2,3	"Prioritering" and "Popularitet" is two similar terms (it is hard to understand the difference).
	13	4	2	2	3	2,3	The fonts used are not the same.
	15	1	1	3	3	2,3	There should be an explanation of what the user gets out of doing the priorities.
	15	1	1	3	3	2,3	There should be links to e.g. more detailed descriptions of the TV-guide (e.g. when hitting one of the priority explanations there should be some further description).
	15	6	2	2	3	2,3	Remember "back" and "home" buttons.
	15	10	2	2	3	2,3	Description of user profile and its possibilities

	17	1	1	3	3	2,3	Missing channel and time
	17	1	2	2	3	2,3	How do I reach earlier messages?
	17	1	3	1	3	2,3	The message cover information on the page. The message should appear in the upper right corner of the screen, this is empty.
	17	1	2	2	3	2,3	The number of dialogues popping up should be kept in a minimum. If this is not possible there should be made another solution (e.g. integrated in the web pages).
	17	2	2	2	3	2,3	It should be possible to get additional information of each message, by e.g. clicking on the message text.
	17	5	2	2	3	2,3	Is there any feedback if I press "God"?
	17	10	2	2	3	2,3	There should be some help in using this function.
	18	1	2	2	3	2,3	Is it possible to ignore the agent?
	18	4	2	2	3	2,3	Maybe the dialog items should be recognizable from the menus.
	18	6	2	2	3	2,3	Remember "back" and "home" buttons.
	1	1	2	3	1	2	The logo is quite large compared to the complete page (I hate remarking this!)
	1	2	2	2	2	2	It would be better to call the "kategori" for "genre"
	1	2	2	1	3	2	What is "brugerprofil"?
	1	3	2	1	3	2	It should be easy visible when selected programs overlap.
	1	6	2	0	4	2	An undo function should be available when e.g. the user regrets a "do not watch" command.
	1	7	0	3	3	2	As said before the system should present shortcuts to often used sequences of commands.
	1	7	1	3	2	2	Missing shortcut to program setup and category setup in user profile
	1	9	1	2	3	2	The system should notify the user if an often seen series suddenly is deselected.
	1	10	1	2	3	2	The system should look in the last commands when the user suddenly asks for help. (Makes it possible to make relevant help directly)
	3	1	1	3	2	2	It is difficult to know that there is a shortcut to a detailed description, when pressing the name of a program.
	3	1	4	0	2	2	A more detailed notepad should be available.
	4	1	1	3	2	2	Missing mouse-over description of symbols like ((S))
	4	1	1	2	3	2	There could be a presentation of which channels the series also is available.
	4	7	1	1	4	2	There is no link to the news providers' homepages.

	6	1	1	2	3	2	What do I get out of a user profile?
	6	1	1	2	3	2	There is no title describing the contents of this page.
	6	1	1	2	3	2	There is only a description of what to do after the entering of the data, not a description of what to do when entering the page.
	6	1	3	3	0	2	Not necessary to have "glemt password" button, just confusing.
	6	5	3	0	3	2	When a new user is created, it should be clearly stated. And the user should be told which user name is created, and then try to login by pressing a button.
	6	5	3	1	2	2	Does the system save my username? (Make some feedback)
	6	10	0	2	4	2	There seem to be no help for this page (besides the text already present).
	7	3	2	2	2	2	The six pages could be represented/supplemented by icons.
	7	10	1	2	3	2	Missing description of the available pages - show some keywords.
	8	1	1	2	3	2	The year of birth could be only "70" etc.
	8	2	2	1	3	2	It is not clear where the next button jumps?
	8	6	1	2	3	2	It not possible to go back?
	8	9	1	2	3	2	What happens when the users enters a wrong email (e.g. without "@")?
	9	1	1	3	2	2	Button to choose channel for your area. It should be possible (e.g. free of choice) to get the channels for the users' local cable net.
	9	1	2	3	1	2	"Online" should be under "Kategorier"
	9	1	2	2	2	2	There should be an explanation of what "efter kategori" means.
	12	1	1	2	3	2	It is not clear whether it is allowed to enter more than one keyword at a time and if it is then how?
	12	9	2	2	2	2	Missing description of "skraldespandsfunktion"
	13	1	2	2	2	2	The table is a mix of checkboxes, text and selectboxes.
	13	3	1	2	3	2	There should be a default set-up, which is well suited for e.g. a 1024x768 monitor resolution and/or an 800x600 resolution (could made by a checkbox or by detecting the current resolution).
	13	10	0	2	4	2	There is no help.
	14	1	1	2	3	2	It should be possible to get a task-oriented help.
	14	3	1	2	3	2	Remember to use many of the same icons and screenshots in the help description.

	14	7	1	2	3	2	There could be links to the most often-visited help pages.
	15	2	1	2	3	2	"UI#4": I do not understand the colour codes on the time bar. How is it marked when two interesting programs are overlapping?
	15	10	2	2	2	2	Description of normal/advanced features (dag,tid...)
	17	2	0	2	4	2	What does "m" stand for?
	18	1	1	2	3	2	It would be nice if the user could make a "search" in the dialogue.
	18	1	1	2	3	2	The help should be task oriented.
	18	5	2	2	2	2	There should be some automatic help if the user takes too long to answer the questions (e.g. "Do you understand the question?", "Is there missing something?", etc.)
	18	7	3	1	2	2	The issues shown could be dependent upon the most popular issues on the web site at the present time.
	18	8	2	1	3	2	The user should be helped/guided if he or she gets to a dead end.
	18	10	2	2	2	2	There should be some additional help/description to this function.
	1	2	3	0	2	1,7	Should it not be "Omtaler" instead of "Omtale"?
	1	2	1	1	3	1,7	What does the "Brugerprofil" menu item (makes a new one, changes the old one or?)
	1	3	0	2	3	1,7	The user interface should remember often-used sequences of commands (e.g. the user often selects some specific channels, categories etc.)
	1	3	0	2	3	1,7	Showview codes should be directly available from a checkbox.
	1	4	1	2	2	1,7	The select-boxes in the filter should have the same size.
	1	5	1	2	2	1,7	It might be good to indicate the time it took to get the information.
	1	10	1	1	3	1,7	It would be a good feature if it was possible to draw a "?" down to the item for which help is needed.
	4	1	1	3	1	1,7	The user would be surprised of the news, web sites, etc. presented here. Why are they here, he didn't ask for them. Maybe they are more useful when he has seen the program.
	4	1	1	3	1	1,7	To much information. I just wanted to know a little bit more about the program, to decide if I wanted to see it. Maybe "Nyheder", etc. should be shortcuts instead.
	4	1	0	2	3	1,7	The complete detailed description should be displayed

							without the need of a scroll bar.
4	1	1	2	2	1,7		Is "Data" a good box-title?
4	1	1	2	2	1,7		There should be links/presentations of similar programs (if any).
4	1	1	1	3	1,7		There is no title on the page.
4	3	0	3	2	1,7		It should be possible to put selected detailed information into the program listing/notepad.
5	1	1	2	2	1,7		There is too much space used for each program compared to how little information that is presented for each program.
5	1	2	0	3	1,7		The user has to go back to the original program listing to find programs that are not already present in the notebook.
5	7	2	2	1	1,7		Maybe there should be a shortcut to the show view codes.
5	9	1	0	4	1,7		An undo function should be available (if the user regrets a de-selection of a program).
6	1	0	2	3	1,7		The columns in the tables are not named correctly.
6	4	0	2	3	1,7		There is no menu at this page.
6	10	1	2	2	1,7		Missing description of what to do when a user with the chosen username already exist.
7	1	1	1	3	1,7		No menu or title on the page.
7	1	0	2	3	1,7		The warning text (about cookies etc.) should be hidden (shown e.g. by hitting a button) or only shown the first time the user enters the page.
7	1	3	2	0	1,7		Have there been changes since last time?
7	9	1	2	2	1,7		Move page listening down below "Brugerprofilen bliver benyttet....."
8	1	2	1	2	1,7		"Din brugerprofil er ..." is repeated from the first page.
9	1	1	2	2	1,7		The two top rows in the table (containing name and checkbox) are not clearly separated from the rest of the columns and from each other.
9	4	0	2	3	1,7		No menu or title or logo.
9	6	0	2	3	1,7		There could be a "quit and save" possibility. (Goes for the previous UI's).
10	1	1	2	2	1,7		The tables seem too big compared to the text inside them.
10	1	0	2	3	1,7		It should be possible for the users to create their own categories (and e.g. associated keywords).
10	1	1	2	2	1,7		Again the two first rows are too similar and the first checkbox looks much like the other checkboxes.

	11	3	0	2	3	1,7	There are missing some program types (e.g. news – channel dependent).
	12	1	1	2	2	1,7	I would like to see some examples of keywords, e.g. advanced keywords.
	12	3	2	2	1	1,7	Again there should be some more icons.
	13	1	1	1	3	1,7	There should be some information about how often data could be expected to be present for the different columns (the "Seertal" is typically not available for every program, specially when the program is e.g. send for the first time).
	14	1	1	1	3	1,7	There should be both an index and search function.
	14	1	0	2	3	1,7	The shown help page should be dependent upon the last visited page (program listing, notepad, etc.)
	14	4	0	2	3	1,7	Again the top and bottom of the page is not the same as for the program listing page etc.
	14	10	0	2	3	1,7	There should be some "help to the help" (?)
	15	3	1	1	3	1,7	It should be possible to sort out programs that do not overlap and programs that overlap.
	17	1	1	1	3	1,7	What if the message is bigger than the window?
	17	1	1	1	3	1,7	It should be possible to select that the message do disappear after X seconds, if the user always want to look at/take action on the incoming messages.
	18	1	1	1	3	1,7	The guide dialogue should be dependent upon the last web page shown to the user.
	1	1	0	2	2	1,3	The GUI seems to have a bit of a list ("slagside") (the filter section is dominating).
	1	1	0	1	3	1,3	The scroll bar under the filter section seems unnecessary.
	1	2	2	0	2	1,3	It can be hard to figure out where to see, e.g. the show view code, maybe user profile should be named "Indstillinger" instead.
	1	3	1	0	3	1,3	There should be links to description of the different channels.
	2	1	0	2	2	1,3	Missing date in "tidspunkt". I do not know what day the program is shown.
	4	1	2	0	2	1,3	There should be displayed more details in the data field e.g. how many is watching the program etc. (it is too big for the small amount of information).
	4	3	1	1	2	1,3	If there is overlap between this program and other selected programs this should be indicated.
	4	4	1	1	2	1,3	The boxes are not the same size.
	5	1	1	1	2	1,3	There is no title on this page.

