Contextual Method for the Re-design of Existing Software Products

Rachel Jones
Natasa Milic-Frayling
Kerry Rodden
Alan Blackwell

30 September, 2004

Technical Report
MSR-TR-2004-96

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

Contextual Method for the Re-design of Existing Software Products

Rachel Jones
Instrata Ltd.
62 Kingston St., Cambridge, UK
Rachel.Jones@instrata.co.uk

Natasa Milic-Frayling
Microsoft Research Ltd.
7 J J Thomson Avenue, Cambridge, UK
natasamf@microsoft.com

Kerry Rodden[1]
Google Inc.
2400 Bayshore Pkwy, Mountain View, CA, USA
krodden@google.com

Alan Blackwell
Computer Laboratory, University of Cambridge
15 J J Thomson Avenue, Cambridge, UK
Alan.Blackwell@cl.cam.ac.uk

ABSTRACT

This paper is concerned with the problem of improving software products and investigates how to base that process on solid empirical foundations. Our key contribution is a user-centered, contextual method which provides a means of identifying new features, to support the discovered and currently unsupported ways of working, and a means of evaluating the usefulness of proposed features. Standard methods of discovery and evaluation, such as interviews and usability testing, gather some of the necessary data but each individually falls short of covering all important aspects. We overcome the shortcomings of these individual approaches by applying an integrated method for collecting and interpreting data about product usage in context. We demonstrate its effectiveness when applied to the discovery and evaluation of new features for standard web clients.

Author Keywords

Design Methods, User Studies, User-Centered Design, World Wide Web.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

It is common that new versions of a software product are released during a product's lifetime. New releases contain bug fixes and frequently extend the functionality of a product by adding new features. New features typically originate from a variety of sources, such as a technical opportunity, a competitor offering, a user suggestion, or simply a good idea. Before features are incorporated into a product, commonly they are tested for their appeal and usability with end users, and then rated based on the users' feedback or by users themselves. During this process, it seems there is no assessment, and therefore no evidence, of the usefulness of a particular feature. In summary, it appears a somewhat ad hoc process for introducing new features into a product.

The development and application of user-centered design methods has focused predominantly on new products; but there has been comparatively little development and application of user-centered methods to understanding how to improve existing products. Effort has mainly focused on evaluation, primarily usability testing and expert evaluation, with some evaluative ethnography, intended to verify or validate a set of already formulated design decisions.

The question arises whether any existing user-centered design (UCD) methods could be usefully employed in either the discovery of new opportunities or the evaluation of proposed features in order to improve the re-design of an existing product. The initial exercise is to identify the aspects of an existing product that should be explored.

We identified four questions that a method needs to answer:

- Whether people are able to make sense of the underlying concepts in a product?
- What features do they use?
- How do existing features support what people use the product for and whether the product could be extended to add value?
- What new features or products could be developed to support discovered and currently unsupported ways of working?

[1] Work done while at Instrata Ltd, working on a project at Microsoft Research, Cambridge, UK.

Although these are broad questions, they help the development team address the usefulness of existing features and explore potential areas where new features would improve the product offering.

In this paper, we examine whether existing user-centered design methods enable us to explore the above aspects. We show how we need to develop a new approach that combines several existing methods and employs them in a novel way. We apply and discuss the effectiveness of the new method in two cases, the discovery of new features and the evaluation of newly designed and developed features.

USER-CENTERED DESIGN METHODS

In this section we give a brief critique of the ability of commonly used user-centered design (UCD) methods to address the broad questions identified in the previous section. A more detailed description of each method can be found in most HCI textbooks, such as [1].

Whether people are able to make sense of the underlying concepts in a product?

To answer this question we could use interviews to probe on peoples' understanding. **Semi-structured interviews** in context gather an understanding of user motivations and activities. They can be unreliable because the information obtained is only what the person tells us, which in turn depends on the user's own awareness of activities and can be oriented to what the user thinks the interviewer wants to hear. Furthermore, the information is described after the event and is post-rationalized.

Alternatively, we could analyze peoples' behavior using logging or field observation, as these are more reliable methods for identifying whether users demonstrate an understanding. **Data logging** involves recording usage of and interactions with a product. Large amounts of data are gathered which are often analysed statistically, although several visualization techniques have been developed to enable the data to be viewed for individuals and aggregately. Logging lacks information about the wider context and thus gives no indication of the person's activities and whether these activities are successful. Logging tends to be used in the evaluation phase of a product, for example, to see which Web pages are most popular or rarely accessed by the visitors to a web site.

Field observation produces detailed descriptions of the "workaday" activities of people within their specific contexts. The gathering of data tends to be prolonged and the interaction detail is not commonly captured. The methods have been used in the discovery and evaluation phases of a project [2]. Rose et al. present a set of practical guidelines based on ethnographic research to be used by designers preparing to evaluate a system for re-design [3].

It seems that neither interviews nor logging nor field observation used individually will provide the necessary data to answer the question.

What features do people use?

This question requires an understanding of the detailed user interaction with the product. For that a diary study or field observation would prove too intensive for the participant or observer respectively. Data logging would capture exactly what features are used.

How do existing features support what people use the product for and whether the product could be extended to add value?

We could use interviews, field observation or a diary study to understand how the product fits with what people do and to identify areas where value could be added.

Diary studies are commonly used to identify people's activities and can vary in the level of detail that is captured. The information is expected to be recorded at the time of the activity, i.e., recorded in the present, although people sometimes forget. The very activity of recording can intrude on current activities, particularly if significant detail is required.

What new features or products could be developed to support discovered and currently unsupported ways of working?

We would need to use field observation to understand peoples' ways of working in the broader context. At the same time we would need logging to capture and understand users' interaction patterns.

A standard method we have not mentioned so far is usability testing. **Usability testing** is used in the evaluation phase of the development lifecycle to verify or validate a set of already formulated design decisions. It is used in the discovery phase to identify usability issues that need to be addressed in the next release. It does not address the usefulness of features.

From this brief discussion, it is clear that none of the standard methods, when used on their own, would be able to cover all four issues that we need to explore. Thus, it is necessary to design a novel approach. In the following sections we first describe the method we developed and then show how it has been used in two studies.

CONTEXTUAL RE-DESIGN METHOD

In this section we describe the method we developed to explore the various issues in the re-design of features for an existing software product, be it an application, a service or a system.

