Understanding the Needs of Hydrologists

Evaluating a web-based, information-heavy system for domain-experts

Linköpings University (LiU), Linköping, Sweden.

Sofia Thunberg, Isabella Koniakowski, Jasmina Jahic,
Sokrates Lamprou, Charlotte Ho

Abstract.
This paper aims to improve the user experience of HYFO by identifying potential problems and amending these with design propositions. These propositions will the customers' time and mental effort while using the system which allows them to make plans and decisions more efficiently and maximise their profits. To find problems in the system we conducted a heuristic evaluation and a contextual inquiry and from their result we carried out a think aloud test and additional interviews. The findings of study 1 resulted in a paper prototype which we evaluated after a pluralistic usability walkthrough. The findings of study 2 resulted in a list of improvement proposals that will be used in future developments of HYFO.

Keywords: **Information system, usability Hydrological forecast,**

I. Introduction

The Swedish Meteorological and Hydrological Institute (SMHI) is an expert agency under the Ministry of Environment and Energy. The organisation provides information regarding weather, water and climate for private parties, public organisations and offers tailor-made products towards businesses in areas such as environment, energy and agriculture.

One of their products, Hydrological Forecast (HYFO), is a web-based information system which presents hydrological and meteorological observations and forecasts. The system provides weather information such as temperature, water bearing and precipitation in forms of maps, tables and graphs. It is a decision making tool that facilitates hydroelectric power companies to make profitable economic decisions in water resource planning.

At the request of SMHI, this paper aims to improve the user experience of HYFO by identifying potential problems and amending these with design propositions, some of which will be implemented in a paper prototype. It strives to minimise the customers' time and mental effort while using the system. As a consequence, it allows them to make plans and decisions more efficiently and eventually maximise their profits.

Our research questions are:

- Which interface problems exists in HYFO?
- How do the users use HYFO?
- Which interaction problems are there in HYFO?
- Which theories and interaction principles can be applied to solve the problems in HYFO?

In this paper we conducted two studies in order to answer our research questions. The first study was conducted in order to identify existing UI problems and the second study aimed to discover which interaction principles could be implemented to solve the problems.

The figure below shows the existing system:

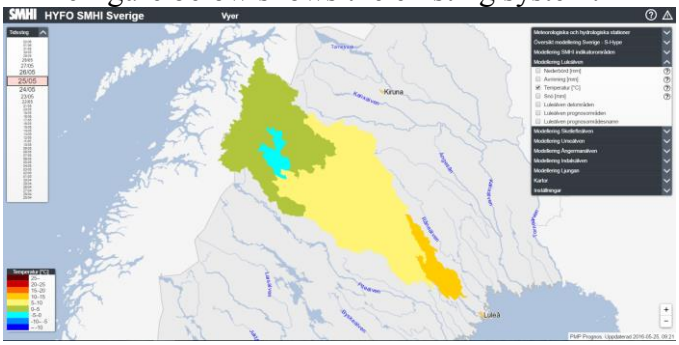

Figure 1. The system HYFO

II. Study I – Methodology

To identify existing problems in the user interface and understand how the system is used we conducted contextual inquiries, a heuristic evaluation and think-aloud tests. Semi structured telephone interviews was later added to further explore the target groups needs of HYFO as well as usability related questions.

Heuristic evaluation

In order to identify problems in the user interface the members of the design team participated in a heuristic evaluation. Heuristic evaluation is described by Nielsen (1994) as a moderately economical and simple usability analysis method for detecting usability errors in a user interface by evaluating Nielsens ten heuristics. Each member inspected HYFO twice to assemble an overview and detailed insights of the system's errors. The resulted in a list of problems connected to the heuristics. In order to assess the severity of the recently discovered user interface problems each issues' significance was estimated on Nielsens five point scale.

Contextual inquiry

This method was selected to examine how and why the users used the system and what problems they encountered in their daily work. This would generate data towards developing a suitable task for the think-aloud test later on. The main goal is to understand the user, how they work, what their needs are and what motivates them in order to create a design that derives from the users themselves. The method uses qualitative data collection and combines semi-structured questions with observations of the users in their natural environments (Beyer and Holtzblatt, 1997). The collected data was later analyzed with an affinity diagram where researchers organised their findings by thematising and categorising through coding. The goal is to get a better understanding of and an insight into the users' activities and purposes for using a service (Arvola, 2014).

