Challenges of structuring data from usability and user studies

Dmitrij Lagutin and Mervi Ranta
PM&RG - Product Modeling and Realization Group
Department of Computer Science and Engineering
Helsinki University of Technology
P.O.B 5400, FIN-02015 HUT, Finland
fax +358-9-451 3293
tel. +358-9-451 4532, Dmitrij.Lagutin@hut.fi
tel. +358-9-451 4807, Mervi.Ranta@hut.fi

Abstract

This papers explorers challenges of structuring data gathered from usability and user studies. Large amounts of data gathered from user studies need to be thoroughly structured and analyzed before conclusions can be made.

This requires a lot of manual labor. In addition it's very hard to present manually processed data in a format, which can be easily understood by other researchers and experts.

Analyzing data gathered from user studies include: conducting interviews, combining all gathered data together, structuring this data by finding relevant and interesting things and creating different views to the structured data for various parties. It's very hard to do this manually, thus there is a need for the structuring tool.

This paper discusses the on-going research on a user data structuring tool from the viewpoint of the challenges that user data and expert users set. The objective of the paper is to provoke discussion on the requirements for such a tool and debate on the potential new working styles of usability and user study experts.

Introduction

Usability and user studies produce large amounts of different kinds of data. This data includes for example interviews, diaries, usage logs and essays. First, all this data has to be captured and processed from the raw sources into various kinds of data sheets and raw documents. Second, the data has to interpret into somehow structured presentations, diagrams and models that can be utilized and shared in the service development. Thus, usage research is very work intensive and though manual labor. Main problems are: how to transfer findings to the other usability experts and other experts involved in the development work? How to access and reuse already collected data?

The user data is collected as spoken, written and filmed material and is far from being formal or even structured. In addition, complex ontologies as well as vague expressions of colloquial language present many problems. A single word may have many meanings and there may be a several words with the same meaning. For example a place has many nicknames and many different persons may have similar names. Furthermore, the data contains complex chains and is inherently ambiguous. Therefore, it is not sensible to search for an automated solution for this problem. A more promising approach is to develop flexible tools for data structuring and modeling.

There are two particular challenges in creating such a data structuring and tracing tool. First, it is not possible to follow the conventional approach of analyzing some set of data in order to define the data structures and models that would form the basis for the software tool. The reason is that the ways to structure the data and the suitable models vary in each usability and user study round according to the user group, the usability and user study expert's habits, the focus of the development work etc. Thus, the user data-structuring tool cannot be based on any fixed models. Second, the structuring tool has the most challenging users that can be found – the usability and user study experts. It is inevitable that it will affect and change the working process and methods of the user data processing. The challenge lies in motivating the usability and user study experts to be creative and help in the process of finding feasible solutions.

Background

User studies in GO-project

Here are the main concepts of usability tests and user studies [1]

- Interview. User study experts conducts face-to-face interview with the subject.
- Diary. User study subject fills out a web-diary where he or she answers set of questions.
- User study consists of interviews, web-diaries and log-file analysis [2].

In this paper we don't look at the challenges of user studies and usability. Instead we concentrate to the user data structuring. More information about user studies is described by Jarmo Parkkinen [1].

Existing tools for structuring

There are some tools for qualitative research like ATLAS.TI [3] and NUD*IST [4]. However these tools are aimed at different problems. These tools are more aimed on text analysis, annotation, associating concepts with parts of text and producing affinity diagrams. ATLAS.TI and NUD*IST are more general purpose tools and the results created with them depend of the user. In addition ATLAS.TI and NUD*IST are more aimed for the single researcher working on some subject. NUD*IST includes support for "auto-coding" text which is useful in many situations.

Functionality of our structured tool is more focused for the specific task. However our tool is easily extendable for other kind of tasks. Structures created with our tool are very similar regardless of the person who uses it. Our tool isn't just aimed for one researcher, one of the main purposes of the tool is to help other people understand the original data. Automation of the structuring tasks can't be used in our work, since many words may have different meanings as mentioned previously.