We selected three standard user-centered design methods to use concurrently to gather the requisite data:

- Semi-structured interviews in context, primarily to ascertain how the product fits with people's activities.
- Data logging to identify detailed interaction with the product.
- Field observation to understand the broader context of people's workaday activities.

We used these methods together to collect and analyze the data. Each method contributed to this process a set of questions or areas of understanding that could be explored by one of the other two methods. It is this hybrid approach of simultaneously and interactively using the three methods that makes it more effective than using individual methods on their own or sequentially. We describe the data gathering in detail before we discuss the analysis.

Data gathering

The method we have developed is for the re-design of an existing system. This has the advantage that we have the possibility to log interaction events; but also the disadvantage that we need to build a logger for the product. The logger needs to capture user interaction events. Traditionally, these will be user interface events, such as button presses, but we could envisage capturing events from sensors in a ubiquitous environment.

Attributes of the logger that we found particularly critical were a description of the interaction sequences, preferably visualized, an indication of the content within the interaction sequence, and statistical utilities that enable discovery and analysis of patterns in collected data. Hilbert and Redmiles [4] survey a number of computer aided techniques for extracting user-related information from UI events.

Some events could not easily be logged, and so we asked participants to record these on a crib sheet, such as when they printed a document.

Various confidentiality issues had to be resolved with regards to data logging. We only presented aggregated results to the organizations in which participants worked, protecting individual results. We did not capture secure parts of the product or parts individuals or the organizations had told us they did not want captured.

Interviews were semi-structured and carried out in context. We began a study by carrying out an initial interview with each participant to introduce ourselves, to explain what we were doing and to find out general information about the participant's job, their role, and their experience. Subsequently, every two or three days we retrieved the data logs and analysed each participant's recording. We focused on specific aspects, such as use of features, routines, user events and problems. The log analysis created prompts for the next interview. Every interview was recorded for analysis and shared with the rest of the team.

The combination of using logs and interviews provided two enormous benefits. Firstly, it resolved a criticism of interviews that participants only mentioned what they were aware of and what they thought the interviewer wanted to know. Secondly, logged information is typically hard to interpret reliably, but the interviews provided the broader understanding that allowed us to make sense of the data.

During the course of the study, we spent half a day with participants observing their "workaday" activities. This was video recorded.

Analysis

The data we gathered was used in three ways in the discovery phase. The initial analysis involved exploring and discovering common themes in the data. The second type of analysis involved testing design hypotheses and the third type of analysis justified the effectiveness of proposed solutions. We discuss each of these in more detail.

Common themes

In the initial analysis, we used the data to explore the following:

- We could find out to some extent how much people understood the underlying concepts embedded in the product.
- We could find out how often and in what contexts particular features were used. We could begin to understand why different participants used some features whilst others didn't.
- We could understand the activities that participants were using the product for. We could understand what participants attempted to do but failed to achieve.
- We could understand new ways of interacting with the system, whether it is a workaround or a completely different pattern that is not supported but emerges from the data. We could understand the parts of the system they accessed and their patterns of access.

Testing design hypothesis

We began to make various hypotheses about the interaction patterns and used the log data to verify them:

- We could test if a pattern did actually occur and how often.
- We could tell how prevalent the pattern was amongst participants and in what context it occurred.
- We could find out more detailed parameters about a pattern.

Later in the development cycle, we used the log data in the similar way to justify the usefulness of developed features. For example, we could ask how much a new feature improves the effectiveness of a product.

Justification of novel features

As we engage in the design of novel features, we used the collected data to verify that the features are likely to improve user's experience if introduced in the product. The analysis of collected data allowed us to estimate the potential effectiveness of the features.

Evaluation of new features

The same approach can be applied to the evaluation of newly designed and developed features. While one can focus on the *common themes* discussed above, looking at underlying concepts, features, activities, and interaction patterns, we can further center our attention on aspects of the new features:

- What is the achieved effectiveness and perceived usefulness of the new feature?
- Did participants adopt the new feature, what are the barriers to its adoption and how do participants' change their behavior over time?
- What problems are incurred due to the design of a new feature?
- What are participants' expectations of a new feature and how does this affect their satisfaction with the feature?

Having described the new method and how the data is gathered and analyzed, the next section describes two studies where we applied the method.

APPLICATION OF CONTEXT RE-DESIGN METHOD IN FEATURE DESIGN

In this section, we present two studies where we applied the new method. The first study uses the method to discover new features and the second study uses the method to evaluate new features.

Study 1: The discovery of features

Aim

The aim of the first study was to investigate the way in which people navigate and search the web in order to inform the next version of a web client. We wanted to answer very varied set of questions, some specific and some more general. They fell into four common themes, related to concepts, features, activities, and interaction with the product, as discussed in the Analysis section:

- In terms of concepts, do people understand the stack model underlying the back button? Do people understand a URL and the way it is structured? If they don't understand, does it reduce their effectiveness in navigating the web?
- What features of the web client do people use and in what situations? Once users have found what they were looking for, do they record it and if so how? How do people return to a site or a page they have been to before?
- How do people make use of information available on the Internet and Intranet as part of their everyday activities?
- What strategies do people adopt to look for information through the client?

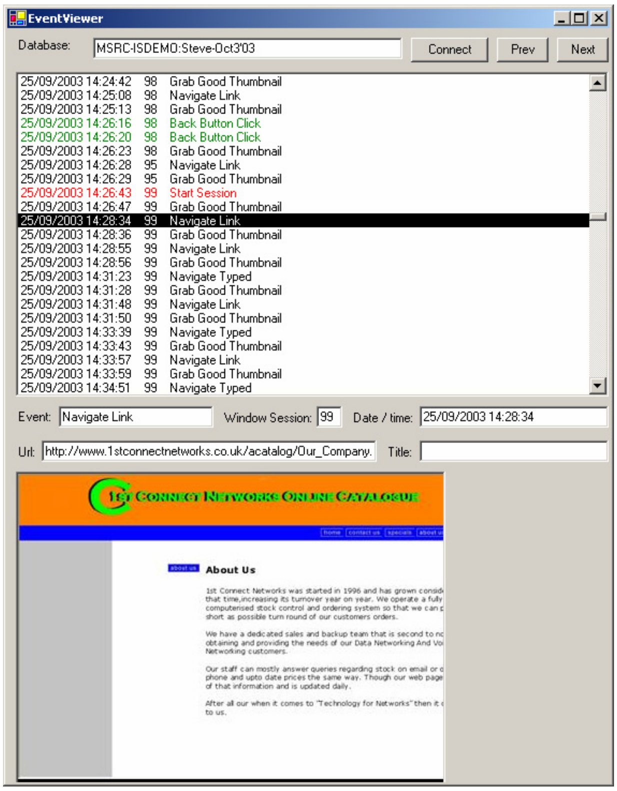
Application of the method

We carried out a study with 9 people in different departments of local government offices, Cambridgeshire County Council. Some participants were knowledge workers and others were administrators. We logged participants' use of Internet Explorer for between 2 and 3 weeks. We retrieved the logs every couple of days, analyzed them and carried out semi-structured interviews with the participants in their workplace.