Two test subjects were recruited to participate in the contextual inquiry. The first test subject was a female from SMHI's hydrology department whose responsibility is to maintain the system on a regular basis. The second subject was a male from Tekniska Verken, which is one of SMHI's customer companies that employ their information systems in their water resource planning. Both test subjects had at least 6 months of experience of the system. The sessions took place at Tekniska Verken in Linköping and at SMHI in Norrköping.

Think aloud test

In the interest of exploring further what possible problems the users may be facing we conducted a think aloud test to examine and determine users' expectations and thoughts of the system. A *think-aloud* test is a technique where the participants are asked to think aloud while they solve problems in a system (Arvola, 2014). This usually results in written transcripts of the verbalisation's which are called *think aloud protocols*. The method is used to evaluate a prototype that the team has built, or an already existing system. The goal is to find faults and problems in the design, as well as misconceptions that the users may have. A participant voicing their thoughts gives the researchers the benefit of answering *why* these misconceptions happen, and why other parts of the design are clear and easy to use.

Coding was selected to analyze the collected data. The focus in coding is to better understand what each member of the design group has heard and seen on the test to help the rest of the group understand the different interpretations. The part conducted is called single-case analysis, in which the participants constructs categories and code interactions to them individually. They then compare their coding two and two which is called check-coding (Goodwin, 2009).

Before conducting the think aloud test we performed a pilot test sessions on four students from Linköping's University in order to test the feasibility of the think-aloud test. Three females and one male were recruited by convenience sampling. The pilot test resulted in four modifications to the think aloud test. The task needed to be more specific, the participants needed to be better informed during the procedure and informed consent needed to be adjusted. In addition the need for better equipment, such as a higher quality recording program, to collect data was discovered.

Five test subjects were recruited for the think-aloud test whereof three were female and two were male. The sessions were conducted at SMHI and Tekniska verken. The goal of the user was to extract the precipitation data for the current month in Holmsjön.

Interviews

The method was selected after the result from earlier data collection stated that the system might not have actual value to the companies using HYFO. As the companies were too far away to travel to, telephone-interviews were conducted. In order to assure the value of the system we interviewed users from Vattenregleringsföretagen (VRF) (2 females, 1 male) and Sydkraft Hydropower (2 males). The interviews lasted between 10 and 20 minutes and each participant was asked about the use and value of HYFO together with usability related questions.

III. Study II– methodology

Paper prototype

In order to evaluate and present the proposed design a paper prototype was created. There are two ways to generate paper prototypes: creating by hand, or using print screens. Since HYFO holds important information to the users it was viewed as appropriate to use print screen to represent the geographical aspects of Sweden and other components of the system instead of hand drawn pictures (Snyder, 2003). The paper prototype was made in Balsamiq Mockup and had 10 frames. The paper prototype was evaluated through a pluralistic usability walkthrough.

Pluralistic Usability Walkthrough

A Pluralistic Usability Walkthrough is most often used to evaluate early prototypes, as it is a method that requires little money and time (Nielsen and Mack, 1994). The method collects unique insights in how the users think in ways others methods are incapable of capturing. There are three kinds of participants in the walkthrough: representative users, developers and designers. During the walkthrough, all participants take on the role as users and write down on printed screenshots what they would do in each situation of a given task. After everyone has chosen what they would like to do and written down their thoughts, a correct answer to the task is given. As the walkthrough is done on hard copies of the prototype, there is only one “correct” path that is predetermined by the walkthrough administrators. The representative users then describe and discuss their usability concerns and once they have finished talking, the developers are asked to continue the discussion. By combining the opinions of these different people, the designers get a picture of what the problems in the system are and what limits the system may have.

Four test subjects participated in the walkthrough. Two were developers of the original system (males, mean age = 37.5) and the other two were representative users of the system, one a recent recruit and the other an experienced user (females, mean age = 32.5).. The walkthrough was split into two tasks; measuring precipitation in the general area of Ljungan and copying a table with precipitation information about Holmsjön. One of the representative users could not attend to the second part of the walkthrough.

IV. Results study I

Heuristic evaluation

The most significant errors from the heuristic evaluation is listed below.