Another difference is that results achieved with our structuring tool are separated from original data. This is important because of the privacy issues results should be usable without looking at original data.

Objectives of the structuring tool

First objective is an ability to reuse vast amounts of data gathered in the user studies. Most of

this data isn't in stored electronically and thus is very difficult to analyze, let alone to manage and reuse it in future research.

The main objective for a user data-structuring tool is to make data interpretation process transparent to others than the individual user study expert that collected and processed it. This means that it is possible to track the sources that lie behind the interpreted models. For example, the technology expert is able to check details from the interview data and point of a diary that were used as sources of a particular service scenario. Likewise, other user study experts are able to review the user data and generate alternative interpretations potentially by combining some new data that they have collected. The possibility to review sources and rationales makes the data comprehensive to others than the user study experts. Furthermore, the user data structuring tool also needs to provide a different views for the different parties, since e.g. user experts, prototype builders and profit analysts each have their own needs and different access rights that ensure the privacy aspects related to sensitive user data.

Tool consists of a separate application is responsible for editing the data and multiple views to that data. Different views of the data can be for showing data to other experts and researchers. All data can also be accessed through the web browsers. In addition it's possible to create graphs (figure 8), which describe various relationships between persons, communications, places and other objects.

The structuring tool is more designed to be an innovation prototype [5] instead of a complete system. Other uses of the tool include researching requirements for various data structures. Right data structures are usually found by trial and error and it's almost impossible to know them beforehand. With the structuring tool it's possible to test various approaches.

Viewpoints of the different users of the system

Structuring tool is designed to be flexible and to be used for different tasks.
Below are listed different types of persons that use the system:

1. User study expert

A user study expert carries out interview and inputs gathered data into structuring tool. Then user study expert analyzes and structures plain text data with the structuring tool and finds relevant connections like communication events. User study expert has full access to whole data and is responsible for deciding which information is private information that shouldn't be shown to others. User study expert isn't interested in the internal file format and other technical aspects of the tool.

After the structuring is completed user study expert will look for the new ideas for service developing. For example if user study experts sees that a subject wants frequently listen music while in train, he realizes that the subject could use wireless music service.

2. Other user study experts

Other user study experts have lesser access to the data. They are looking at the potential service ideas as the user study expert mentioned previously. In addition they can use already gathered data for their own user studies without having to start all over. User study experts can also find new ideas for improving the structuring tool.

3. Other researchers (software, hardware, network and business experts),

Other researchers are interested in rationales for their prototypes, they want to see the reasons behind them. They may have created some interesting device and want know wherever the users have a need for this kind device. In addition they are looking for ideas for future prototypes. As other user study experts, they don't have full access to the data.

4. Research and project partners

Research and project partners include researchers from other universities and companies working together on projects. They can see simplified version of the structured data without real names and other private information. This kind of data can be delivered to the public. They are interested in rationalizations for service ideas.

5. Developers of the structuring tool

Developers add new data structures and functionality to the structuring tool. They are interested in technical details and not in the data itself.

Example scenario: usage study on HUT students

In this scenario [1] target group consists of students of Helsinki University of Technology. Students are studying in degree programme of information networks. Usage studies consist of interviews and web diaries.

Interviews are conducted face-to-face and recorded. In addition students fill out a web form where they answer set of questions that consider the movements, events and communication during the day.

User study expert then processes the data and tries to find interesting parts from the interviews and communication events from the text.

Example:

Subject answers daily a set of questions through a web page, for example:

- *Who did you met and where did you meet them?*
- *With who did you have a contact and how?*

User data expert gets following answers:

- *11-12 I was eating at Innopoli with Kalle. I called Hessu and asked him for a coffee.*
- *12-12.30 Drinking coffee at Spektri with Hessu*
- *17-18 Drinking coffee at Leena's*

Figure 1. Questions and answers from the web form

After the questions are answered, user data expert structures the diary and finds items of interest like locations, persons and communication events.