Data logger

We used data logging software to record user's interaction with the product. In the section on related work we discuss comparable types of logger that have been developed. We developed our own logger mainly for convenience, and not because our logger is particularly different from others.

The logger is programmed to capture particular events in the product, such as button presses. The logged data is stored in a SQL database, that can be queried and deliver statistical results. The log viewer shown in Figure 1 allows us to quickly scan through the events. Events of particular interest are highlighted for easy identification. Metadata about each event is recorded and displayed in different fields. Thumbnail images of the content that the user viewed are shown. In this manner the LogViewer can be used as an event-based replay of the user interaction with the product. We can experience the linear exposition of pages that the user viewed along with an indication of the type of events they invoked.



Database: MSRC-ISDEMO:Steve-Oct3'03

```
25/09/2003 14:24:42   98   Grab Good Thumbnail
25/09/2003 14:25:08   98   Navigate Link
25/09/2003 14:25:13   98   Grab Good Thumbnail
25/09/2003 14:26:16   98   Back Button Click
25/09/2003 14:26:20   98   Back Button Click
25/09/2003 14:26:23   98   Grab Good Thumbnail
25/09/2003 14:26:28   95   Navigate Link
25/09/2003 14:26:29   95   Grab Good Thumbnail
25/09/2003 14:26:43   99   Start Session
25/09/2003 14:26:47   99   Grab Good Thumbnail
25/09/2003 14:28:34   99   Navigate Link
25/09/2003 14:28:36   99   Grab Good Thumbnail
25/09/2003 14:28:55   99   Navigate Link
25/09/2003 14:28:56   99   Grab Good Thumbnail
25/09/2003 14:31:23   99   Navigate Typed
25/09/2003 14:31:28   99   Grab Good Thumbnail
25/09/2003 14:31:48   99   Navigate Link
25/09/2003 14:31:50   99   Grab Good Thumbnail
25/09/2003 14:33:39   99   Navigate Typed
25/09/2003 14:33:43   99   Grab Good Thumbnail
25/09/2003 14:33:57   99   Navigate Link
25/09/2003 14:33:59   99   Grab Good Thumbnail
25/09/2003 14:34:51   99   Navigate Typed
```

Event: Navigate Link Window Session: 99 Date / time: 25/09/2003 14:28:34

Url: http://www.1stconnectnetworks.co.uk/acatalog/Our_Company. Title:

About Us

1st Connect Networks was started in 1996 and has grown consid that time,increasing its turnover year on year. We operate a fully computerised stock control and ordering system so that we can p short as possible turn round of our customers orders.

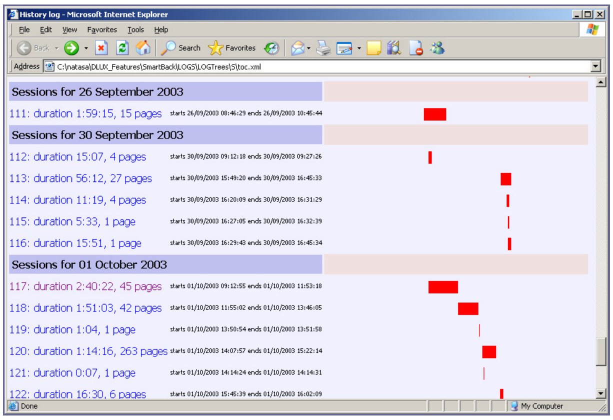
We have a dedicated sales and backup team that is second to nc obtaining and providing the needs of our Data Networking And Vo Networking customers.

Our staff can mostly answer queries regarding stock on email or c phone and upto date prices the same way. Though our web page of that information and is updated daily.

After all our when it comes to "Technology for Networks" then it c to us.

Figure 1. Log viewer

The detailed view of user's navigation steps in Figure 1 is complemented by a higher level view of the navigation history. The viewer in Figure 2 shows the user's navigation for each observed date. It provides easy access to individual window sessions and details about the duration and length of the navigation paths. This is expressed in numeric values and through visual displays (red bars indicate the duration of each product session). Following the hyperlinks of the individual window sessions activates the detailed graphical view of individual window sessions as shown in Figure 3.


History log - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back
Search
Favorites
Address
Done
My Computer

C:\natasa\DLUX_Features\SmartBack\LOGS\LOGTrees\S\toc.xml

Sessions for 26 September 2003

111: duration 1:59:15, 15 pages starts 26/09/2003 08:46:29 ends 26/09/2003 10:45:44

Sessions for 30 September 2003

112: duration 15:07, 4 pages starts 30/09/2003 09:12:18 ends 30/09/2003 09:27:26

113: duration 56:12, 27 pages starts 30/09/2003 15:49:20 ends 30/09/2003 16:45:33

114: duration 11:19, 4 pages starts 30/09/2003 16:20:09 ends 30/09/2003 16:31:29

115: duration 5:33, 1 page starts 30/09/2003 16:27:05 ends 30/09/2003 16:32:39

116: duration 15:51, 1 page starts 30/09/2003 16:29:43 ends 30/09/2003 16:45:34

Sessions for 01 October 2003

117: duration 2:40:22, 45 pages starts 01/10/2003 09:12:55 ends 01/10/2003 11:53:18

118: duration 1:51:03, 42 pages starts 01/10/2003 11:55:02 ends 01/10/2003 13:46:05

119: duration 1:04, 1 page starts 01/10/2003 13:50:54 ends 01/10/2003 13:51:58

120: duration 1:14:16, 263 pages starts 01/10/2003 14:07:57 ends 01/10/2003 15:22:14

121: duration 0:07, 1 page starts 01/10/2003 14:14:24 ends 01/10/2003 14:14:31

122: duration 16:30, 6 pages starts 01/10/2003 15:45:39 ends 01/10/2003 16:02:09

Figure 2. Session viewer

The interaction trail viewer (Figure 3), breaks up the user's interaction flow into navigation trails, represented as separate tree structures, and shows thumbnails of screenshots. This interaction trail viewer allowed us to scan quickly through a navigation session and observe the structures that resulted from the user's interaction with the product. Further details about individual navigation steps and target pages can be viewed in the web page viewer, shown in Figure 4. This view is activated by clicking on the individual thumbnail images.