	5	2	1	1	2	1,3	Tid -> Tidspunkt
	5	3	1	0	3	1,3	Overlapping programs should be indicated and suggestions should be made accordingly (if video recorder is available).
	6	1	0	0	4	1,3	What does "[FrontPage Save Results Component]" mean?
	6	2	2	1	1	1,3	Is "Opret bruger" a good term?
	6	7	0	1	3	1,3	There should be a "clear fields" function.
	6	8	1	3	0	1,3	Maybe suggestions to other user names.
	7	7	0	2	2	1,3	The advanced user would miss some shortcuts (are not interested to visit this page)
	8	1	2	0	2	1,3	Maybe "Nulstil" should be renamed and instead say e.g. "get from database".
	8	4	0	1	3	1,3	Missing title and menu bar and logo.
	9	3	1	0	3	1,3	There should be some short presentation/description of each TV-channel.
	18	1	2	2	0	1,3	Would maybe be OK for the profile but annoying when checking todays TV-program
	1	3	0	0	3	1	Maybe the duration of the programs should be showed.
	4	1	1	0	2	1	How has others evaluated the program?
	4	7	1	2	0	1	There should be a "next"-link that presents details for the next selected program. (This will make a need for a program-listing-button).
	5	1	1	1	1	1	It should be possible to see old programs from my notepad.
	7	4	2	1	0	1	The "inTelly.com" name is written with lower case letters.
	7	10	0	0	3	1	It could be mentioned that many Web sites makes cookies without even ask the user!
	9	1	1	0	2	1	Small thing: are there not any American channels (are they all English?)
	9	7	1	0	2	1	To channel descriptions.
	10	3	1	0	2	1	There should be some icons indicating the different categories.
	12	1	2	0	1	1	"Tilføj" could be "Gem i database"
	13	1	0	0	3	1	The "Reklamelængde" is a good feature.
	14	1	1	0	2	1	Online ask function. Indicated if someone is online with a green and red indicator.
	1	1	0	2	0	0,7	Not interested to know how many programs there are presented.
	2	2	2	0	0	0,7	To show the sequence in which the user should be

							able to see what to see in the telly, there could be put a 1. at the filter, 2. at the List and 3. at description.
	4	1	0	0	2	0,7	What is (160) – irrelevant information, nope, number of the serieal.
	4	1	1	0	1	0,7	Maybe it should say "Nyheder ifm. Strengt fortroligt"
	4	1	0	0	2	0,7	A discussion forum is wanted.
	4	1	0	0	2	0,7	The picture is also too small for the box.
	5	4	2	0	0	0,7	I don't think it should be possible to sort this table.

Table 1: Rated comments from heuristic evaluation.

Test Plan

Assessment Test

(version 0.2)

Responsible: kk, lkr, pmj
Test date: 20th and 21st of February 2001
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|Iterative|Installation]

Test plan

This paper will describe the assessment test plan for the inTelly version 0.2. In this test the participants are asked to perform some tasks on the inTelly system and asked how he or she "feels" (positive and negative considerations) about selected parts of the design.

Introduction

The test plan contains the following items:

- Purpose
- Objective
- User profile of test participants
- Method
- Accessories
- Task list
- Test monitor role
- Evaluation measures
- Report Contents and Presentation
- Description of non-implemented functions and features

Furthermore our test plan will include:

- Possible outcomes of the test
- Actions to be taken according to the possible test outcome

The following sections will comment on the different items in the test plan.

Purpose

The purpose of the test is to get ideas and critics from future users by letting a number of participant's tryout the version 0.2 of the inTelly system.

Version 0.2 is the first implementation of functionality in the system, where version 0.1 only contained the user interface design. It should be noticed that not all functionality is implemented in version 0.2.

Objective

The objective of this test is to point out the strengths and weaknesses of the current version of the system. This concerns both functionality (implemented and non-implemented) and user interface aspects. The user will be asked to do some specific tasks, to state an opinion about some system features and to describe how some tasks could be performed by the use of intuition. The results will be used in the design of the

next version. It should be noticed that this test also should evaluate whether there is a correspondence between the inTelly version 0.2 and the user's conceptual model of the product.

Another important issue about this test is that it should focus on the "new" things in the inTelly system like the personalisation part. Even though the user profile is only partially completed this should not give any problems with testing the personalisation because the most important parts like channels and categories are implemented.

User Profile of Test Participants

The users performing the test will be selected from the final user group. This means that the participants will be advanced Internet users. The participants will be 5 employees from MINDPASS A/S.

Method

The method used in this test will be to ask the participants to:

- Solve some given task on the inTelly version 0.2.
- State his or hers opinion about some of the features on the inTelly version 0.2 (primary uncommon features).
- Describe how some tasks could be performed by the use of intuition.
- Describe the overall experience with the system.

The following list will describe the outline of the test:

- The first thing to be done is to ensure that the participant recruited is an advanced user. This means that the participant is asked some questions in order to determine whether he or she is an advanced user.
- Secondly the user should be informed about the overall purpose of the system: It is a personalised TV-guide.
- The method will be that the test monitor asks the participant to perform the tasks (given in section: Task List) one at a time.
- When the test monitor has ensured that the participant has understood the task, the participant should start act upon it (could be done by a description and/or by doing some operations).
- The participant will be asked to inform the test monitor about why and what he or she is doing during the exercise. This will include good and/or bad things about the functionality and the layout.
- A second test monitor will write down relevant actions performed by the participant and also the utterances of the participant. This task will be rather demanding, but it will be allowed to interrupt the participant in order to get opinions and actions clarified and to get a further description of the comments uttered by the participant.

- The test participant should be debriefed, about his/hers overall experience in using the inTelly system.

The test monitor must make it clear that the questions are made in order to evaluate the performance of the inTelly system and not the participant. This means that the participant should not be frustrated or stressed when some task is hard to get right (this only indicates a major problem in the inTelly design). The duration of the test will be approximately 30 minutes. This duration is considered appropriate, because the time is not too long so that the participant becomes impatient. The order of the questions will be the same for all participants.

It should be noticed that pilot tests are made in order to evaluate the test itself and make corrections on the system and/or the tasks to be performed before the final test.

As mentioned above it had been chosen to include two test monitors in this test. The main reason is that the work of writing down the participant's actions and comments, to communicate with the test participant, to ensure that the system runs perfectly etc. is a too big task for one test monitor.

Accessories

The accessories to be used are:

- The inTelly version 0.2 (this includes web-files, agent framework, databases, jdk1.3, Tomcat 1.2, Java Servlets, MSXML version 3.0 installed on the browser).
- A computer with an Internet connection and a browser installed.
- A list of questions to be asked.
- A list of descriptions of non-implemented functions in order to give the participants the same description every time they try to use a non-implemented function (See also the section: Description of Non-implemented Functions).
- The test environment will be an office-like room.

Task List

The tasks the participant should go through are given by the different questions listed below. The reason for asking the questions is placed after each question.

- 1) What is your first opinion about the start page (how about the layout, how is the purpose of the page presented (intuitive or?), what are your expectations to this site, etc), are there anything you do not understand? The result of this task should indicate whether the user's mental model corresponds to the purpose of the system. Furthermore this task should give some indication of where there are usability problems in the inTelly system.
- 2) Try to use the system without logging in (e.g. find out what to see on TV tonight). This task is meant to give the user a chance to try some of the simple features of

the system. Furthermore this task should show how the first time user performs in using the system.

- 3) Try on your own to make an account on the inTelly web site. Please fill out the channel and category pages in the user profile (it is possible to break off the user profile creation at any time, but this is a test situation and it is important to evaluate all the web-pages in the user profile). This task is primary used to compare the generation of a user profile in the inTelly system to other TV-guides evaluated in an earlier test (in the Observation test).
- 4) Try to describe the different columns in the program listing to the test monitor. (Focus on how the participant understands the "new" features such as priority and AI). The result of this task should give an idea of how the new features are interpreted by new users.
- 5) Ask the test participant to evaluate today's programs in order to get a feeling for the notepad. After the evaluation the participant should look at tomorrow's program listing in order to get an idea of what is meant by this type of personalisation. (Remember to get the immediate response from the participant – The user must be logged-in in this question). Also remember that more user evaluations are needed in order to predict the programs the user likes to see. The reason for performing this task is to both evaluate the notepad (if not already done) and to present the principles in the personalisation.
- 6) The user should also have tried out the different types of filtering, sorting and detailed description in this test (ask the user to do some filtering and sorting if he/she has done this in a previous task). The result of this task should show how the different filtering facilities perform and if they are understandable.

Test Monitor Role

The role of the test monitor will be the following:

- Presentation of the superior tasks to be performed by the participant (remember to state that it is the inTelly system that is evaluated and NOT the participant).
- The test monitor should state clearly that the participant should try to abstract from the slow speed of the system. The version of the system is still a prototype and there are much more important aspects about the system to be tested at this state of the development (e.g. the user's conceptual model of the system).
- If the participant ends up in a situation where he or she is trying to use a non-implemented function the test monitor should immediately interrupt the participant and describe what is meant to happen in future versions by reading the corresponding document describing the non-implemented function (see also Description of Non-implemented Functions). Remember to get the participant's opinion about the non-implemented function.
- The test monitor should ask the participant to act upon the different tasks.
- When the test participant is having problems with a question, the test monitor is allowed to help the participant.

- During the test the test monitor writes down the actions and statements made by the participant.
- The test monitor can ask the participant to comment on specific and common actions/statements.
- The test monitor can stop a task and jump to the next question if the participant is having serious problems.

Evaluation Measures

The evaluation measures will be the comments (both positive and negative) given by the participants during the test and the tasks (or actions) they perform in order to act upon the questions.

Report Contents and Presentation

The outcome of the assessment test will be a discussion of the comments and tasks performed by the participants. In the end of the discussion there will be a conclusion.

There will also be an overall presentation of the good and bad things about the inTelly system.

Description of Non-implemented Functions

The list below contains a short description of the non-implemented functions.

- Search feature is not implemented. This feature will in future allow the user to make a search in the TV-programs. The presentation of the search results will be a new program list similar to one displayed in the normal program list where there are added some information about the search result.
- The detailed description window of a TV-program is not implemented (available by clicking a program name in the program listing). There is implemented a default detailed description which is always presented to the user when he or she clicks on some TV-program name in the program listing. The test participant should try to abstract from the fact that program presented in this window does not correspond to the one that was clicked on in the program listing and try to evaluate the detailed description in general (the layout, news presentation, etc.)
- There are no messages from the system to the user available in this version of the system. An example could be a message to the user when a Repeater that the user usually likes to watch suddenly is moved to another time of day. This feature will probably be implemented in a future version of the system.
- There is an error message that sometimes pops up when using the system. The user should just click 'ok' and continue as nothing has happened. It should be

stated that the error indicated has no affect on the use of system and it will of course be corrected in a later version.

- The user profile is only partial implemented which means that only channels and categories are implemented. So there is missing the following parts of the user profile: personal data, Repeaters, keywords and settings.

Possible Outcomes of the Test

Some possible outcomes of the test are presented below:

1. The non-implemented functions seem to annoy the participant (indicating the test should have been performed at a later time when the functions have been implemented or the questions could be formulated in a way that ensured that the participant did not need the non-implemented functions).
2. The participants do not understand the descriptions of the non-implemented functions.
3. The system fail to present the works of the integrated notepad and the system priority in an intuitive understandable way.

Actions to be taken according to the possible test outcome

In the previous section a number of different outcomes of the assessment test was presented. In this section the actions to be taken according to the outcomes are presented. This includes both actions to be taken before and after the test in order to avoid the situations described.

1. If non-implemented functions annoys the participant at the time of the final test nothing can be done expect from taking in into account in the evaluation of the test. If the pilot test indicates that the non-implemented functions will be a problem in the final test it should be decided whether it is necessary to implement the functions and/or adjust the task list in order to avoid the non-implemented functions. Another possibility is to accept the problem and comment on it in the final evaluation of the test.
2. If the participants do not understand the presentation of the non-implemented functions in the pilot test the description should be changed accordingly. If the descriptions only show to be a problem in the final test the test monitor should try to explain the non-implemented functions as good as possible and it should be considered whether it is necessary to adjust the description of the non-implemented functions to be used in the following tests. It should be noticed that if the descriptions are changed during the final test it could affect the final test result. How this situation influences on the final result should be considered in the evaluation of the test results.