Raw data:

11-12 I was eating at Innopoli with Kalle. I called Hessu and asked him for a coffee.
12-12.30 Drinking coffee at Spektri with Hessu

Items of interest:
Locations:
- *Innopoli*

Persons:
- *Kalle*
- *Heikki (Hessu means propably brother Heikki)*

A communication event:
- *A phone call to Heikki*
- *Meeting with Kalle*

Movement:
- *Walked from Otaniemi to Innopoli*
- *Walked from Innopoli to Spektri*
- *From Spektri back to Otaniemi on Heikki's car*

Figure 2. Raw data and found items of interest

User study expert knows that Innopoli is a building near Otaniemi and thus marks it as location. Kalle and Hessu are names, and thus they are marked as persons. Hessu is a typical nickname for Heikki. There are two communication events present, a lunch meeting with Kaler and a phone call to Heikki concerning the coffee. User study expert concludes that subject moved from Otaniemi to Innopoli, then from Innopoli to Spektri and then back to Otaniemi.

Overview of the structuring process

In the figure below is shown structuring process with a data structuring tool. Structuring process consists of gathering data, inserting data into structuring tool, structuring it and analyzing the results.

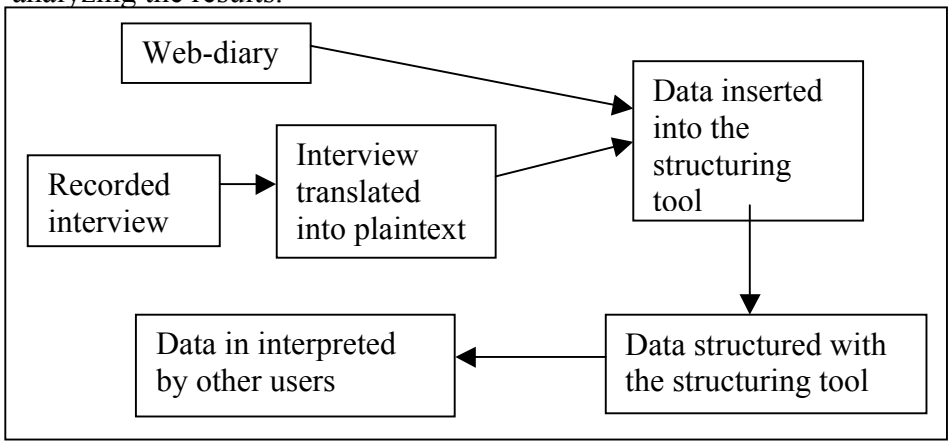

Figure 3. Structuring process with a data structuring tool

Steps in the structuring process:
1. Subject is interviewed or he or she fills out a web form
2. If necessary, interview is translated into plaintext by user study expert
3. User study expert inserts a plaintext data into the structuring tool
4. User study expert uses structuring tool to structure data
5. Other experts interpret structured data