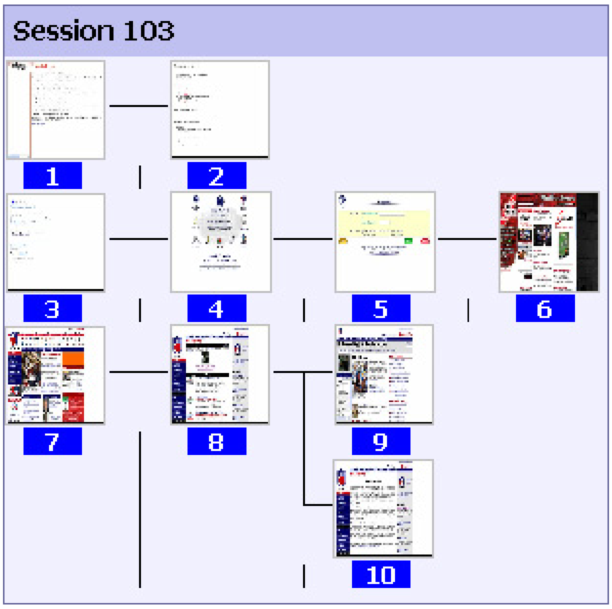
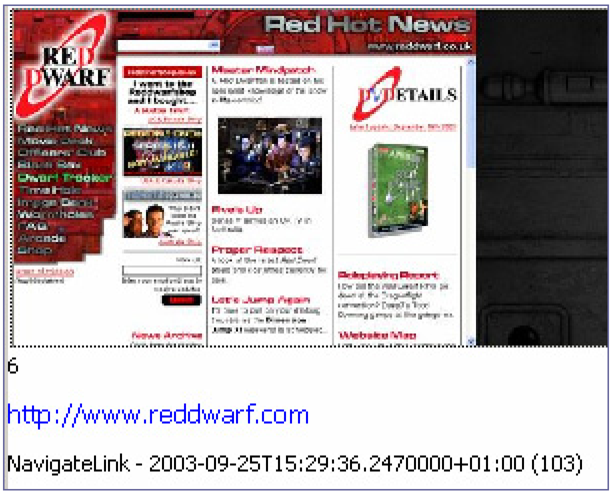


Figure 3. Interaction trail viewer

The logger could be used to provide detailed transcripts of the user's navigation. Transcripts were used to verify the windows involved in a pattern that had been identified through statistical analysis. They were particularly important in preparing for interviews because the context could be used to prompt the user about particular events. They also enabled us to better understand the comments that the user was making about the circumstances under which the system or a feature was used.



6

http://www.reddwarf.com

NavigateLink - 2003-09-25T15:29:36.2470000+01:00 (103)

Figure 4. Web page viewer, showing details of the page 6 in the navigation trail shown in Figure 3.

Findings and initial analysis

For each of the four questions on our aim list, we here discuss examples from the analysis.

Concepts. In terms of concepts, participants expected to return to pages by traversing navigation paths using the back button and were confused when some pages were not accessible. It was apparent that few of the participants understood the stack model that underlies the back button.

We observed in the logs that participants were able to edit URLs appropriately in order to navigate, for example, from an individual page to the site home page. Participants had gained some understanding of the structure of URLs.

Features. We found out how often particular features were used. For example, the back button was used 22.7% of navigation actions whereas the forward button was used 0.2%. Delving into the log data we noticed that the forward button was effectively used as an undo for the back button.

Activities and information access. We used the interview data to identify user activities. Compared to other studies of web activities [5], the monitoring activity seemed more prevalent. We identified several types of monitoring; participants would check the daily reports, such as the online newsletter, participants would visit consumer sites to find out the latest releases, participants would monitor filed documents, such as planning applications, to assess how heavy the workload would become, and participants would monitor web sites for the arrival of documents, such as the budget statement from central Government, which appears on line when the chancellor stands up to address parliament.

We noticed patterns of activities through the day and week. Participants would check on bank accounts and do pre-shopping just before lunchtime. They would check on holiday locations before the weekend. We noticed people persistently trying to do something but failing to achieve it, such as trying to find out the address of the local store in a national chain. From the logs we were able to understand how common these activities were.

Strategies for information access. In our observations, we noticed that people had regular sites which they commonly visited by typing the name and using the auto-complete feature. Interestingly, participants rarely used on-line search for a topic; only 6.6% of all navigation activities were related to search.

We recognized the essential difference between transitional and intentional re-visitation of pages. Transitional pages result from hub and spoke navigation, where the user is required to traverse to the hub in order to navigate another spoke. Other studies of web navigation do not distinguish between transitional and intentional back navigation [6].

Hypotheses

We illustrate some of the hypotheses related to the usage patterns that we drew from the findings and which we tested on the data. As discussed in the Analysis section, the data enabled us to hypothesize on the occurrence and prevalence of the patterns as well as on the parameter values that characterize the pattern, such as periodicity within various time intervals.

Use of Favorites. From the logs, we knew that only 2.9% of navigation actions were accessing Favorites and only three out of nine participants made regular use of Favorites to return to pages. In interviews, participants expressed guilt at not using them more. On further probing, we found most participants used favorites to store URLs that were difficult to find or hard to remember, whilst a few used them as a quick access list open in bar at the side of the client. On further investigation of the logs, we found that frequently accessed pages were often not recorded as Favorites. Participants said they rarely organized their favorites and the logs showed that a quarter of stored favorites were out of date. We identified a need for automating frequently accessed pages and making them available at the time they are needed, which resulted in us developing a feature we called SmartFavorites, comprising TrueFavorites and Prediction links [17].