3. If the special facilities in the inTelly system (primary the notepad and the system priority) fail to work by some intuition of the test participant it should be considered how to increases the performance in later versions of the system. The results of the test should indicate the problems with the new facilities and the result should be used in the making of future versions.

Test Result

Assessment test

(version 0.2)

Responsible: kk, lkr, pmj
Test date: 20th and 21st of February 2001
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|**Iterative**|Installation]

Test Result

The test results of the assessment test will be presented in this document. First there will be a section presenting the most interesting immediate results (evaluated results), next there will be a discussion of the results followed by a conclusion on both the results and the test itself. In the last section of this document there are some lists containing the immediate results.

Evaluated Results

In the assessment test five participants were presented to the inTelly version 0.2. The objective of the test was to evaluate both functional and usability aspects of the inTelly system primary concerned the integrated notepad, the system priority and the user evaluation of TV-programs.

The test participants contributed with many comments concerning the usability and the functionality. In the presentation below only the most important and most interesting results are mentioned. It should although be noticed that all comments are evaluated and used in some form in the further development of the system.

The selected comments in the list below are both concerned with usability and functionality.

Comments

- There is in general too much text presented to the user in the help pages and in the presentation of the user profile creation. The test participants did in general not bother to read all the text.
- There are no sub contents in the help document, which was mention by one of the test participants. The sub contents could help the user speed up the search in the help document.
- Some of the test participants would like some kind of a wizard that could guide them through the generation of the user profile. One method could be a indication of how many steps there are in the generation of the user profile and which number the user is currently carrying out. (E.g. "step 2 of 4")
- In general there are missing feedback in a number of situations:
 - When a user has been created.
 - When the user profile has been saved.
 - When the user profile has been retrieved from the database.
 - When the user has performed an evaluation.
 - When no programs are shown in the TV-program list.
- There could be a problem when using cookies on a computer used by several persons.

- The works of the evaluation, notepad and system priority is not immediately intuitive. But the participants quickly understood the idea if they were told a few words about the principle.
- The system does not indicate that an account should be created in order to use the features of the system. There is no specific help to the first time user.
- There is no pattern in the evaluation method used by the test participants: Some use primary negative feedback and others primary positive feedback.
- Several test participants feels trapped by the pre-defined time slots in the filter menu. They miss some advanced filtering options. Could be a user defined time slot and/or a user defined channel selection.
- There is missing some search feature.
- An indication of TV-programs overlapping in time would be a good feature.
- The colours should be tested on colour-blind people. The colours in version 0.2 were difficult to separate for one of the test participants.
- There was missing some indication of how far in the program list the user has got.
- Some of the test participants mentions that it is possible that he/she would have different TV-program desires dependent on the weekday, time of day etc.
- Two participants think that the evaluation radio buttons are reversed according to their intuition. They think that the negative evaluation should be the first radio button.
- The TV-programs explicit selected by the user is displayed in the program list independent on the settings in the filter menu.
- It should be possible to deselect the system priority.
- Missing an undo function to be used when a program is removed from the list because of a negative evaluation.
- Some test participants would like a specific training page instead of the current integration between the program listing and the evaluation feature.
- It would be nice if the user explicit could start the generation of the system priorities.
- There should be some popup indication to be used in special situations (e.g. when a film the user for sure would like watch is available in the TV-program database).
- It is easy to mix-up the duration of a program and the duration of a program.
- The short description popup (activated on "mouse over") disappears too quickly.
- A feature that makes it possible to add keywords is missing.
- Suggestions based on users that have approximately the same user profile are considered to be a good feature.

Discussion

In this section there will be a discussion of how the comments listed in the previous section can be interpreted and suggestions to solutions and/or future actions to be taken.

Several of the comments are suggestions to additional functions and improved usability, which has already been considered by the design team but it was not possible due to time constraints to implement it all in version 0.2. A question could be raised whether it is waste of valuable test time spending time on aspects already known by the design team. Nevertheless it was chosen to let the test participants state their opinions about the features in order to be able to evaluate where to put focus in the future system development. It should be noticed that this contribution from the test participants was not a directly part of the test objective, but it has not taken time from original task list; the result was instead that the test duration typically was 45 minutes (Expected duration was 30 minutes).

One of the most important test objective was to see how first time users interprets the evaluation, integrated notepad and the system priorities. The result was that first time user did not catch the new features immediately, but there was required a minimum of information in order to understand them correctly. It will be necessary in the next version of the system to present some kind of messages to the first time users presenting the ideas of the features. The amount of information needed has shown to be minimal and there is reason to believe that it is possible to present the works of the features to the first time user in a few sentences.

The above comments brings us on to the next important comment: The test participants did not bother to read much text, both in help situations and in the presentation of the user profile generation. This result could point in two directions. If the result can be generalised to all advanced Internet users then it means that all textural information should be removed from the inTelly web site. On the other hand it is possible that the test participants are not representative to the advanced users on the Internet. This is probably the situation because the test participants were collected among the employees at MINDPASS A/S and these users are probably extra impatient when using the Internet (will not waste time on web sites that do not seem to be immediately intuitive). The conclusion must be that even though the amount of text in some situations seems too extensive for the test participants it should still have its place on the web site in order to satisfy other users that might want more text to read about the system. The presentation of the help facility and creation of the user profile will although need a work over (use of wizards etc.).

The test participants were missing system feedback in several situations. There can be no doubt about the importance of feedback to the user and the problems notified by the test participants must be solved in the next version of the system.

Another comment about the test results will be the test participant's different use of the three evaluation measures. Some of the participants preferred to make negative evaluation and others preferred a positive evaluation. This indicates that the order should be clearer to the users.

The final point in this discussion will be a few words about the test form. The task list presented in the test plan has not always been followed strictly. In some situations that test participants has made some of the tasks on their own and sometimes the order of the tasks has been altered due to the test participants actions. This could indicate that the pilot tests has not proved to point out the possibility of letting the participant more or less control the test. And only let the test monitor ask specific questions when a participant misses some features that has to be tested. It could be discussed whether the task list should be formulated different. But the test monitors succeeded in the letting the test participants control the test and only break in when necessary. It has been considered to be a good idea letting the test participant control the test and making some observations on their own without some given tasks to be performed strictly.

In several situations in the test there were interesting discussions about the system between the test monitors and test participants. In the test plan it was stated that the test form would be the user think aloud and there was not said anything about discussions between the test monitors and the participants. This resulted in a test duration of typically 45 minutes (30 minutes in the test plan), but the design team benefited from the discussion because many of the future ideas where discussed to be used in the development of the next version of the system. But there should be no doubt about this possibility should have been mentioned in the test plan.

Conclusion

The assessment test performed has given a number of constructive inputs to the further development of the system.

Some of the problems pointed out by the participants where already known by the design team, but it was considered valuable to discuss the possible solutions with potential future users even though this was not explicit included in the test objective.

The most important result is considered to be the non immediately understanding of the notepad, evaluation and system priority. At a first glance this could indicate a serious problem, which should result in other suggestions to these major design issues. But as the test showed it was easy to describe the features to the participants (in a few words) and the features seemed natural and uncomplicated to use after a very limited time of use. The conclusion that has been made on this result is that the features is kept in more or less the same form in the next version of the system, but there will be added some message popup or alike describing the new features to the new users as they are needed.

Another result is the pointing out the importance of feedback after user and system actions. Nearly all participants mentioned some missing system feedback when the system. The design team has been aware of some of the missing feedbacks but the test

pinpointed the need and where they should be placed. The necessary feedback could be given in different ways. In some situations the needed feedback could be given by just "fading" a button text, in other situations some minimum of text or message is to be considered in order to present a new and typically unknown feature.

As already mentioned in the above discussion the test plan had a few weak points:

- It did not take into account the possibility of letting the participants decide the order of the tasks on their own.
- Further more it did not mention the possibility of having discussions between the participants and the monitors about the current design and ideas to future versions. (It only describes that the test participants should state their opinions).

Immediate Result

These immediate results are the raw data/comments, which came from the test participants during the test. The numbers of test participants for this test was five and they are all listed below.

Test person: #1

Test monitors: Kenneth Kammersgaard and Peter Michael Jensen

Date: 20th of February 2001

1. Thinks that the order of the programs is odd. They are sorted by the reverse time of day (last programs first).
2. Observes that "The rest of the day" cuts out programs that have been shown.
3. Thinks that the program duration is irrelevant.
4. Tries out some different filters and notifies that the order is odd.
5. Is in doubt about which programs that are available (should have been explained before the test start).
6. TV Danmark 1 and 2 are not found under the Danish channels.
7. Test participant asks how to change the order of the TV-programs. Tries out the help function but there is too much text (do not bother to read) and no sub-contents. Is finally told how clicking on the columns in the program listing can alter the order.
8. Misses a link to the TV-program listing when in the help page. There are only referred to the different pages in the help facility in the left-placed help contents. (The link to the TV-program listing is only present in the top frame).
9. Is surprised about the check box "show hidden programs" when not logged in. The reason is that system has registered him as logged when he actually has not logged in (because of a non deleted cookie).
10. Hits the user profile button.
11. Is in doubt whether he is on the front page, which is the situation.

12. Clicks on "Get user profile" and nothing happens. There is no feedback and no indication of the works of this button.
13. Goes to the channel user profile.
14. Makes some marks on the interesting TV-programs and goes on to enter categories and do the same.
15. Hits "Save user profile".
16. Misses an indication that says that the user profile is saved.
17. Thinks that the page shown after saving the user profile should be the TV-program listing.
18. Mentions the cookie problem when more users use the same computer.
19. Remarks that the filter has changed to "My channels" and "My categories".
20. Clicks on "Show short description".
21. Do not know how to prioritise (What gets high and low priority?). He is told which is high and which is low and starts evaluating.
22. Does not click on the TV-programs he wants to see – makes only negative feedback.
23. The 6 hours intervals are problematic – what if one wants to see what is on TV from 19:30 – 23:00?
24. Misses a search feature.
25. The categories column is placed far most to the right and should be more to the left. The category is used by the test participant to determine what to see and what not to see.
26. Would like if the order of the columns were user defined.
27. When does a program end – this is not easy to see in the program listing.
28. Misses an overlap indication (some indication when two or more programs overlap).
29. Misses an indication of how long a program has been going on in order to evaluate whether it is possible to watch an already started program.
30. Misses some advanced filtering (e.g. only want to see TV 2, DK4 and Discovery).
31. Would like to have more than one profile per person.
32. The system could make suggestions for e.g. a whole family from some summarized user profile.
33. Tries out "Show hidden programs".
34. Feels trapped by the drop down menus. Would like some advanced feature.
35. The ordering by colours is not intuitive (green, yellow and red).
36. The differences in colour are minimal (the test participant is colour blind).
37. "Create user". Works all right. Should go directly to the user profile instead of back to the program listing.
38. Creates a user profile and when finished ("Save user profile") the participant is not put back to the program listing as expected. ("Front page" is a misleading term).
39. Selects "Show short description".
40. Evaluates a number of programs.

41. Are in doubt about the effect of the evaluation (no feedback and is not told what they can be used for).
42. A reprise of a high prioritised program appears later in the week (not good because the participant probably already has seen the program at that time).