Implementation

Architecture overview

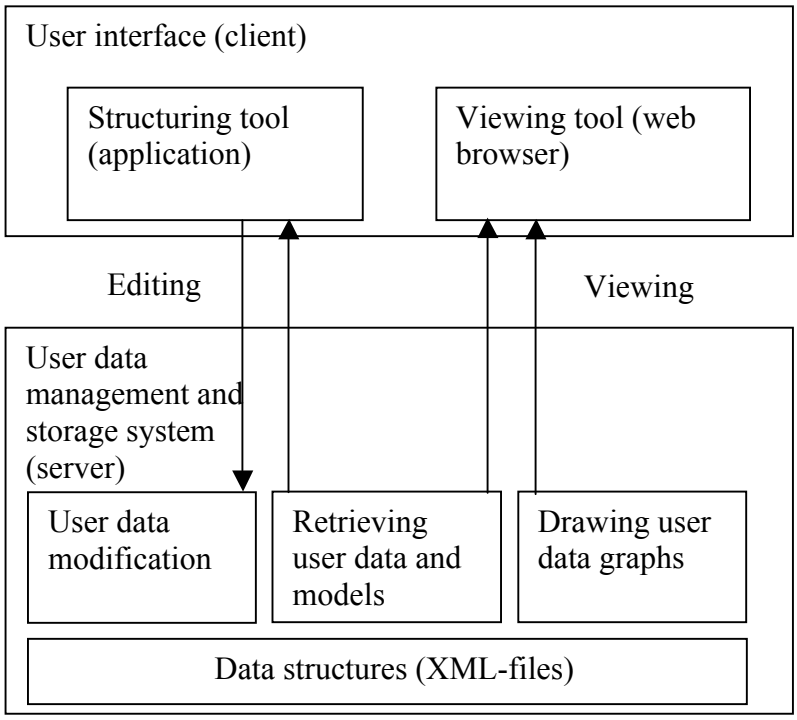

Figure 4. Architecture overview of the system

Figure above show the architecture of the system. Arrows are showing dataflow in the system. Arrows on the left side show dataflow during the editing data and arrows on the right side show dataflow when viewing data.

Division of responsibilities between user interface and data management and storage system

In order to make user interface easier to implement, most of data management functionality resides on server side (see figure 4). For example in order for user interface to display communication event and related persons it needs to get information about communication event and related persons. To achieve this user interface requests from the data management and storage system communication event object with related links.

Another solution would be to send whole available data with links to user interface and let it figure relations. However this approach would make user interface more complicated. Since data should be accessed through the browser, it makes more sense to create needed functionality on the server side, instead of re-implementing it several times.

In short, server send only necessary data to the user interface, user interface shows it in

dialogs to make structuring easier and then user interface sends back updated data to the server.

Data and relations (links) are stored on the server side together with three major modules responsible for managing data as showed in figure 4. These modules are:

- User data modification module is responsible for adding and editing data and links.
- Retrieving module is responsible for sending data to the user interface and web browser.
- Graph drawing module is responsible for producing various graphs that can be viewed through the web browser.

Client side (figure 4) contains user interface for the user experts, which is used for structuring process. This client presents data to the user experts and by communicating with Update-module adds or updates data. In addition web browser can be used for accessing information and viewing graphs.

User interface for the structuring tool and server-side applications are written in Java to ensure compatibility with different software and hardware platforms. Communication is handled using HTTP protocol.

Data structure

Data format was designed to be very simple and flexible. It can be easily extended. Thus data and relations between data (links) are separated from each other.

Stored data consists of following:

- Categories contain similar objects like places, persons, communication events and diaries.
- Objects contain information about something specific, like a person called Matti or Leena's web-diary. Each object belongs to one category. Internal format of objects varies in different categories.
- Links define relationships between objects. Link has a starting and ending point, which refers to the objects and categories. Links can be also used to create hierarchies. For example: Area -> City -> Suburb.

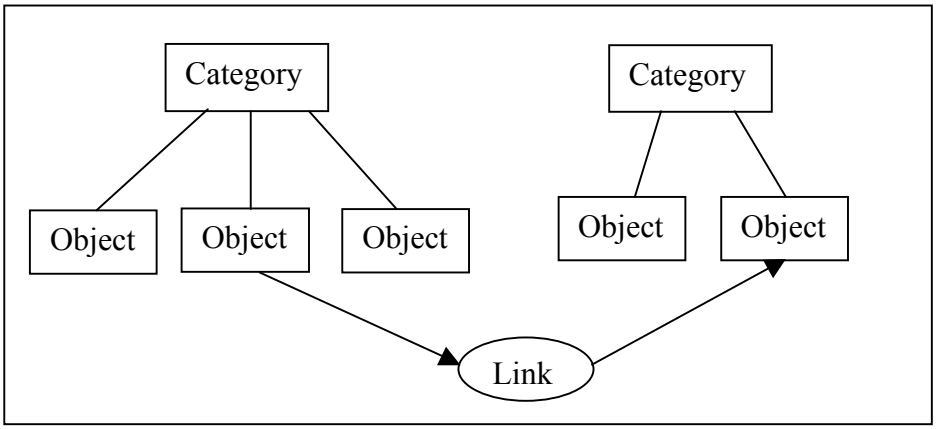

Figure 5. Data structure

An example about stored data:
A communication event object contains following data:

- Name of the event.
- Group, way, category, place, time and matter of communication.