The nature of a semi-structure interview provided an opportunity for the users to express their opinion freely, followed by their own train of thoughts and associations. This revealed very valuable information about general design preferences. For example, one of the participants spontaneously expressed concern with an existing feature, articulating how unhappy she was with intrusive "push" technologies, such as a new email alert. This led us to consider "light push" in SmartFavorites in form of very subtle suggestions in the toolbar rather than intrusive interruption through pop up windows or similar.

Hub and spoke. We noticed the extensive use of hub and spoke navigation; 8.4% of pages become hubs and 28% of page visits are to hubs. We looked at key pages within navigation trails to identify the characteristics of hubs. This led to the design of SmartBack, a feature that allows users to jump directly back to a hub [7][17].

Retracing navigation sequences. We observed that people navigate directly to a site and then follow a path to the page they wanted. We wondered how long these trails were, how common they appeared and whether they were shared. Table 1 shows statistics on the repeated navigation sequences for the nine participants as identified from the logs.

Quite a number of two step navigation sequences are repeated by participants. As these are probably the result of executing a link on the page, the opportunity for improving "forward" navigation is more in the realm of sequences consisting of 3 or more steps. As can be seen, almost all the participants engaged in at least couple of repeated sequential navigations. A closer look at the statistics reveals that there are about 30 distinct 3-step sequential patterns that were observed among the 9 participants.

	Number of observed repeated sequences per user								
Length	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
2	8	9	15	19	23	32	6	74	12
3	2	6	7	8	4	6	4	24	7
4	0	1	5	1	4	3	2	3	4
5	0	0	4	1	1	0	2	3	2
6	1	0	0	1	0	0	0	1	0

Table 1. Statistics on the repeated navigation sequences traversed by the participants in the first study

These observations led to one of the algorithms comprising SmartFavorites features which captures and learns from the repeated navigation sequences. As the user access a page from one of the recurrently accessed 3-step sequences, the link to the end-of-sequence page would subtly be displayed on the Link Bar.

Justification

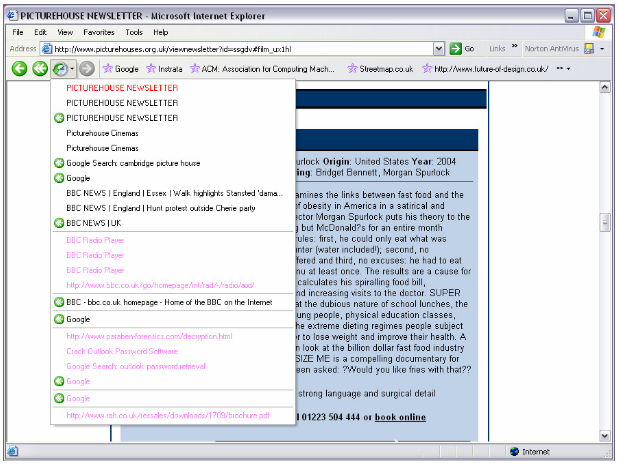
We arrived at several potential features but we needed to know how much it would improve the effectiveness of the product if they were introduced and used by the users. For example, if we introduced the SmartBack feature, would it significantly reduce navigation actions? We found out that 64% of all back clicks are to hubs, and so the remaining 34% of backward navigation steps have been used for navigation through branches. Thus a SmartBack feature that allows the user to access the hubs directly would cut down on the transient re-visitation of pages through the regular Back button.

In this section, we have shown how we used the method to discover new features through a series of analyses.

Study 2: The evaluation of features

Aim

The aim of this study was to assess the usefulness and effectiveness of four proposed features for a new version of the web client: SmartBack, Session Overview, and SmartFavorites that proactively suggested links that the user has seen frequently and relatively recently. The Overview feature is a drop-down list containing a specified number of links to pages the user has visited (default 30 links) [17]. Links are presented in the order of visit and icons added by the side of the list to indicate whether a link is a hub, a Favorite or a typed URL, in order to flag potentially key pages to the user. Figure 5 shows a screen dump of the new features. The SmartBack feature has a double arrow pointing leftwards. The Overview is the drop-down list shown. Four to six TrueFavorites and Predictions appear as single and double purple stars, respectively, on the toolbar, with a drop-down list on the right, showing additional links that do not fit on the toolbar [17].


PICTUREHOUSE NEWSLETTER - Microsoft Internet Explorer
File
Edit
View
Favorites
Tools
Help
Address
http://www.picturehouses.org.uk/viewnewsletter?id=ssgdv#film_ux1hl
Go
Links
Norton AntiVirus
Google
Instrata
ACM: Association for Computing Mach...
Streetmap.co.uk
http://www.future-of-design.co.uk/
PICTUREHOUSE NEWSLETTER
PICTUREHOUSE NEWSLETTER
PICTUREHOUSE NEWSLETTER
Picturehouse Cinemas
Picturehouse Cinemas
Google Search: cambridge picture house
Google
BBC NEWS | England | Essex | Walk highlights Stansted 'dama...
BBC NEWS | England | Hunt protest outside Cherie party
BBC NEWS | UK
BBC Radio Player
BBC Radio Player
BBC Radio Player
http://www.bbc.co.uk/go/homepage/int/rad/-/radio/aod/
BBC - bbc.co.uk homepage - Home of the BBC on the Internet
Google
http://www.paraben-forensics.com/decryption.html
Crack Outlook Password Software
Google Search: outlook password retrieval
Google
Google
http://www.rah.co.uk/ressales/downloads/1709/brochure.pdf
Internet

urlock **Origin**: United States **Year**: 2004
ing: Bridget Bennett, Morgan Spurlock

amines the links between fast food and the
f obesity in America in a satirical and
ector Morgan Spurlock puts his theory to the
g but McDonald?s for an entire month
rules: first, he could only eat what was
inter (water included!); second, no
ffered and third, no excuses: he had to eat
nu at least once. The results are a cause for
calculates his spiralling food bill,
nd increasing visits to the doctor. SUPER
at the dubious nature of school lunches, the
ung people, physical education classes,
he extreme dieting regimes people subject
r to lose weight and improve their health. A
n look at the billion dollar fast food industry
SIZE ME is a compelling documentary for
een asked: ?Would you like fries with that??

strong language and surgical detail

l 01223 504 444 or book online

Figure 5. New features for a web client

Application of method

The study involved interviewing and logging participants in a workplace and a family for a month. The workplace chosen was a Chartered Accountancy firm, an SME with 7 employees including the owner. The family comprised 4 members: Mum, Dad, a boy aged 16 and a girl aged 11 years.