Test person: #2

Test monitors: Kenneth Kammersgaard and Peter Michael Jensen

Date: 20th of February 2001

1. Tries out some different filters.
2. Finds out the feature of title ordering (column click).
3. Logging in by creating a user.
4. Problem that the user is put directly back to the program listing.
5. Goes to the user profile. Too much text again.
6. Selects some programs.
7. Insecure whether the information disappears when hitting "Save user profile".
8. Selects categories.
9. Does not understand the term "front page".
10. Are in doubt about the terms "priority" and "AI". But finds out quickly about the evaluation principle.
11. Do not know what happened when evaluating a program (it disappeared!)
12. First thought was that the order of the evaluation was reversed.
13. Evaluates different programs.
14. Selects "DR 2" in filter.
15. The test participant is not sure where he is in the list (misses some indication).
16. Finds out what "Show hidden programs" does.
17. The test participant mentions the possibility that a user might want to see something completely different some other day (dependent on the time of day, the time available, the mood of the user etc.)
18. Does not use the "Might see" evaluation.
19. After entering the evaluation the test participant do not know what to do.
20. Misses a function that marks all non-evaluated programs as "Will not see".
21. The TV-programs listing could be static until the user hits some icon or button (and then starts the evaluation).
22. Misses a back button in the detailed description.
23. Misses feedback from the system (e.g. what has happened after the evaluation).
24. There could be some real time evaluation (some button).
25. There should be some week overview where it is possible to find out what to see for the whole week.
26. There should be some pop-up when some important program is in the TV database.

Test person: #3

Test monitors: Kenneth Kammersgaard and Peter Michael Jensen

Date: 21st of February 2001

1. Do not know what the purpose is and what to do when entering the web page.
2. Maybe a headline for each of the three frames.
3. In doubt about why there are programs when not logged in. The first think that should was setting up preferences.
4. There should be a starting page saying that one should create a user profile in order to get the optimum of the system.
5. Clicks on "create user". The terms could be "wanted user name" and "wanted password".
6. Too much text in general – do not bother to read it.
7. Clicks on "create user confirm" - there is no message saying that the creation went well and what to do next.
8. Is helped to get to the user profile page.
9. There should be a presentation of "step X of X" in the user profile generation. Or some other wizard guide.
10. Reads a bit text on the user profile start page.
11. The "next page" button should be both on the top and in the bottom of the channel and categories user profile pages.
12. Again there is too much text in the pages (user profile)
13. There could be some icons instead of channel language presented by text.
14. Is in doubt about the headlines in the user profile and the menu on the left (are they corresponding or what – they are and it should be more clearly presented).
15. Get and save user profile buttons are always present which they should not be.
16. Thinks that categories should be selected before channels.
17. What will happen when "Save user profile" is hit? It does not say, but the test participant is guessing right.
18. Goes to the TV-program listing.
19. Surprised about the evaluation/training features.
20. The priorities are reversed compared to the opinion of the test participant (negative first).
21. Thinks it is wrong that the programs selected are always present (also when using the filter menu).
22. Does not catch the notepad facility.
23. It should be possible to disable the priority facility.
24. The test participant finds the detailed description of the programs.
25. Undo facility is missing when removing a program with the negative evaluation.
26. Thinks that the training of the system should have a separate interface.
27. Likes the idea of the system suggesting TV-programs.
28. The programs should not jump without informing the user first!
29. Will rarely use negative evaluation.

30. The system should perhaps only show a top 10 of programs each day instead of all non-hidden programs.
31. Do not need a notepad – would rather trust the system.
32. The system should indicate time overlaps in selected programs.
33. A suggesting of separating the interface in three main parts: training, notepad and normal listing.
34. Give a message when no programs are displayed.

Test person: #4

Test monitors: Kenneth Kammersgaard and Peter Michael Jensen

Date: 21st of February 2001

1. Mentions the two scroll bars (one of them seem unnecessary).
2. Think that the filter with the weekdays should be static (always Monday, Tuesday, etc.)
3. The time intervals in the filter are too big – it should be possible to create a user defined time interval.
4. The program duration column and the start time column are easy to mix up.
5. "Create user" - no problems
6. User profile creating – no problems.
7. Is insecure about the "AI" – term.
8. Finds out the column sorting on his own.
9. The evaluation boxes should be reversed (negative first).
10. Surprised that a program disappears when performing an evaluation.
11. Is in doubt where the "Show hidden" is presented, but finds out quickly. (Ok)
12. Tries out to find a film for today by using the filter menu. Works fine.
13. The detailed description popup disappears too quickly.
14. Good feature to make it possible to "View short description".
15. Finds out quickly the evaluation of programs.
16. Thinks that the radio buttons used for evaluation has a "normal" button functionality, which could be hard to understand.
17. There could be a button that forces the system to train according to the evaluation instead of it happening without the user knowing it. The system priority feature is almost totally hidden for the user, which might mean that the user never finds out what the system is able to do.
18. Are not aware of that "My categories" in the filter corresponds to the selected categories in the user profile.
19. The "AI" number could be replaced by a number of icons.
20. Thinks that the middle evaluation radio button is unnecessary.
21. Detailed description is fine.

Test person: #5

Test monitors: Kenneth Kammersgaard and Peter Michael Jensen

Date: 21st of February 2001

1. Tries out the filter and popup menus without logging in everything is found ok.
2. The time indication in the program list misses an ":".
3. "Create user" is performed very quickly without comment.
4. "User profile". Some of the text is read but there is too much and cancels the reading.
5. Selects channels and categories without problems, but is in doubt about the relation of "underholdning" and "series".
6. "Saves the user profile".
7. Asks a question about the priority column, but catches the point immediately after.
8. Should remove all programs with the same title on a day, when removing one of them.
9. Likes the notepad and the evaluation system very much.
10. Primary concentrates on the first columns when finding out the rating of a program.
11. Would like to click on important words in the program description (which then could be added to the user profile).
12. Maybe it should be possible to remove channels from the user profile but doing something on the program listing.
13. Good feature about the hidden programs (removal of irrelevant information).
14. Thinks that the automatic ordering of evaluated programs is good.
15. The time of day should be a part of e.g. the hidden user profile.
16. Thinks it is wrong that the notepad is not manipulated by the filters (is static).
17. It would be a good feature if there were some suggestions dependent on users that have a profile similar to the user.

Test Plan

Heuristic Evaluation II

(version 0.3)

Responsible: kk, lkr, pmj
Date: 4th to 18th of April 2001
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|**Iterative**|Installation]

Heuristic Evaluation

This paper presents the test plan for the heuristic evaluation. The heuristic evaluation is based on [JN1, p. 115-163], [JN2] and [JN3].

Although there is a difference between a user evaluation and a user test [JN1, p. 157] it has been decided to use the same test plan template for both heuristic evaluations and tests. The main difference is that in a test situation the test participant does not evaluate e.g. the user interface directly but the observer or test monitor interprets the participants' responses/actions. In an evaluating situation the participant is responsible for the analysis of e.g. the user interface.

The above differences do not result in any problems in the making of the evaluation plan because the template is usable in both situations due to the specification of every detail in the test/evaluation (this includes the role of the test monitor, if present).

Introduction

This evaluation plan contains all the aspects of the heuristic evaluation. The usability heuristics used in this evaluation are presented in detail in *Appendix B*.

The evaluation plan contains the following items (no changes from the test plan template):

- Purpose
- Objective
- User profile of test participants
- Method
- Accessories
- Task list
- Test monitor role
- Evaluation measures
- Report Contents and Presentation

Furthermore our test plan will include:

- Possible outcomes of the test
- Actions to be taken according to the possible test outcome

The following sections will comment on the different items in the evaluation plan.

Purpose

The purpose of this second heuristic evaluation is to evaluate the user interface due to the usability heuristics. The reason to perform the evaluation is to find the usability problems in the chosen design.

Objective

The objective of the second evaluation is to determine how the user interface performs due to the usability heuristics given by [JN1, p. 115-155]. The user interface is based on the following issues:

- The results of the exploratory tests performed in the initial design phase (questionnaire and observation test).
- The first heuristic evaluation (I).
- The second exploratory test on version 0.2.
- Project group ideas.
- Inspiration from existing products (see also *Appendix J*).
- Ideas given by other people (e.g. MINDPASS A/S employees, supervisor etc.)

User Profile of Test Participants

The test participants used in the heuristic evaluation could in principle come from any of the following categories (used by [JN1, p. 161]):

- Novices with computer knowledge in general, but no special knowledge about usability aspects.
- Single experts who are usability experts, but not specialized in the domain of the interface.
- Double experts who both are usability experts and experts in the kind of interface being evaluated.

The results differ dependent upon which type of the groups that is used for the evaluation. According to [JN1, p. 161] will double experts be best at finding usability problems, while novices find the smallest number of usability problems. This indicates that double experts are to be preferred in heuristic evaluations. In this heuristic evaluation the participants are taken from the "single experts"-group and the "double experts"-group.

The number of participant should be between 3 and 5 due to some cost benefit calculations made in: [JN3]. In Figure 2 the number of usability problems as a function of number of evaluators is presented. It can be seen that the curve flattens out when there is more evaluators than 5 and a number below 3 results in a significant lower number of usability problems found.

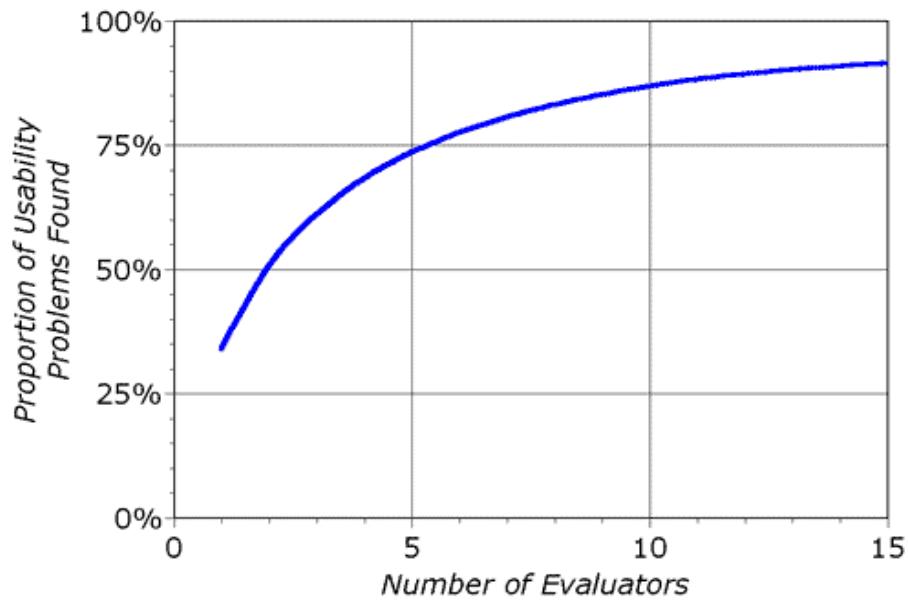


Figure 2: The number of usability problems found as a function of the number of evaluators [JN3].

In the second heuristic evaluation it has been decided to use usability experts from MINDPASS A/S. The primary reason for using the MINDPASS A/S experts in this evaluation is that it has been considered that the project group will have difficulties in making an impartial evaluation at this stage of the development due to the fact that the project group has worked very intensively with the design. It is also considered important to have "new eyes" to look at the design before the validation test is carried out in order to solve as many usability problems as possible before the test.

Method

Jacob Nielsen originally presents the evaluation method used. The method is used to determine usability problems in a user interface so that they can be attended to as part of an iterative design process [JN1, p. 155].