In addition following links are related to the communication event:
- Links to the persons present in the communication event.
- Link to the origin of the communication event (web-diary, interview or some other source).
- Possible links to the related communication events.

This kind of structure make easy to see various connections between people.

It is also possible to create various different views from the data for different parties. In addition of showing data in plaintext, it's possible to create automatically various graphs describing relationship between objects such as people, places and communication events (figure 8).

Internal format of various objects is fixed for specific use. This ensures that structuring process is easy and end results won't depend much on user-study experts. However internal format of objects can be easily changed for different user studies.

Functionality

Structuring user interface

In order to speed up the structuring process, client of the structuring tool contains various dialogs for filling various information like communication events.

Following functionality is implemented:

Creation of communication events:
User opens a web-diary using the structuring tool. User selects a line from the diary where some communication is mentioned, opens a communication event dialog and fills the necessary information about communication event.
Structuring tool creates a new object into communication event category and creates necessary links to the persons present in communication event and to diary, which is the source of communication event.

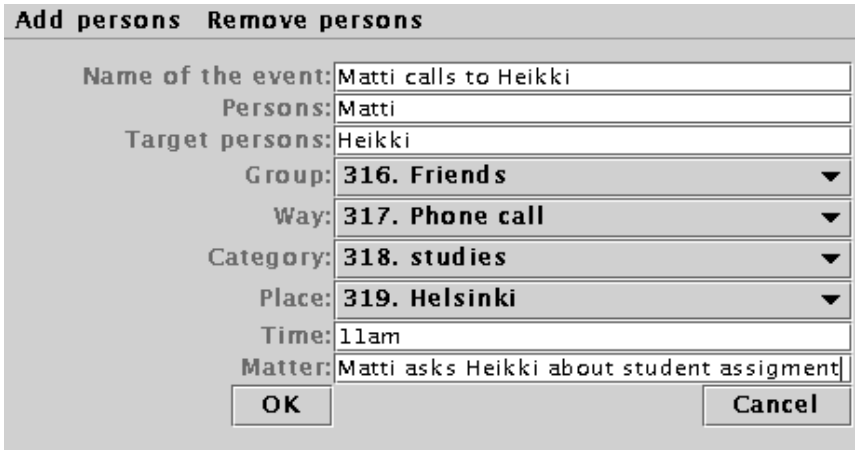

Figure 6. Dialog for creating and editing a communication event

Editing persons information:
User opens a web-diary and selects the line where some information is mentioned about the person. User selects person from list and open dialog to edit person's information.

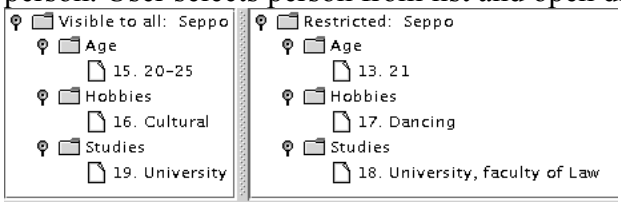

Figure 7. Dialog for editing personal information

In order to protect privacy, there is two separate instance of personal information. On the left side is information that is visible to all and shouldn't contain anything specific that can be used to identify user. On the right side there is information, which is visible only to the user expert. When looking at the person's information from browser, it's possible to choose to view full or only partial information.

In addition it's possible to create links (relations) between various objects (e.g. Otaniemi and Espoo, since Otainemi is a suburb of the Espoo).