We focus our observations and analysis on the four specific themes for evaluation of new features described in the Analysis section: effectiveness, adoption, design issues, and satisfaction, linked to the user's original expectations of the novel features.

Effectiveness. In terms of characterizing the effectiveness of novel features, statistical data on the usage of features was essential. Such data was considered as an indicator of the feature effectiveness and carefully analyzed by the decision makers in development groups who had to weigh the potential benefits and decide on the adoption of the feature for future product offering.

Interestingly, in terms of perceived usefulness, some participants could see the value in a new feature even if it was not 100% effective all the time. For example, SmartBack takes the user back to hubs, typed URLs and Favorites but sometimes it jumps over the page the user wished to go to. However, the participants did not mind having to 'correct' the overly jump by following links to the desired page, as could be seen from the logs. It seems that the users saw a great value in not having to press the back button repeatedly.

Feature adoption and design issues. We identified barriers to a features' adoption. We give three examples. Firstly, a couple of participants tried to use SmartBack but because it was not apparent how it worked and it was not successful on the occasions it was first used, the participants gave up trying.

Secondly, some participants adopted the new features straightaway, whilst others continued to use existing features. It was not clear that the latter would change. For example, some participants used the back drop-down menu to check they had not missed a page of interest when they were finishing an activity. The back drop-down menu is similar to Overview, but it removes items off the list as the user navigates using Back button. Thus it provides an automatic 'clean-up' operation for the navigation stack. This is rather different from the behavior of the Overview and the users who prefer it are unlikely to adopt the Overview.

Thirdly, the features like SmartFavorites automate and interpret people's activities. Some participants expressed concern about them seeming "big-brother-ish". It became clear that we could overcome some of these barriers if we could make it apparent how these features worked. We could anticipate such a reaction from the logs since it was apparent that the new feature was used. Indeed, it was successful in predicting habitual user visits to a set of sties. However, the interviews provided insights into the user's thoughts beyond superficial observations by enabling the participants to reflect on the design and express their concerns.

We came across new situations when a feature was used that we had not predicted. For example, we noticed that pop-ups windows, if recorded in the Overview, quickly clutter the list. We also had to consider an alternative algorithm for SmartBack to avoid accidental termination of secure sessions which typically start with user login. From logs and interviews, we were able to find sufficient details about the types of interaction and redesign the feature to support such situations appropriately.

Expectations and satisfaction. Participants who became adept at using the new features were surprised and irritated when a feature did not work how they expected. For example, in order to identify a hub, we need the user to navigate away from the hub and then return to it. It is not apparent or obvious to users how SmartBack works and participants expected it work on the first navigation away from the hub.

People expressed concern about too many features in the web client and yet people used features in different ways. The owner of the SME uses Favorites as a quick list and when we asked him to rank the web client features, he ranked Favorites first because he was concerned that we would propose removing it. A major issue in re-design seems to be upward compatibility versus clarity of interaction.

In this section we have shown how we used the method to evaluate new features. The following section reflects on our use of the method in both the discovery and evaluation studies.

COMMENTS ON THE CONTEXT RE-DESIGN METHOD

In this section, we comment on the application of the method in the discovery and evaluation studies. We outline where the method was effective and how it could be improved.

As we pointed out earlier, the combination log analysis, interviews, and field observation provided two benefits. Firstly, it resolved a criticism of interviews that participants only mention what they are aware of and what they think the interviewer wants to know. For example, participants did not tend to mention activities that were unsuccessful, and yet failures were of interest to us. In the first study, we were keen to know what they were trying to achieve. In the second study, we were also interested in whether the features were working correctly. We could see problems had occurred in the logs and we could probe in the interviews. In addition, we could see patterns developing in the logs, such as someone's use of Favorites or Overview, and we could probe the participant for greater detail.

Secondly, logged information is typically hard to interpret reliably, but the interviews and field observation provided a broader understanding that allowed us to make better sense of the data. The visualization tool in the logger was extremely helpful. However, it was still incredibly time-consuming to go through participants' logs. The interviews added structure to log data and made analysis of the logs easier. For example, from the interview, it was clear there were daily patterns in people's activities and we used the logs to explore these for closely. Participants told us they visited particular sites, such as the daily newsletter, in the morning. We could see this acted as a portal site with participants performing hub and spoke navigation. As a result, we began to discern between transitional and intentional back navigation. The logs showed a surprisingly low use of search tools and so we explored how people navigated to sites. This led to our discovery of trails.

As an illustration, we provide in Tables 2-4, a detailed profile of the user and user's experience for one of the participants in the feature evaluation study. This record of issues and observations sheds more light on the potential and benefits of the hybrid approach we recommend.

On reflection, it is difficult to discern how an issue was raised because it was an organic and iterative process. However, it is clear that the exchange of knowledge gathered in the observation, interviews and logs enhanced the data we collected and provided an excellent source of quantitative and qualitative data that we could use in further analysis.

We asked participants to manually record specific events, which we were unable to capture automatically. It was common that participants forgot to complete the crib sheets. It reinforced our earlier view that we would impose too much on participants' activities to ask them to note interaction events and that logging was necessary to capture the interaction detail.

It would have been helpful to have carried out cultural research at the outset, specifically on local government organizations. Cultural research would have greatly informed our findings and enabled us to explore their generality. It is unclear how prevalent say monitoring is as an activity across other types of organization.

People either adopted the new features straight away or not at all. Participants tended to repeat themselves in interviews, searching for something to say. It is suggested that an initial interview and an interview when usage patterns have emerged are all that is needed in applying the method to the evaluation phase. We tested all the new features together. Some of them offer repeated functionality, e.g., SmartBack and Overview. It is not clear whether it would have been better to test the features individually, and if we did so, how we could combine our findings.

The data we collected was found extremely useful to the design team and was used extensively over a three year period, both within the research and development teams.

RELATED WORK

In this section, we focus on outlining methods that have extended or combined interviews, field observation and logging.