The heuristic evaluation is performed by letting a number of participants evaluate a number of user interfaces according to the usability heuristics given below (described in detail in *Appendix B*):

- Simple and natural dialogue
- Speak the users' language
- Minimize the users' memory load
- Consistency
- Feedback
- Clearly marked exits

- Shortcuts
- Good error messages
- Prevent errors
- Help and documentation

The participants are evaluating the user interfaces individually. The results can be collected in several ways (written by the evaluator, made by an observer etc.) In this evaluation it has been chosen to let the participants make a written document as their contribution to the evaluation. This gives the participants the freedom to decide when they want to perform the evaluation (in a given period). To ensure that the documents are somewhat easy to evaluate and to ensure that the participants remembers all the usability heuristics it has been decided to make an evaluation scheme to be filled out by the participant (one scheme for each user interface). An extract of a filled-out scheme is given below as an example:

Usability heuristic	Comments
Simple and natural dialogue	#1 The sequence of tasks to perform when opening a file is too difficult. #2 There are too many colours on the user interface (makes it confusing). #3 ... #4
Speak the users' language	#1 I don not understand the word "inTelly". #2 ...

When the schemes are filled-out the design team makes a new scheme containing all the usability problems found by the participants. The next thing to be done is to make a prioritising among all the usability problems. The reason is that it might not be possible to correct all the problems listed therefore it is necessary to have an idea of which problems are considered as major problems and which are minor problems. The method used to determine the rating between the usability problems is based on a "Severity Rating" given by [JN3]. The rating is done by sending out all the usability problems to not less than three and not more than five of the original participants, asking them to evaluate each usability problem according to the scale below:

- 0 = I do not agree that this is a usability problem at all.
- 1 = Cosmetic problem only: need not to be fixed unless extra time is available on the project.
- 2 = Minor usability problem: fixing this should be given low priority.
- 3 = Major usability problem: important to fix, so should be given high priority.
- 4 = Usability catastrophe: imperative to fix this before product can be released.

The severity of a usability problem is a combination of three factors [JN3], which should be taken into account by the evaluation participants:

- The **frequency** with which the problem occurs: Is it common or rare?
- The **impact** of the problem if it occurs: Will it be easy or difficult for the users to overcome?
- The **persistence** of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

When the ratings are returned from all the participants the mean is found for each usability problem and they are sorted accordingly. It is now possible to determine which problems that are considered most important to fix etc.

The user interfaces to be evaluated are a combination of WebPages, GUI shells etc. All the user interfaces considered in this evaluation are presented in *Supplement O*.

Accessories

The accessories to be used in this evaluation are all the user interfaces to be considered. Each participant should be given the user interfaces to be evaluated and the same number of heuristic schemes to be filled out. Every user interface should be given a unique identification, which is also used to mark each scheme in order to simplify the collection of the usability problems. Further more the description of the usability heuristics is given to each participant (see also *Appendix B*).

Task list

The tasks to be performed are not given by this evaluation plan, because the tasks to be considered only consist of filling out some heuristic schemes.

Test monitor role

The role of the test monitor is to handout the user interfaces and the evaluation schemes to the participants. This will include explaining that there belongs a specific scheme to each user interface.

Evaluation Measures

The evaluation measures will as already described be the severity ratings given by the participants. As a starting point all usability problems with a rating of 3.0 and above should be corrected.

Report Contents and Presentation

The presentation of the evaluation result will be a list of the means of the usability problems ordered with the major problems at the top of the list.

Possible Outcomes of the Test

The following outcomes are considered in this and the following section:

- The final list of usability problems is very long and indicates several severe problems.
- The list contains usability problems that, if corrected, could result in new usability problems.

Actions to be taken according to the possible test outcome

The actions to be taken according to the above listed results are respectively:

- If the list of usability problems is very extensive it might be necessary to increase the rating indicating that a problem should be corrected (default set to a mean of 3.0). This will be acceptable in the first heuristic evaluation due to the fact that it will be possible to catch the minor usability problems later in the design phase. Furthermore it is possible that the correction of the major problems will result in a more or less automatic correction of some of the minor problems. If the second heuristic evaluation shows severe problems with the design there must be made some decisions about whether it is possible to redesign the user interfaces giving the problems and/or the validation test should be carried out with the problems indicated by the second evaluation. This will strongly depend on the severity of the problems and time available to correct the problems.
- If there are potential new usability problems in the correction of some of the problems in the list it should first of all be considered how to avoid this problem. If the potential problem is considered to be critical, a total change of the UI is a possible solution (acceptable in the first heuristic evaluation, but if present in the second evaluation it might not be possible to redesign the user interface due to time constraints). If it is impossible to avoid some potential new usability problem it should be evaluated whether this problem will be less significant compared to the problem corrected. And if this is the situation it should be corrected anyway.

Test Result

Heuristic evaluation II

(version 0.3)

Responsible: kk, lkr, pmj
Test date: 4th to 18th of April 2001
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|**Iterative**|Installation]

Test Result

In this document there will be a description of the results of the second heuristic evaluation. The first section will contain an evaluation of the results followed by a discussion of some the most severe usability problems and then there is a conclusion on the whole evaluation and its results. Finally there is a presentation of the immediate results.

Evaluated Results

The evaluation of the results will first of all consist of a presentation of the most severe usability problems according to the rated usability problems made by the test participants. The severity of the usability problems is evaluated in the order of the average rating. This means that the high-prioritised usability problems will get the primary attention and the low prioritised usability problems will get less attention. Nevertheless all usability problems listed in the Immediate Results will be taken into account and discussed in the design of the final version of the system (version 1.0).

The most important overall usability problems in version 0.3 are presented in the lists below:

- The navigation on the inTelly web site sometimes gives rise to usability problems. Primary the navigation in the user profile and the browser back button.
- The help facility is not considered to fulfil the needs. The structure and the contents of the help facility are too difficult to understand.
- The user interface presenting the Repeaters seems to give usability problems.

For a complete list of the usability problems found in this evaluation the reader should see Immediate Results. The usability problems will be discussed in the following section.

Discussion

The discussion is divided in the three main areas of the test.

Navigation

It is considered very important, both by the test participants and the project group, that the navigation in the user interfaces are intuitively structured and that they work similar in all situations. The heuristic evaluation showed that the browser back button is not working properly and the navigation in the user profile seems difficult to fully understand (there is a consistency problem because there is a "go-to-next-page-button", but no button that makes it possible to go back). This indicates that this issue needs some considerations in the final version of the system (1.0) to be developed on the basis of

this evaluation. There can be no doubt that the works of an essential function such as the browser back button must work as the user expects.

Help

The help facility still contains too much text and it seems that the complete user interface needs a work over. The solution will include to minimize the amount of text presented to the user and to structure the contents differently: E.g. by the use of more bullet points, more hyperlinks, etc.

Repeaters

The user interface presenting the Repeaters seems to be inconsistent compared to the program list. The overall basics of the two interfaces are similar (presenting a list of programs, which are possible to evaluate) and there should therefore be some consistency between them. There are difficulties in understanding why and how the priority facility works in this user interface. Therefore it will be necessary to find a solution how to present the similarities and differences between the two interfaces in an appropriate way. Another possibility would be to make the user interfaces more or less identical, but this might also give rise to usability problems because there is an increasing risk of mixing-up the two user interfaces.

In general all the suggested changes presented above should include an extensive evaluation of pros and cons in order to avoid severe usability problems in version 1.0. This is extra important in this phase of the design due to the fact that there is only one test (the validation test) scheduled to be performed in the rest of the project period. The validation test will be on the final version of the system and there will not be made corrections to this version. There is always a risk of creating new usability problems, when solving some usability problem.

Conclusion

The Evaluated Results and Discussion presented above has put focus on the most important results of the second heuristic evaluation. This does not mean that these are the only usability problems in version 0.3 of the inTelly system, but they are indicating the major problems in the design and further they are representative to the type of usability problems that are found in version 0.3. As already stated the uncommented usability problems will also be considered in the design of version 1.0.

The results are considered to be useful and version 1.0 will be based upon them. The list of usability problems is relatively long and the time constraints will naturally be a factor in the decision of what to improve/change and what to ignore in the final version.

Immediate Result

Below the immediate results from the second heuristic evaluation is presented. It contains all the comments from the three test participants together with their rating of the problems and the average of their rating.

UI : User Interface number
 He : Heuristic
 #1 : Test person #1
 #2 : Test person #2
 #3 : Test person #3
 Av. : Average

User interfaces

- 1 : Create user and login.
- 2 : User profile.
- 3 : Program list.
- 4 : Detailed description of a program.
- 5 : Repeaters.
- 6 : Help.
- 7 : All user interfaces in general.
- 8 : General comments.

No	UI	He	#1	#2	#3	Av.	Comments
1	2	9	4	4	4	4,0	Missing back function.
3	3	2	4	4	4	4,0	It is not clear what "Installer XML" means
4	7	1	4	4	4	4,0	Popup message dissapear to quickly.
5	7		4	4	4	4,0	The primary concept of the tv-guide is not clear.
6	7		4	4	4	4,0	Back button in browser, does not work properly.
7	1	8	4	3	4	3,7	Error message when failed login is not removed when trying login later.
8	2	4	4	3	4	3,7	"Gem brugerprofil" and "Hent brugerprofil" is disabled when written some text.
9	2	5	4	3	4	3,7	I would like to know where the userprofile is saved (at user computer or at server)
17	5	1	4	3	4	3,7	It is not clear how to delete a Repeater.
11	7		4	3	4	3,7	It should be possible get the last message repeated.
12	1	9	4	2	4	3,3	It is possible to create usernames with æøå, ", "Skriv dit navn"
13	2	1	4	2	4	3,3	It would be nice if the userporfile were saved automatically.
14	3	3	3	4	3	3,3	The graphic representation of the time is difficult to relate to the program data.
15	3	4	4	3	3	3,3	No feedback when prioritising.
16	3	10	3	3	4	3,3	When searching the short description disappears.
21	5	1	3	4	3	3,3	It is difficult to see which title corresponds to which time and channel.
23	5	4	4	2	4	3,3	It is not clear that it is possible to deselect Repeaters because of the term <i>Abonner</i> .
19	7	3	4	3	3	3,3	Popup message makes funny things, jitter, reapear, etc.
20	2	3	3	3	3	3,0	Repeaters is not part of userprofile which ruins the userprofile layout.