1. VISIBILITY OF SYSTEM STATUS	
4	Too little notification when there is an update.
4	It is unclear when the system updates, the update date and time does not correlate to the update button.
2. MATCH BETWEEN SYSTEM AND THE WORLD	
4	Warnings should be more salient and noticable. This could be fixed with a notification and/or colour.
3. USER CONTROL AND FREEDOM	
4	It is not possible to press the name of an area and move to there on the map. This is only possible by choosing which information that should be displayed.
4	To choose display options you need to press the word or the checkbox, not the whole row.
4. CONSISTENCY AND FREEDOM	
4	No difference in which graph is shown first depending on what is being displayed on the map.
4	No differentiation between checkboxes and radiobuttons.
9. HELP USERS RECOGNISE, DIAGNOSE, AND RECOVER FROM ERRORS	
4	The error message "there is an error" is not sufficiently informative. It occurred because the computer lost internet connection, and then gained it again.

Figure 2. Table of the errors found in the heuristic evaluation that were rated 4.

Contextual inquiry

The contextual inquiry resulted in two separate affinity diagrams with eleven categories each. The words under each category corresponded to some part of positive or negative activity that the user did in the system. The categories for each diagram were: header, menu, map, timeline, legend cursor, help guide, graph, table, customer and user. The first affinity diagram showed the results from SMHI.

Some of the keywords associated with the categories were attention, quick overview, warning, season, transparency, descriptions, automatic, no shortcuts, surveillance, transparency, drag to compare color, use of arrows, weekly check, not interactive, experience and system standby. The second affinity diagram showed the results from Tekniska Verken and some of those keywords were update, monthly rainfall, restrictions, colour confusion limitations in time, transparency, lack of representation, accumulated values, copy to excel, FTP, long range forecasts and less reliable than web. These keywords were then used to create the tasks that were later used.

Think aloud test

The results from the think aloud tests that contributed to prototyping are categorised into four groups: graph, map, menu and timeline. It was found that the graphs did not match the selected variables in the menu, meaning that the same graphs were shown no matter which options in the menu were chosen. In the graph legend, one needed to click on the text instead of the icon in order to make a variable on the graph disappear. One suggestion was to share the functionality of text with the icon. One fundamental problem that was common among the participants was that they could not notice what was shown on the map. A search function and/or drop-down list was recommended by participants so that users could find what they needed effectively. In the menu, all the options looked like check boxes, even though some of them had the function of radio buttons. The timeline was also suggested to be redesigned, since the original timeline was hard to understand and was found illogical.

Interviews

The interviews showed that all the participants thought HYFO was useful for the decision making in their line of work.

V. Results Study II

Paper prototype

The paper prototype resulted in 12 frames and three of them are presented below. Figure 3 displays the opening screen for HYFO. The loading icon is located in the lower right corner. The “Vyer” menu is located in the menu bar on the right side and the timeline has been rearranged as a horizontal timeline. Figure 4 displays a selected model for precipitation over Ljungan. One new feature is to reach this state is by using the search function situated on top of the menu bar. The values in the menu bar have also been divided up into checkboxes and radio buttons. Figure 5 displays opening a graph over Holmsjön and checkboxes have been inserted in order to clarify the options for the graph.

Figure 3. Shows the menu.

Figure 4. Shows the area Ljungan.

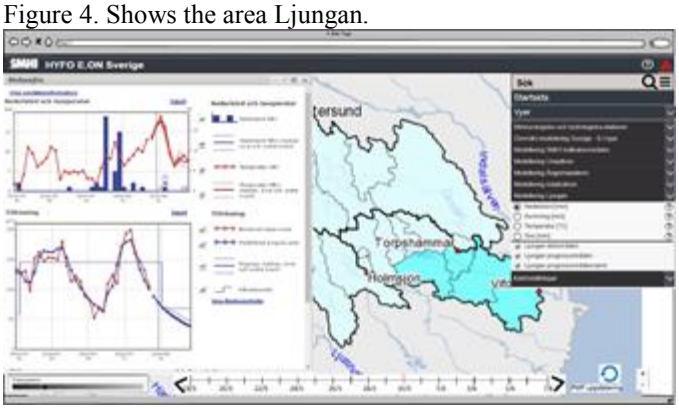

Figure 5. Shows a graph over Holmsjön.