Beyer and Holtzblatt built on the interview method to develop Contextual Inquiry, which is a specific type of interview for gathering field data from users [8]. Interviewees are interviewed in their context, when doing their tasks, with as little interference from the interviewer as possible. This allows the interviewer to observe participants

LYN **Employee, 29, Chartered Accountant and Independent Financial Adviser**	
From observations and interviews	**From the logs**
Web activities She has three main activities: 1. Using specific work-related sites, such as product searches on investment sites 2. General Internet research on behalf of James, the practice owner, for clients 3. Personal use. She often servers a number of clients simultaneously, each of them for some period of time. The set of clients changes as she completes the work for individual clients.	• Lynn browses the Web most of the time during the day, for business and entertainment. On average, she visits about 86 pages and runs 8 windows sessions per day. In fact, her daily navigation may involve up to 22 sessions a day and 215 page visits per window session. She routinely visits personal entertainment sites first thing in the morning and during the lunch time. • During the day she engages in finding information from the professional sites and through Google search. Professional sites often involve logging in and may have their own search facility. Her general search activities are focused on finding home pages of organizations that she is researching.
Comments on features Lynn doesn't use Favourites and feels guilty about it. "I know I don't use my Favourites as much as I should, because they're just all in my head and it's just usually quicker just to type it all in than it is to… because it remembers them anyway. You tend to do what's quickest for you at any one time, don't you? […] There's a couple in there I have in there just because I can't remember the web site address." She has so many sites to go to, she says, "it would take longer to find it than just to type in the site name." True Favourites appeals to her. "I do use [True Favourites]. […] I've been using them quite a bit actually." When asked, if she uses True Favourites or types in the URL, she says, "If I can see them on the buttons, I'll use the buttons, because that's quicker, isn't it." She has used the prediction links on non-work browsing and says they are "quite useful". She says, "It does know me quite well. It feels slightly big brother-ish actually. It has decided what I spend my lunchtime doing, which is slightly disconcerting. It would be quite easy to get a bit worried that it was keeping a close eye on you".	***Detailed analysis of logged features*** • In 36.9% of all sessions and 11.7% of sessions from the Overflow menu. (This is compared with the use of normal back in 32% of all sessions or typing in a full URL in 24% of sessions, for example.) o Lynn's list of Favourites includes only 11 URLs. None of the links accessed by True Favourites are in her Favourites list. Thus, True Favourites provide access to links that she did not mark as Favourites although she revisits them frequently. o She is using True Favourites regularly to access 2 personal sites and the Google search page. o Her usage of True Favourites increases with time, from 32.8% of all sessions in the first three weeks to %36.9 for all four weeks. o She used Predictions from the Link Bar on two occasions, both times to access the personal entertainment page that she typically finds on the Link Bar as True Favourites.
"I haven't used [SmartBack] as much. I think perhaps one time I tried to use it and it didn't do what I expected it to so that put me off	• Lynn used SmartBack only on a couple of occasions over the 4 week period. o When searching, Lynn typically finds the relevant page with her first choice of link from the result page. Only occasionally she explores a couple of sites from the same result page. This diminishes the value of SmartBack in the search scenarios. o Furthermore, since SmartBack was not tuned to for use with sites that involve login, her first try of using SmartBack on the professional sites was not successful. It took her too far back to the login page. She did not use the SmartBack after that instance.

Table 2. Record of the user experience analysis from multiple sources.

LYN **Employee, 29, Chartered Accountant and Independent Financial Adviser**	
From observations and interviews	**From the logs**
Comments on features She thinks of Overview as a list of pages she would want to go to, "You can see the stars. If it's in the stars I use that, if not I go to [Overview]". She doesn't seem to want to go back to pages. "For me, it would be more useful just to have the site you had been to, rather than breaking it down to where you had been in that particular site. […] [The Overview list] is what I was doing at lunch but if I wanted to go back to something I was doing this morning, it is no longer on there. So if you have less detail, it would be able to keep a longer record and so it is more likely to be of use. I do tend to jump around, working on any number of clients on the same day but still need to go back there, sometime later, so if it's kept 15 different websites on record rather than that level of detail, it would be easier to find stuff." The thumbnails appear after she has chosen to click on a link. Observations of the Lyn's working habits provide explanation for her specific request on the Overview design. As Lyn works simultaneously on a number of clients she visits a number of distinct sites. They need to be revisited periodically during a period of time. If they disappear from the Overview they are not easily accessible any more.	***Detailed analysis of logged features*** • Lynn consults the Overview and clicks on the links in the Overview from time to time. She used to do that right at the beginning of a new session. That was of no use initially, since the first version of the Overview did not contain links from the previous sessions. With the change of the Overview configuration she started to use it more frequently. o Indeed, her use of Overview increased over the last week, relative to the usage in the previous three weeks and relative to the use of other features. o In the first period, the log recorded opening the Overview in about 10% of sessions, which increased to 15.6% of sessions in the last week. Also, in the last week she actively used the Overview by clicking on the link in 8.9% of sessions while before she did that in only 3.4% of the sessions. o If we look at all navigation activities other than normal link execution or auto-complete, opening the Overview menu accounts for 5% of all activities over the first 3 weeks and increased to 10% for the last week. This is contrasted with the drop down in the fully typed URLs, from 16% to 12%.

Table 3. Continued record of the user experience analysis from multiple sources.

Interpretation - LYN
Her work and personal Web use involve some of the same sites but lots of work related browsing involves new sites, typically accessed by finding the home page through Google search. True Favourites appeals to her need to easily access the same sites. She finds that typing the other sites is quicker than looking through Favorites to identify them. On the other hand, her list of Favourites is very limited and Favourite URLs rarely used.
Predictions work for her on two occasions but she is concerned about the privacy implications. This may have to do with the fact that TrueFavorites and Predictions display her personal entertainment sites. Neither of the two personal sites that she uses frequently on a daily basis is included into Favourites. Before were probably accessed by typing or using auto-complete but now that is replaced by a click on True Favorites.
She had a bad experience the first time she used SmartBack, which has put her off using it. It would benefit some of her searching activities, but this illustrates how important it is that something demonstrates its benefit in the first couple of tries.
She would prefer Overview to be a list of sites rather than detail each page within a site. It would give her a longer visible record which would give her access to sites she went to half a day ago.