44	5	1	3	4	2	3,0	Bad term "Repeaters"
25	6	10	3	3	3	3,0	Navigation picture in help is blurred.
26	7	3	3	3	3	3,0	It should be possible to disable the popup messages
27	1	5	3	3	2	2,7	Username should not be default "Skriv dit navn"
28	2	2	3	2	3	2,7	Too much technical info.
29	2	4	3	2	3	2,7	No consistency between menu and Create user wizard.
30	3	1	2	4	2	2,7	Bad term "AI"
31	3	3	2	4	2	2,7	Is popup message saying "3. Brug TV-guiden intelly.dk" necessary?
32	5	3	3	3	2	2,7	Missing detailed description of each repeater.
46	5	10	3	3	2	2,7	There is no help to the prioritise functionality at the repeaters list.
34	6	2	2	4	2	2,7	Under "System-beskeder" the language is bad.
49	6	3	3	2	3	2,7	Change layout to more bullets, paragraphs, etc.
35	1	3	2	3	2	2,3	It should be possible to get a mail with forgotten password.
36	2	1	2	3	2	2,3	Email is presented to soon, maybe present under text.
37	2	1	2	3	2	2,3	"Gem brugerprofil" at the Repeaters page should be above the Repeater link.
38	2	1	2	3	2	2,3	Too much information, some possibilities off screen due to too much text.
39	2	5	2	3	2	2,3	Repeater link is very slow.
40	3	4	2	3	2	2,3	When "Kort beskrivelse" is selected the popup description should be turned off.
41	3	10	2	3	2	2,3	Why is "Installer XML" present here
42	3	10	2	3	2	2,3	No text on printer icon.
43	4	4	2	2	3	2,3	Prioritising is not same layout as in programlist.
61	5	1	2	3	2	2,3	It is difficult to intuitive know that the title is a link.
62	5	3	3	1	3	2,3	It is not clear why it is D22possible to prioritise Repeaters
22	5	3	2	2	3	2,3	Bad term "Abonner"
45	5	4	2	3	2	2,3	Repeaters list do not automatically update the user interface like the program list.
24	5	5	2	3	2	2,3	Missing popup description of the prioritising radiobuttons.
47	6	1	2	3	2	2,3	Two first sections of help frontpage should be under Create user.
50	6	4	2	3	2	2,3	Missing a next-button.
51	6	10	3	2	2	2,3	Programlist help too complex.
52	7		2	3	2	2,3	There should be context dependent help available.
53	1	2	2	2	2	2,0	There can be some confusion about if "Afbryd" means "Fortryd"
54	1	5	2	3	1	2,0	It would be nice with a popup message when the user is just created.
55	1	10	1	3	2	2,0	Remove silly text from popup message "Det er da nemt :-)"
56	2	1	1	4	1	2,0	Bad language and spelling errors.
57	2	1	2	3	1	2,0	Frontpage of userprofile confusing.
58	3	1	1	3	2	2,0	Bad term "Prioritet"
59	3	1	2	2	2	2,0	Bad term "Vis skjulte"
60	3	3	3	3	0	2,0	The programlist is too long
33	6	1	2	2	2	2,0	Help navigation seems misplaced.
48	6	1	2	2	2	2,0	Missing grouping in help menu.
63	6	4	2	2	2	2,0	Missing popup in help menu.
64	6	10	2	2	2	2,0	Programlist help figures missing popup.

65	3	1	1	3	1	1,7	Annoying that the date is presented before the time.
66	3	1	1	3	1	1,7	The date is not necessary when only one day is showed.
67	3	8	1	3	1	1,7	Bad message "Hvis der ikke vises nogen programmer, kan det være..."
68	4	1	2	3	0	1,7	Interface is confusing, very wide.
69	4	1	1	3	1	1,7	More space between "Abonner"-button and program data.
70	4	11	1	3	1	1,7	Missing additional information about actors, instructors, etc.
71	5	3	3	2	0	1,7	Repeater list too long
72	5	6	2	3	0	1,7	The back button is placed far down the page.
73	6	10	1	3	1	1,7	Intro text in help is not help
74	6	10	2	1	2	1,7	Missing help wizard/task oriented help.
75	6	11	1	3	1	1,7	Missing search function in help.
76	7		3	2	0	1,7	Separate training of intelligence from notepad selection.
77	1	8	1	2	1	1,3	Remove "!!!!" in error message.
78	2	1	1	2	1	1,3	Bad term "Brugerprofil"
79	2	1	1	2	1	1,3	Bad term "Data"
80	3	1	1	1	2	1,3	Bad term "Avanceret" in filter
81	3	4	2	1	1	1,3	"Kort beskrivelse" is more than short description.
82	3	10	1	2	1	1,3	I do not need "Antal programmer i alt: 235"
83	4	4	1	2	1	1,3	Wrong font
84	2	9	2	0	0	0,7	Is checkbox on email necessary?
85	8					0,0	Good feature with navigation screen.
86	8					0,0	"Abonnér" is a good feature.

Table 1: Usability problems – heuristic evaluation II.

The table below presents an overview of the possible actions that could be made on the different usability problems. Most of them will be corrected in the final version of the system.

No	Action
1	Put in buttons with text.
3	Remove.
4	Make it possible to see last message.
5	Make a cartoon illustrating the principles and make it a part of the creation of a new user.
6	Make unique page ID's.
7	Correct.
8	Automatically saving of user profile.
9	Add in help.
17	Add in help.
11	Make it possible to see last message.
12	"Skriv dit navn" should be removed and the comment should be removed.
13	Automatically saving of user profile.
14	Make a box around the program and graphics.
15	Correct the error.

16	Put all XML in one frame when initialising.
21	Add category.
23	Rename link to "Automatisk prioritering" but keep the Repeaters in text.
19	Correct the error.
20	Remove from user profile.
44	Rename link to "Automatisk prioritering" but keep the Repeaters in text.
25	Make the picture clearer.
26	Will not be corrected because they are only present for the first time of use.
27	Remove text.
28	The text will be rearranged.
29	Fix menus and insert text in buttons.
30	Will be explained in the cartoon.
31	Will be removed and replaced with a link to the cartoon.
32	Will not be implemented.
46	Make popup like in the program list.
34	Correct and update it (use more bullets, links, etc.)
49	Correct and update it (use more bullets, links, etc.)
35	Will not be implemented.
36	Will be considered
37	Automatically saving of user profile.
38	Work over and correct.
39	Fetch data directly from database.
40	Turn of popup when short description is selected.
41	Remove.
42	Make popup.
43	Will not be corrected.
61	Will not be corrected.
62	Rename link to "Automatisk prioritering" but keep the Repeaters in text.
22	Rename link to "Automatisk prioritering" but keep the Repeaters in text.
45	Update automatically and maybe insert checkbox to disable automatic update.
24	Make popup like in the program list.
47	Correct.
50	Will not be corrected.
51	Correct and update it (use more bullets, links, etc.)
52	Consider making it.
53	Correct.
54	Consider making it.
55	Remove text.
56	Correct.
57	Correct it.
58	Will not be corrected.
59	Will not be corrected.

60	Will not be corrected.
33	Will not be corrected.
48	Will not be corrected.
63	Add it.
64	Add it.
65	Consider making it possible to select columns in the user profile.
66	Correct it.
67	Will not be corrected.
68	Make table.
69	Fix the layout.
70	Will not be made.
71	Make filters.
72	Will not be corrected.
73	Remove.
74	Correct and update it (use more bullets, links, etc.)
75	Will not be implemented.
76	Will not be implemented.
77	Remove.
78	Will not be corrected.
79	Remove.
80	Will not be corrected.
81	Make a popup description.
82	Will not be corrected.
83	Fix all fonts to be the same.
84	Will not be corrected.
85	Only comments – not usability problems.
86	Only comments – not usability problems.

Test Plan

Validation test

(version 1.0)

Responsible:

kk, lkr, pmj

Test date:

9th to 10th of May 2001

State:

Done

Project:

inTelly.dk – an intelligent TV-guide

Phase in project:

[Pre-analysis|Initial|Iterative|Installation]

Test plan

This test plan will present the validation test that will be performed on the final version of the system (1.0).

Introduction

The test plan contains the following items, which are taken from the test plan template:

- Purpose
- Objective
- User profile of test participants
- Method
- Accessories
- Task list
- Test monitor role
- Evaluation measures
- Report Contents and Presentation
- Possible outcomes of the test
- Actions to be taken according to the possible test outcome

The following sections will comment on the different items in the test plan.

Purpose

The purpose of the test is to evaluate the usability of the final version of the inTelly system (1.0). The usability goals will form the basis for this test.

Objective

The main objective of this test is to test and evaluate the usability goals presented in *Main report - 14. Testable goals*. It should be noticed that the evaluation and test of the functional requirements and the performance goals are not presented in the Test Report. The reason is that these tests will not involve test participants and a test plan made from the Test plan template will therefore be inappropriate.

User Profile of Test Participants

The user profile of the test participants will be advanced Internet users. The advanced users that gave their email address in the Questionnaire are contacted and asked to participate in the validation test.

Method

The method of testing the usability goals will be by letting some potential end users tryout the system more or less on their own. The test participants will be sitting in a room where their acts and reactions are recorded on videotape for later evaluation.

The method used in this test will be to ask the participants to:

- Solve some given tasks on the inTelly system version 1.0.
- After using inTelly: The participants are asked to fill out an evaluation scheme that contains some statements concerning the acceptability and enjoyableness of using the system. The statements are taken from the usability goals listed in *Main report - 14. Testable goals*. The Likert scale is used in the evaluation.

The following list will describe the outline of the test:

- The user should be informed about the overall purpose of the system: It is a personalised TV-guide.
- The method will be that the test monitor hands out the task list given in section Task List and user should perform the tasks one at a time in the given order. The user should think aloud during the use of the system.
- The test participant should be debriefed, about his/hers experience in using the inTelly system by filling out the Evaluation Scheme.

The test monitor must make it clear that the tasks and questions are made in order to evaluate the performance of the inTelly system and not the participant. This means that the participant should not be frustrated or stressed when some task is hard to get right (this only indicates a problem in the inTelly design). The duration of the test will be approximately 30 minutes. This duration is considered appropriate, avoiding that the participant becomes impatient. The order of the tasks and the statements in the evaluation scheme will be the same for all participants. There has been made pilot tests in order to evaluate the test form and make corrections to this test plan where appropriate.

Accessories

The accessories to be used are:

- The inTelly version 1.0 (this includes web-files, Agent Framework, databases, jdk1.3, Tomcat 1.2, Java Servlets, MSXML version 3.0 installed on the browser).
- A computer with an Internet connection and a browser installed.
- A list of tasks to be performed by the participants.
- An evaluation scheme to be filled out by the participants after he or she has tried out the system.
- The test environment will be a home environment.

- There will be a camcorder in the room in order to record the user when he or she is using the inTelly system. The camcorder will be placed so that the face of the participant is recorded in order to evaluate the attitude of the participant when performing the tasks on the system. The utterances of the participant will also be recorded by the camcorder. This will make it possible to review the comments said by the user when he or she is thinking aloud during the test.
- There will be used a screen capture software (Camtasia version 2.1.0) to record the screen during the test. This makes it possible to review the all the actions performed by the user on the user interface. The sound is also recorded by the screen capture software, which makes it possible to synchronise the video recording and the screen recording after the test.

Task list

The tasks the participant should go through are listed below. The arguments for the different tasks are given for each task after the task.

1. Find out the TV-programs send on TV3-Denmark tonight. The user should perform this simple task in order to feel comfortable by starting out easy. Furthermore this task will show how intuitively the user interface is to use for the first time user. It should be noticed that the user is not logged in at this task.
2. Create a user on the inTelly web site. This is the first step for a new user that wants to personalise the TV-guide. The task is considered simple.
3. Create a user profile on the inTelly web site. The user profile is an important feature of the system. The user profile has shown to be an important factor in the evaluation of an Internet based TV-guide (see also the Observation test). The usability goals (*Main report – 14. Testable goals*) stated that the creation of a user profile must maximally take 5 minutes with a maximum of 1 error for min. 70 percent of the users.
4. Find out what to see on TV tonight. The purpose of this task is to find out whether the TV-guide fulfils the usability requirements stated in the *Main report – 14. Testable goals*. This will include both time and error measures.
5. Is there anything so far there is difficult to understand – Explain. It is considered important that the user states if there are difficulties in understanding the TV-guide.
6. Prioritise 9 TV-programs with more than one line of description equally distributed in the different categories: "Want to see", "Maybe want to see" and "Do not want to see". This task is used to find out how well the user prioritising and integrated notepad performs. The task is explicitly defined because the user in a later task should evaluate the principle in the system prioritising.
7. Find out what automatically prioritising means ("Automatisk prioritering"). This feature is also one of the new features in TV-guide and it is considered important the users are able to determine the overall principle.