Pluralistic usability walkthrough

The pluralistic usability walkthrough resulted in confirmation of the validity of certain elements that had been added to the prototype and some new areas of improvement. The result is based on discussions and on the questionnaires that contained estimation questions and open questions. The estimation

questions resulted in figure 6. and figure 7. The figures depicts two groups: developers and users. The developers were more positive towards the system than the users who felt unsure in the system. However the developers expressed that the system was inconsistent and would be hard to learn in contrast to the novice user who thought it would be easy to learn. Further, the open questions revealed that the update- and warning- buttons were now visible and urged to click on it. The participants also thought the timeline was now more logical than the previous one.

	DEVELOPERS			
	Task 1		Task 2	
	D1	D2	D1	D2
Easy to use	1	1	2	1
The functions work good together	1	1	2	1
Not consistent	5	4	5	5
Easy to learn	2	1	2	1
Difficult to use	4	5	4	5
Safe and secure	2	1	2	1

Figure 6. The developers' answers from the estimation questionnaire.

	USERS			
	Task 1		Task 2	
	U1	U2	U1	U2
Easy to use	2	2	2	-
The functions work good together	3	3	1	-
Not consistent	4	2	4	-
Easy to learn	2	4	2	-
Difficult to use	5	4	5	-
Safe and secure	3	4	3	-

Figure 7. The users' answers from the estimation questionnaire.

VI. Discussion

Method

The result of the heuristic evaluation could have been affected by that the evaluators is not design experts.

In the contextual inquiry there was an inability to install recording software on the test subject working computers led to an obstacle in order to understand their natural behaviour this could have had an impact on the result.

In a think aloud test, to properly conduct studies with few participants one should do tests in iterations, preferably once a week and if we had the time and resources we might have been able to find more design problems.

When analysing the interviews it was discovered that two of the recordings' quality was better on the caller and that it was hard to hear what the participant said, which could have affected the result if the analyzer missed something of importance.

The presented paper prototype have low depth, which means that users are more delimited to certain functions in the system, however, this could be improved by using highly interactive HiFi. During the pluralistic usability walkthrough, one of the representative users was forced to leave after the first task, which makes the results from the second task of the walkthrough less valid than that of the first task.

Result

Some of the design proposals were implemented in a paper prototype and this prototype was received positively by the participants of the pluralistic usability walkthrough. Through the contextual inquiry we found out that users use the system to gather hydrological data, and they do not use HYFO very often. The telephone interviews told us that other customers use the system often and extensively. Some of the interaction problems that we encountered were the design and position of the timeline, the position and design of "Vyer" which was an invisible menu and that users did not notice the system update button. In this discussion we bring up how these problems can be lessened or solved by taking into account several design theories.

Limitations

In the study our main limitation is the small number of participants. With more participants, there might be implications for information-heavy systems in general and stronger implications for HYFO specifically.

Further Research

In the study, we had a small number of participants. Given more money and time, more participants from different parts of the country could be included in the evaluation of the system, further evaluating which customers that need the system and which would benefit from another service. This could further illuminate ways that the interaction in HYFO could improve, and how HYFO could be developed to cater to the customers' needs. A study on a larger scale could be done using a SUS-survey, and sending this to all of the users of the system.

Implications

To SMHI our recommendations is a list of design improvements.

We also recommend business staff and developers in general to conduct broader testing and needs analyses before implementing new systems and services. This is because of the result from the contextual inquiry at Tekniska Verken, which showed that they have more functions than they need and could use a more simple kind of service at a lower cost.

By extension these implementations will make the end-users work more efficiently, reducing their workload and making their decisions easier. This will lead to greater profits for the companies buying the system.

REFERENCES

Arvola, M. (2014). *Interaktionsdesign och UX*. Lund: Studentlitteratur.

Beyer, H., Holtzblatt, K. (1998). *Contextual Design: Defining Customer-Centered Systems.* San Francisco: Morgan Kaufmann Publishers.

Snyder, Carolyn. (2003) *Paper Prototyping: The Fast and Easy Way to Define and Refine User Interfaces*. Morgan Kaufmann Publishers. Burlington, Massachusetts.

Goodwin, K. (2009). *Designing for the digital age*. Indianapolis, IN: Wiley Pub.

Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.

Nielsen, J. & Mack, R. (1994). *Usability inspection methods*. New York: Wiley.