Table 4. Synthesis of the evidence provided from multiple perspectives.

carry out activities. Interviewers are encouraged to do little or no analysis but to collect raw data. In contrast to this approach, we used the interviews to probe on issues that had arisen in prior analysis of the logs. We used field observation to collect field data.

Logging has been used mainly for evaluative purposes, with some loggers simply aggregating results, whilst others aiming to predict user patterns. Ivory and Hearst [9] review the state of the art in automating usability evaluation of user interfaces.

Most web servers log page requests, making server log analysis popular. Most of these tools produce aggregate reports, such as the number of transfers per date and the most popular pages. However, access to server logs is often restricted to the owners of the servers. Further, they are only able to log server interactions and not local client interactions, such as access to cached pages. Two better known server log analyzers are WebVIP [10] and WET [11]. WebVIP was specifically built to run usability tests. However, it requires a local copy of an entire site and instruments each link with special identifiers and event handling code. WET (Web-event logging Technique) requires less modification to sites but still requires each page on the server to be modified. Client-side logging requires special software and is usually operating system and web browser specific. Vividence Clickstreams is a commercial usability tool for visualizing individual and aggregate user paths through a web site that uses client-side logging [12].

Some have developed proxy-based logging to overcome some of the issues with choosing client-side or server-side logging, e.g. Web Quilt [13].

However, loggers have one main problem; they give no indication of what people are trying to do and whether they are successful. Some loggers offer participant recruitment and online surveys, such as NetRaker, to try to find out people's goals [14]. However, a survey does not allow for any exchange of knowledge between the logger and the questions in the survey.

Loggers have been used to identify behavioral patterns. Siochi and Ehrich analyzed repetition in logs to try to identify interaction patterns [15]. They indicate that the system they developed highlighted some usability problems but not the most important issues. Chi et al. [16] present a system for the analysis and prediction of user behavior and web site usability. They integrate research that has been done on human information foraging theory, information visualization and longest repeated sequence, to enable the exploration of hypotheses about complex interactions of user goals, user behaviors and web site designs. They are able to identify "way points" in navigation patterns, well-traveled paths, information needs in these paths, and predicted destinations. Their aim is to develop a system that informs the re-design of web sites. Our approach is different in several respects. We have used interviews and field observation in addition to logging to identify people's activities and their context in the use of the web. We have used the logs to identify patterns in order to develop features to support users rather than inform web site designers design around them. We have different but complementary objectives.

CONCLUSIONS

Our key contribution is a new method for the re-design of software products. We have showed how the method has been applied in two studies, to identify new features and to evaluate new features. Although the studies involve the identification of the features and then evaluation of the developed features, the method could be applied to evaluate the usefulness of new features however they are discovered.

Because the method uses logging as a data collection technique, we do not envisage the method being useful for identifying completely new products outside the scope of the interaction detail gathered. In addition, where re-design involves work re-design, better use of other methods, such as ethnography, would be more effective.

We have not compared the effectiveness of the method to other methods in terms of the findings. The findings enabled us to develop four novel features to a web client and evaluate their usefulness. It is clear that using a combination of standard methods in novel ways provided far more interesting material than if the methods had been used individually.

We applied the method to a specific product, a web client. We believe there was nothing specific about a web client that would limit use of the method to other types of product. However, we have begun applying the method to a different set of products and we will be able to report on its generality in future publications.

REFERENCES

1. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. *Human-Computer Interaction*. Addison-Wesley, Reading MA, USA, (1994).

2. Hughes, J., King, V., Rodden, T., and Anderson, H. Moving Out from the Control Room: Ethnography in System Design. *Proc. CSCW 1994*, ACM Press (1994), 429-439.

3. Rose, A., Shneiderman, B. and Plaisant, C. AN Applied Ethnographic Method for Redesigning User Interfaces. In *Proc. DIS 1995*, ACM Press (1995), 115-122.

4. Hilbert, D.M. and Redmiles, D.F. Extracting Usability Information from User Interface Events. *ACM Computing Surveys 32*, 4 (2000), 384-421.

5. Sellen, A.J., Murphy, R. and Shaw, K.L. How knowledge workers use the web. *Proc. CHI 2002, ACM Press* (2002), 227-234.

6. Tauscher, L. and Greenberg, S. How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies* 47, 1 (1997), 97-137.

7. Milic-Frayling, N., Jones, R., Rodden, K., Smyth, G., Blackwell, A. and Sommerer, R. SmartBack: Supporting Users in Back Navigation. In *Proceedings of the Thirteenth World Wide Web Conference*, ACM Press (2004), 63-71.

8. Beyer, H. & Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. San Francisco: Morgan Kaufmann Publishers, ISBN 1-55860-411-1 (1998).

9. Ivory, M.Y. and Hearst, M.A. The State of the Art in Automating Usability Evaluation of User Interfaces. *ACM Computing Surveys 33*, 4 (2001), 470-516.

10. NIST WebVIP. 1999. http://zing.ncsl.nist.gov/webmet/vip/webvip-process.html.

11. Etgen, M. and Cantor, J. What Does Getting WET (Web Event-Logging Tool) Mean for Web Usability? *Proc. Fifth Human Factors and the Web Conference*, (1999).

12. Vividence. Vividence Browser 2000. http://www.vividence.com/.

13. Hong, J.I. and Landay, J.A. WebQuilt: A Framework for Capturing and Visualizing the Web Experience. *Proc. WWW10*, ACM Press (2001), 717-724.

14. NetRaker Corporation. NetRaker Suite. (2001) Http://netraker.com/.

15. Siochi, A. and Ehrich, R. Computer Analysis of User Interfaces Based on Repetition in Transcripts of User Sessions. *ACM Transactions on Information Systems 9,* 4 (1991), 309-335.

16. Chi, E.H., Pirolli, P. and Pitkow, J. The Scent of a Site: A System for Analyzing and Predicting Information Scent, Usage, and Usability of a Web Site. In *Proc. CHI 2000*, ACM Press (2000), 161-167.

17. Milic-Frayling, N., Jones, R., Rodden, K., Smyth, G., and Frayling, A. Designing for Web Revisitation: Exploiting Structure from User Interaction and Navigation. *Microsoft Technical Report* MSR-TR-2004-97 (2004).