8. Try to prioritise some TV-programs in the list presented in the automatically prioritising. This task is meant as a supplement to the previous task.
9. Lets say it is the next day and therefore go to the program list and select the day of tomorrow in the program list filter. The system has at this state evaluated the user priorities given in task 6 and the AI column is updated accordingly. The reason for this task and the following tasks is to get some feedback on the system priority feature.
10. Sort the list according to the AI column. See task 9.
11. Select "Show hidden". See task 9
12. Comment on what has happened. See task 9

When the user has performed the tasks above he or she is asked to fill out the evaluation scheme presented below:

Evaluation scheme

Please put only one cross in each row in the statements below.

1 : Strongly disagree
 2 : Disagree
 3 : Neutral
 4 : Agree
 5 : Strongly agree

Statement	1	2	3	4	5
1) The TV-guide is comfortable to use					
2) The TV-guide is fun to use					
3) The TV-guide is intuitively to use					
4) The layout of the user interface is nice					
5) Each page of in the user interface is well arranged					
6) It is easy to navigate in the TV-guide website					
7) The TV-guide offers an effective way of finding out what to see in TV.					
8) The TV-guide helps me in remembering what to see in TV					
9) This is a TV-guide that I would use in the future if possible					

Comments

The statements in the evaluation scheme are primary taken from the *Main report – 14. Testable goals* in order to be able to evaluate the usability goals.

Test monitor role

The role of the test monitor is to keep out of focus in this test. If there are severe problems for the test participant the test monitor can and must help the participant.

The role of the test monitor will be the following:

- The test monitor should ensure that the test participant accepts that the session will be recorded and used in the evaluation of the test.
- The test monitor reads up the text presented in the briefing below. This briefing ensures that the test monitor remembers to give the test participant the relevant information and that all test participants gets the same information.
- The test monitor will sit behind the participant and will see to that the test progresses as planned.
- The test monitor should ensure that the browser is at the inTelly start page and that the recorders are running before the start of the test.
- After the test participant has used and performed the necessary tasks on the inTelly system the test monitor handout the evaluation scheme presented in the previous section.

Briefing

The test monitor will read out the following text to each test participant before the start of the test:

First of all you be aware that we are testing the system performance and not you.

The total duration of the test will be approximately 30 minutes. A camcorder and a screen capture tool will monitor the test and your utterances if you agree to it.

You should think aloud when you are using the system and tell when there is something you do not understand. You should just "talk to the monitor" and explain why and what you are doing.

You are free to use the help in the system if you are in doubt of anything. You can also ask me if the system fails to help you.

It should be noticed that there is only a limited number of TV-programs available and the TV-program categories are not working properly.

There are a number of tasks that you should perform. You should tell when you think that you have performed a task.

You should think that you sit at home and is connected to the Internet and want to find out what to see in TV. You have just found this TV-guide and have just entered the start page.

Evaluation Measures

The evaluation measures will primary be the video recordings, the screen recordings, sound recorded and the filled out evaluation schemes.

Report Contents and Presentation

The outcome of the validation test will be a discussion of each of the subtest performed on the inTelly system version 1.0. There will also be an overall conclusion on the validation test.

There will also be an overall presentation of the good and bad things about the final inTelly system plus suggestions to future versions (additions and improvements).

Possible Outcomes of the Test

Two possible outcomes of the test are:

1. The concept of the new facilities in the TV-guide is not clear (personalization and integrated notepad).
2. The system fails to pass the usability goals (e.g. more than 30 percent of the participants cannot generate the user profile in 5 minutes).

Actions to be taken according to the possible test outcome

The corresponding actions to be taken are not actions that will be performed in this project, but will/could be taken in a future version of the system. The reason is that in this project the system is considered finished in version 1.0 and there will not be made corrections to this version.

1. If the concept of the new facilities is still not clear in version 1.0 of the system then it should be considered whether it is possible to make the concepts clear before a final release of the TV-guide. If this is considered impossible then the release of the system should be reconsidered. It would not be appropriate to release a system that is not clear to understand for the future users. It should be reviewed whether this situation could have been foreseen at an earlier state of the development and use this experience in the future.
2. If the system fails to live up to some of the usability goals then the result must be evaluated. It must be considered whether the result is acceptable anyhow. This could be the situation if the result is relatively close to the 70 percent demand. If this is not the case then it must be considered how correct the problem(s) before

the release of the system. This estimation will depend on the specific result and the possible solution to the problem.

Test Result

Validation test

(version 1.0)

Responsible: kk, lkr, pmj
Test date: 9th and 10th of May 2001
State: Done
Project: inTelly.dk – an intelligent TV-guide
Phase in project: [Pre-analysis|Initial|**Iterative**|Installation]

Test Result

The result of the validation test will be presented in this document. The first section will present the evaluated results. In following section there is a discussion of the evaluated results. The third section will conclude on the validation test. The last section will present the immediate results. This will include a summary of the most important events registered concerning the users performing the tasks given in the Test plan, a presentation of the evaluation schemes filled out by each user. There will be references to video clips located at the project CD-ROM where there has been observed especially important events during the test (both positive and negative events).

Evaluated Results

In the validation test six participants were asked to perform some specific tasks on the inTelly system version 1.0 and to fill out an evaluation scheme. The objective of the test is to evaluate the usability goals presented in the *Main report – 14. Testable goals*.

The following comments are based on the results presented in Immediate Result.

- Most of the users (5 out of 6) catch the primary concept of the new features presented in the system. See Videos 1,2,3,4,7 and 8.
- Most of the participants do not know what to do after they have created a user profile. But the other navigation in the user profile works very well. See Videos 9 and 10.
- The principle in the Repeater list is somewhat difficult to completely understand. (The effect of prioritising programs in this list is not immediately understandable). See Videos 5 and 6.
- The principle in that some of the programs are hidden is easy to understand for the participants. See Videos 7 and 8.
- During the test some minor errors in the system where discovered. See Video 11.
- It was easy and quick for the participants to use the help in the system. But one participant would like a context sensitive help. See Videos 14 and 15.
- The column containing the system priority (AI) was not immediately understandable. See Video 15.
- The speed of the system was generally considered to be slow.
- The facility of prioritising the TV-programs was in several situations immediately understandable without any introduction (neither the general help nor the program list introduction). See Videos 2,3 and 4.
- Only 67% of the participants were able to find out what to see in TV for one day in less than 3 minutes. The usability goal was a percentage of at least 70.
- All statements in the evaluation scheme have passed the test (83% or above have answered “neutral”, “agree” or “strongly agree” to all statements).

Discussion

In this section there will be a discussion of how the comments listed in the previous section can be interpreted.

The usability of new facilities in the TV-guide is considered to be of major importance. The participants were all able to use these facilities after trying out all the tasks. It should be noticed that there is still some doubt about the possibilities and the works of the features at the beginning of the use. One of the test participants used the introduction to the TV-guide and was immediately familiar with the principle in the integrated notepad and the possibility of prioritising the programs (see Video 1). This could indicate (and only indicate!) that a solution to the problem could be to make the introduction more visible to new users. In version 1.0 the introduction is step three in the create-user-wizard, but can also be found in the help system. In the test only one person actually read the messages in the wizard and this is not satisfactory. It should be noticed that the test situation is an artificial condition and it is likely that more of the end users will read the wizard messages due to the fact that in a normal use there will not be handed out a predefined list of tasks that the user must try to perform. One reason for the missing use of the wizard and messages in general could therefore be that the test participants primary concentrate on solving the tasks instead of reading the wizard properly.

There was discovered one major usability problem when the user has created the user profile and was in doubt about what to do next (see Video 9). This would be an important thing to correct (e.g. adding a button/link to the program list) before a release of the product, both because 5 out of 6 users had problems in this situation and because the users were seriously in doubt about what to do. The buttons and menu worked fine (see Videos 9 and 10).

The Repeater list was somewhat difficult to understand. It was not clear for several of the participants that it is possible to prioritise repeating TV-programs in this list and that it will effect the program list (see Videos 5,6, 7 and 8). This should be stated clearly in a release version of the system. One problem might be that the Repeater list and the program list are almost identical and the user might therefore not be aware of the differences. A solution could be to make them distinct and in that way indicate to the user that the two lists do not work in the exactly same manner. This solution could although disturb the consistency in the web site. Another solution might be to show an additional explanation and/or illustration when the user enters the Repeater list for the first time. But too much text has showed to annoy the user (see also the Assessment test).

The errors that were discovered during the test should naturally be corrected before a release of the system.

The help in the system has proved to work satisfactory (see Videos 14 and 15). But one participant suggested that the help could be context sensitive and this should be considered as an improvement of the system. This should be considered in a release version of the system. The navigation in the help system has proved to be very effective and the users quickly found what they needed.

The speed of the system was generally considered to be slow. One participant suggested that the system should run on a computer with a faster Internet connection. This will not solve the problem because the primary reason for the lack of speed is the speed of the computer. The task of sorting e.g. the program list is a demanding task and a solution could therefore be to use a more powerful computer (than a PII – 266MHz with 392Mb Ram). It should also be noticed that one main reason for the relatively slow system speed in the test is that the screen capture software (Camtasia version 2.1.0) uses a significant part of the system resources. This will naturally not be a problem in a normal use of the system.

Only 67% of the users did find out what to see in TV for one day in less than 3 minutes. This means that the corresponding usability goal was not reached. This could be considered as a major usability problem, but the two participants that used over 3 minutes was relatively close to the demand (3:45 and 3:27). Therefore it is not considered to be a usability problem. It should also be noticed that this duration would be minimized when the TV-guide personalises the TV-programs for the user when he or she has prioritised some TV-programs in the programs and Repeater list.

The result of the evaluation scheme indicates that the system is successful concerning the acceptability and enjoyableness. It has nevertheless been decided to comment on statement 2, 3 and 9, where one user in each statement disagrees about the statement:

- 2: The TV-guide is fun to use. It could be discussed how the TV-guide could be made more fun to use in order to meet with this demand. But one explanation to the negative result could be that the test participant did experience a few errors during the use of the system: The user did accidentally drag the mouse in two situations where the system changed the size of the frames and jumped to a detailed description. These two situations did point out two errors in the design that must be corrected before a release of the system. The negative evaluation might be a result of these errors, which naturally must annoy any user.
- 3: The TV-guide is intuitively to use. If the users do not think that the TV-guide is intuitively to use, then the system has failed one of the major usability requirements. This is especially important in web design, because users of Internet services generally tries out new web sites shortly and if they fail to fulfil the user's needs the will go away and never went back. One solution (as already mentioned above) could be to make the introduction more visible to new users and perhaps more or less force them to see the relatively short introduction, which proved to be a very efficient way to learn the principles in the TV-guide.

- 9: This is a TV-guide that I would use in the future if possible. This statement was not taken from the *Main report – 14. Testable goals*, but was considered to give a hint about the success of the final version of the system. The participant that disagreed in the statement noted that the person never used TV-guides in general and this has been considered to be the reason for the answer in this statement.

Conclusion

The system has passed the usability goals listed.

The validation test although showed that the system is not perfect and corrections should be made before a release of the system. The corrections are more or less included in the above discussion and will not be explained here further. The task of correcting the errors and usability problems are considered simple and it should therefore not be a problem to correct them before a release of the system.

