ECIR 2012

Barcelona, Spain

April 1st, 2012

Searching4FUN

Workshop of the 34rd European Conference on Information Retrieval

Organised by:

David Elsweiler, Max. L Wilson and Morgan Harvey

Preface

These proceedings contain the papers presented at the ECIR 2012 Searching4Fun Workshop, that took place on 1st April, 2012 in Barcelona, Spain.

People spend more and more time online, not just to find information, but with the goal of enjoying themselves and passing time. Research has begun to show that during casual-leisure search, peoples' intentions, their motivations, their criteria for success, and their querying behaviour all differ from typical web search, whilst potentially representing a significant portion of search queries. This workshop will investigate searching for fun, or casual-leisure search, and aims to understand this increasingly important type of searching, bring together relevant IR sub-communities (e.g. recommender systems, result diversity, multimedia retrieval) and related disciplines, discuss new and early research, and create a vision for future work in this area.

There are lots of other open questions relating to searching for fun and the papers presented at the workshop deal with issues such as:

- Understanding information needs and search behaviour in particular casual-leisure situations.

- How existing systems are used in casual-leisure searching scenarios.

- Use of Recommender Systems for Entertaining Content (books, movies, videos, music, websites).

- Interfaces for exploratory search for casual-leisure situations.

- Evaluation (methods, metrics) of Casual-leisure searching situations.

- The role of Emotion in Casual-leisure search

We would like to thank ECIR for hosting the workshop. Thanks also go to the programme committee and paper authors, without whom there would be no workshop.

April 2012

David Elsweiler
Max L. Wilson
Morgan Harvey

Organisation

Program Chairs

David Elsweiler (University of Regensburg, Germany)
Max L. Wilson (University of Nottingham, England)
Morgan Harvey (University of Erlangen, Germany)

Program Committee

Pertti Vakkari, Tampere, Finland
Elaine Toms, Sheffield, UK
Ryen White, Microsoft Research, USA
Leif Azzopardi, Glasgow, UK
Bernd Ludwig, University of Regensburg, Germany
Ian Ruthven, Strathclyde, UK
Daniel Tunkelang, LinkedIn, USA
Pablo Castells, Madrid, Spain
Richard Schaller, Erlangen, Germany
Stefan Mandl, Augsburg, Germany
Amund, Tveit - Atbrox, Norway
Michael Hurst - Loughborough University, UK

Table of Contents

Preface……..……I

Organisation…….II

Table of Contents……...III

Keynote Lecture

Finding without Seeking, Retrieving without Searching…….....VI
Elaine Toms (University of Sheffield)

Presentations

Session 1: Mobile Search

Rethinking mobile search: towards casual, shared, social mobile search experiences ……....1
Sofia Reis (Telefonica), Karen Church (Telefonica) and Nuria Oliver (Telefonica)

Out and About on Museums Night: Investigating Mobile Search Behaviour for Leisure Events ……...5
Richard Schaller (Erlangen-Nuremberg), Morgan Harvey (Erlangen-Nuremberg) and David Elsweiler (Regensburg)

The Information Needs of Mobile Searchers: A Framework …….....9
Tyler Tate (TwigKit) and Tony Russell-Rose (UXLabs)

Session 2: Emotion

Role of Emotion in Information Retrieval for Entertainment. ……...…12
Yashar Moshfeghi (Glasgow) and Joemon M. Jose (Glasgow).

Searching Wikipedia: learning the why, the how, and the role played by emotion …….....14
Hanna Knäusl (Regensburg)

Rushed or Relaxed? -- How the Situation on the Road Influences the Driver's Preferences for Music Tracks ..16
Linas Baltrunas (Telefonica), Bernd Ludwig (FAU-EN) and Francesco Ricci (Bozon-Bolzano)

Session 3: Browsing for Reading

Serendipitous Browsing: Stumbling the Wikipedia ……..21
Claudia Hauff (Delft) and Geert-Jan Houben (Delft)

A Diary Study of Information Needs Produced in Casual-Leisure Reading Situations. ……..25
Max L. Wilson (Nottingham), Basmah Alhodaithi (Swansea) and Michael Hurst (Loughborough)

In Search of a Good Novel: Examining Results Matter ……...29
Suvi Oksaenen (Tampere) and Pertti Vakkari (Tampere)

Keynote Lecture – Elaine Toms

Finding without Seeking, Retrieving without Searching

In information retrieval we tend to focus on the process from specific information need to desired solution that follows a lockstep path from start to finish. Yet a rich part of our information world is in the unfocused, accidental encounter with information that leads to novel findings, and enriched experiences that maybe more about the journey than the destination. This is very true of how we approach information spaces in our leisure activities and how we use our unplanned time in digital worlds. This talk will focus on the accidental encountering of people with information, how systems support (or not) the orienteering and foraging that people tend to do, and how information retrieval might provide more optimal solutions.