Creation of Scenarios:
Scenarios are similar to communication events, but they contain several issues and scenes. For example, scenario may be about a person looking for the new job. In first scene person is interviewed by the phone, in the second scene person is meeting employer personally.

Graph generation

In addition to view data in text format, it's possible to create various graphs describing relationships between objects. These graphs can be viewed through the browser. Currently supported graphs are:

- Graph showing communication events and other persons related to the selected person
- Graph showing communication groups (work, friends, hobbies) related to the selected person
- Graph showing persons related to the selected communication group
- Graph showing locations related to the selected location
- Graph showing communication event chain (example showed below)

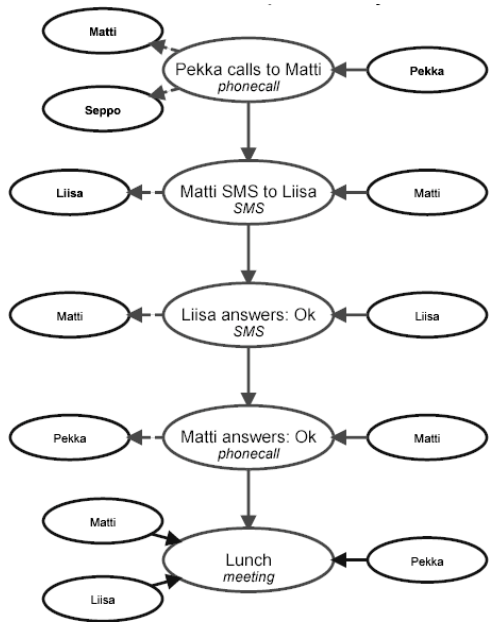

Figure 8. Communication event chain

This picture describes a chain of events in which some friends (Matti, Seppo and Liisa) organize a meeting to have a lunch together.

Middle row contains communication events. Under the name of the communication event, way of communication (phone call, meeting, SMS, e-mail or something other) is present. Left and right rows represent persons involved, persons name is shown in bold when the persons is mentioned for the first time during communication event chain. Arrows between and dashed lines represent the order of communication. Dashed line between person and communication event means that person is target of communication. Normal line means that person has initiated communication. In the beginning Pekka calls both Matti and Seppo to organize a meeting.

All this information is gathered by looking at structured communication events and links between communication events and persons. This allows generation of the above-mentioned graph automatically without user intervention. Graphs are generated in a SVG format and can be converted to the JPEG format for older browsers.

Alternative views to user data

The same data that can be viewed through the structuring user interface can be also viewed through the web browser using XSLT transformations. This is only for the viewing, editing data isn't possible.

This makes easier for other researcher and partners access data, since they don't need to learn a user interface.

Conclusion and future work

The work is still in progress. We have found the structuring tool to be promising and flexible. Ability to change used data structures easily is very important, since in many case first choice

was the wrong one.

User experts are quite demanding users. They are quite unwilling to learn to use new user interfaces, thus structuring tool needs to be easily extendable. Principle of the structuring should be the same in different kind of user studies.

In the future, user interface needs to be improved to make structuring process very intuitive and easy to learn.

Web-diary was chosen example case, because it provides some initial structuring. Future work includes support for more complex data and structures.

References

1. Parkkinen, J. 2003. Mobiiliviestimen käytettävyyden arviointimenetelmiä.
2. Parkkinen, J. and Ranta, M. 2003. A zero prototype as a usability tool. In proceedings of COST269 Conference, Helsinki, Finland, 3-5 September 2003. COST.
3. ATLAS.ti, a powerful workbench for the qualitative analysis of large bodies of textual, graphical, audio and video data. http://www.atlasti.de
4. NUD*IST, a software for powerful management and flexible analysis of text data http://www.qsr.com.au
5. Ranta, M. and Soisalon-Soininen, E. 2001 GO-PROD Product concepts and user aspects.