The most important usability aspect is how the user grasps the new ideas in the system. The time taken to understand the new features is considered to be acceptable. But it is not ideal and there should be made adjustments to the design in order to minimize the time taken to learn the new features presented in the system.

Immediate Result

The immediate result consists of four parts. The first part presents clips from the video files recorded in the test; these clips contain some of the most important observations made. The next part presents a summary in text of the most important events observed when the users have performed the tasks. There will be a reference to the video clips located on the project CD-Rom. The next part will present the filled out evaluation schemes. The final part presents the result of the usability goals listed in *Main report – 14. Testable goals*.

Video files

The videos recorded in the Validation test are edited to highlight the most important observations made in the Validation test. The videos are the screen captures with the recorded sound.

The video clips are located at the CD-ROM, in directory:

X:\Validation test

Video_01---Program_list_intro---Testperson_5_2

Video_02---Prioritising---Testperson_6_2

Video_03---Prioritising---Testperson_2_4
Video_04---Prioritising---Testperson_1_1
Video_05---Repeater_prioritising---Testperson_1_5
Video_06---Repeater_prioritising---Testperson_3_6
Video_07---Primary_concept---Testperson_2_8
Video_08---Primary_concept---Testperson_1_6
Video_09---User_profile---Testperson_3_1
Video_10---User_profile_menu---Testperson_4_2
Video_11---Change_frame_size---Testperson_4_1
Video_12---Popup_description---Testperson_4_5
Video_13---Accidently_detailed---Testperson_4_6
Video_14---Help_repeaters---Testperson_4_7
Video_15---Help_AI---Testperson_5_8
Video_16---Sorting---Testperson_6_1

There will be made references to these videos in the form Video 1, Video 2, etc.

Observations

Test person: #1

Date: 9th of May 2001

- 1) Use priorities without having read the Introduction to the Program list. See Video 4.
- 2) Deselects programs without thinking about it.
- 3) Uses the prioritising of the Repeaters immediately and naturally. See Video 5.
- 4) Understands the primary concept. See Video 8.

Test person: #2

Date: 9th of May 2001

- 1) Would like a "Continue" button in the user profile instead of the buttons indicating the name of the previous and next page.
- 2) The test participant is in doubt about the principle in the keywords. Enters TV-program categories and not keywords as intended.
- 3) Misses some information about what to do (and where to go) when the last page of the user profile has been filled out.
- 4) Uses the prioritising of the programs immediately and naturally without looking in the help. See Video 3.
- 5) Ignores the messages in the message box.
- 6) Misses some sort function in the program list. Are not aware that it is possible to sort the list by clicking on the column names in the top of the list.
- 7) Thinks that it is wrong that it is possible to select "Film"-category in the list of Repeaters.

- 8) Understands the integrated notepad, AI and user priority. See Video 7.

Test person: #3

Date: 10th of May 2001

- 1) Misses a button to the program list at the final page in the user profile. See Video 9.
- 2) In the detailed description a prioritised program is not marked in the radio button.
- 3) Ignores the messages.
- 4) Remarks that the speed of the system is slow.
- 5) Is in doubt about the priority and AI columns.
- 6) In the Repeater list the test participant uses the prioritising feature immediately. See Video 6.
- 7) Cannot relate the filter frame to the Repeater list.
- 8) Misses some extra information in the detailed description.

Test person: #4

Date: 10th of May 2001

- 1) Drags the mouse and accidentally changes the size of the frames. See Video 11.
- 2) Uses the menu in the user profile instead of the navigation buttons. See Video 10.
- 3) Goes immediately to the program list after the user profile has been completely created.
- 4) Thinks that the default sorting of the programs in the program list is illogical (sorted after reverse time of day!).
- 5) Thinks that the short popup description is good (much better than visiting a new page). See Video 12.
- 6) Marks a program title (dragging the mouse) and accidentally enters the detailed description of the program. See Video 13.
- 7) Uses the help to the Repeater list and finds out principle that way. See Video 14.
- 8) Prioritises some Repeaters without problems.
- 9) Thinks that the term "Rest of the day" is funny when looking at tomorrow's programs.
- 10) Notifies the comfortable effect selecting and deselecting the Repeaters.

Test person: #5

Date: 10th of May 2001

- 1) Reads the four-step-wizard before starting to create a user.
- 2) Goes directly to the introduction of the program list without creating a user profile (this means that the default values in the user profile is used). See Video 1.
- 3) The prioritising of programs seems natural and the principle is immediately understood.

- 4) Creates the user profile after the test monitor has asked the participant to do it (the result is needed in order to evaluate the user profile).
- 5) Have difficulties in finding out what to do after the user profile has been created.
- 6) Have difficulties in understanding the AI column in the program list.
- 7) Misses some context dependent help.
- 8) Finds help to AI in the help immediately and understands the principle. See Video 15.

Test person: #6

Date: 10th of May 2001

- 1) Uses the column sort facility immediately. See Video 16.
- 2) Uses the prioritise facilities immediately. See Video 2.
- 3) Is in doubt about what happens to programs that are prioritised: "Do not want to see".
- 4) Have difficulties in understanding the principle in the AI.
- 5) Marks a program title (dragging the mouse) and accidentally enters the detailed description of the program.
- 6) Is in doubt about the colour codes.

Evaluation schemes

The immediate results of the evaluation schemes are presented in this section. The evaluation scheme below is repeated from the test plan.

Evaluation scheme

Please put only one cross in each row in the statements below.

1	:	Strongly disagree
2	:	Disagree
3	:	Neutral
4	:	Agree
5	:	Strongly agree

Statement	1	2	3	4	5
1) The TV-guide is comfortable to use					
2) The TV-guide is fun to use					
3) The TV-guide is intuitively to use					
4) The layout of the user interface is nice					
5) Each page of in the user interface is well arranged					
6) It is easy to navigate in the TV-guide web site					
7) The TV-guide offers an effective way of finding out what to see in TV.					
8) The TV-guide helps me in remembering what to see in TV					

9) This is a TV-guide that I would use in the future if possible

--	--	--	--	--	--

Comments

Test person: #1

Statement number	1	2	3	4	5
1				X	
2			X		
3			X		
4				X	
5				X	
6			X		
7				X	
8			X		
9			X		

Comments (translated into English)

The system is easy to use. The speed of the system is rather slow. Good functions such as "AI" that can help you to find out what to see in TV.

Test person: #2

Statement number	1	2	3	4	5
1				X	
2			X		
3				X	
4				X	
5			X		
6			X		
7					X
8					X
9				X	

Comments (translated into English)

It would be nice with faster Internet connection because the speed of the system is slow. Good idea for people like myself that does not subscribe to newspapers and magazines. There could be some reminder function via. SMS/"popup" on PC in order to remember to

watch the TV-programs. Maybe the buttons in create user should be named "Continue", "Cancel" to us people that never gets away from Bill Gates/Windows.

Test person: #3

Statement number	1	2	3	4	5
1				X	
2			X		
3				X	
4			X		
5				X	
6				X	
7				X	
8				X	
9			X		

Comments (translated into English)

Everything is said during the test.

Test person: #4

Statement number	1	2	3	4	5
1				X	
2		X			
3				X	
4					X
5			X		
6					X
7					X
8					X
9					X

Comments (translated into English)

All links should be underlined otherwise it is a bit confusing. AI would probably not make sense to everyone. Beware of using frames. Tomorrow – "Rest of the day"? I would normally use "My channels". The system seems to load the computer.

Test person: #5

Statement number	1	2	3	4	5
1				X	
2			X		
3				X	

4					X
5				X	
6				X	
7			X		
8				X	
9		X			

Comments (translated into English)

I do not normally use TV-guides.

Test person: #6

Statement number	1	2	3	4	5
1			X		
2			X		
3		X			
4			X		
5			X		
6				X	
7				X	
8			X		
9			X		

Comments (translated into English)

None.

Results of Usability Goals

This section will present the results compared to the usability goals listed in the *Main report – 14. Testable goals*.

Usefulness

At least 70 percent of the first time users must be able to find out what to see in TV for one day without using the help facility of the system.

All participants (100%) were able to find out what to see in TV for one day without using the help facility. It should be noticed that the introduction to the program list is not considered to be a use of the help facility, because this page is explicitly suggested as a step in the create user wizard. This means that the users have been strongly encouraged to go to this page.

It should be noticed that some of the participants used the priority facility while others did not use it in this task. This was free to decide for the participants.

At least 70 percent of the first time users must be able to create a user profile without using the help facility of the system.

100% of the participants did fulfil this task without using help.

Effectiveness

The first time users should be able to find out what to see in TV for one day in maximum 3 minutes for 70 percent of the users.

Test participant	Duration
1	2:30
2	3:45
3	3:27
4	1:16
5	0:34
6	1:25

The above table shows that 67% of the participants succeeded in finding out what to see in TV, which is just below the usability goal. The time is measured from the point they enters the program list until they say that they have solved the task.

The creation of a user profile should at maximum take 5 minutes for 70 percent of the users (first time users).

Test participant	Duration
1	2:32
2	7:24
3	1:45
4	1:48
5	4:51
6	1:49

The result is that 83% of the participants have created the user profile in less than 5 minutes, which means that the usability goal has been fulfilled. The time does not include the extra time that many participants used to get out of the user profile. The reason is that this usability has been mentioned in a previous section and therefore taken into account.

The first time users should catch the primary concept of the new features in the system in a maximum of 30 minutes for 70 percent of the users.

The new features are the possibility of prioritising TV-programs, the integrated notepad and the system priorities (AI).

Test participant	Catches the primary concept
1	Yes
2	Yes
3	Yes
4	Yes
5	No
6	Yes

83 percent of the participants did catch the primary concept of the new features of the system. The evaluation of whether the participants have understood the primary concept is based on the participant's actions, their utterances and the filled out evaluation scheme (primary the comment). One test participant did not catch the idea of the AI.

Errors

The task of finding out what to see in TV for one day must maximum result in 1 error for 70 percent of the first time users.

Test participant	Number of errors
1	0
2	0
3	0
4	1
5	0
6	0

100 percent of the participants perform the task with a maximum of 1 error. The usability goal is therefore fulfilled. The error made by participant number 4 is an accidentally marking of a program title with the mouse where the system jumps to the detailed description (this error is described elsewhere in the immediately results).

The user profile must be generated with a maximum of 1 error for 70 percent of the users (first time users).

Test participant	Number of errors
1	0
2	0

3	0
4	0
5	1
6	0

Again 100 percent of the test participants perform the task with a maximum of 1 error. The test participant that makes an error in this task hits the create user button before entering anything in the username and password fields.

Acceptability and enjoyableness

70 percent of the users should be neutral, agree or strongly agree upon statement 1-8 in the evaluation scheme.

The total result of the evaluation scheme is presented in the following table. Both the percentage and the count value are listed. The last column presents the percentage and number of participants that have answered neutral, agree or strongly agree.

Statement	1	2	3	4	5	Total
1	-	-	17% (1)	83% (5)	-	100% (6)
2	-	17% (1)	83% (5)	-	-	83% (5)
3	-	17% (1)	17% (1)	67% (4)	-	83% (5)
4	-	-	33% (2)	33% (2)	33% (2)	100% (6)
5	-	-	50% (3)	50% (3)	-	100% (6)
6	-	-	33% (2)	50% (3)	17% (1)	100% (6)
7	-	-	17% (1)	50% (3)	33% (2)	100% (6)
8	-	-	33% (2)	33% (2)	33% (2)	100% (6)
9	-	17% (1)	50% (3)	17% (1)	17% (1)	83% (5)