Rethinking mobile search: towards casual, shared, social mobile search experiences

Sofia Reis
CITI
Universidade Nova de Lisboa
2829-516 Caparica – Portugal
se.reis@campus.fct.unl.pt

Karen Church
Telefonica Research
Plaza de Ernest Lluch i Martín, 5
08019 Barcelona – Spain
karen@tid.es

Nuria Oliver
Telefonica Research
Plaza de Ernest Lluch i Martín, 5
08019 Barcelona – Spain
nuriao@tid.es

ABSTRACT
The mobile search space has witnessed phenomenal growth in recent years. As a result there has been a growing body of research aimed at understanding why and how mobile users search the Web via their handsets and how their mobile search experiences could be improved. However, much of this work has focused on addressing the many challenges of the mobile space. In this short position paper argue the need for more casual, shared, social mobile search experiences. We outline a number of open and challenging research questions related to shared, social mobile search. Finally, we present our ideas through a proof-of-concept mobile paper prototype designed to support causal mobile search and information sharing with co-located groups of friends.

Categories and Subject Descriptors
H.5.2 [**Information Systems**]: Information Interfaces and Presentation – *User Interfaces.* H.3.3 [**Information Systems**]: Information Storage and Retrieval – *Information Search and Retrieval.*

General Terms
Design, Human Factors.

Keywords
Mobile search, mobile internet, mobile web, social search, social context, casual search, shared search, collaborative search

1. INTRODUCTION
Mobile phones, once deemed as simple communications devices, have now evolved into sophisticated computing devices, offering users the ability to access a wealth of online information, anytime and anywhere.

As mobile Internet usage has increased, there has been a growing body of research aimed at understanding why and how mobile users search and browse the Web via their mobile handsets as well as how their mobile search and browsing experiences could be improved [2, 4–9, 13, 17]. However, much of this work has focused on addressing the challenges of the mobile space and enabling mobile users to find the information they need as quickly and effectively as possible.

While past research has shed key insights into mobile Web behaviours and lead to a number of great advances in mobile Web services, recently there has been a shift in the mobile world, which we believe will force the community to *re-think* the mobile Web and mobile search space. In the past mobile meant on-the-move, portable, personal and dynamic. However recent research has highlighted that (1) more and more users are accessing the mobile Web in non-mobile settings like at home or at work [2, 13] (2) mobile users are often motivated not by an exact need or urgency, but rather curiosity, boredom and even social avoidance [2, 17] and (3) mobile web access, and mobile search in particular, is often a social act, carried out among groups of people, rather than while the end-user is alone [2, 5, 18]. Given these findings, we believe it's time to devote some effort to enable mobile users to search the Web in a more casual, social setting.

In this short position paper we motivate and argue the role of shared, social search experiences in the mobile space. We highlight what we think are important and fruitful areas of research related to this new direction in mobile search. Finally, to illustrate our ideas we present examples of a proof-of-concept mobile paper prototype, which is designed to support causal search and information sharing with co-located groups of friends via their mobile handsets.

2. BACKGROUND & MOTIVATION
The gaining momentum of mobile Web and mobile search usage has also resulted in a growing body of interesting research related to understanding mobile users, mobile information needs [3, 16] and mobile Web behaviours [2, 4–6, 9, 13, 17]. In this section we highlight key takeaway messages extracted from this past work that we believe motivate a rethinking of the mobile search experience we provide to users.

2.1 Mobile does not always mean on-the-move
Recent findings suggest that mobile users often access online content in non-mobile settings. For example, a one week diary study of mobile Web access carried out by Nylander et al. [13] shows that mobile Internet access occurs mostly at home (31%). A more recent study by Church & Oliver shows that > 70% of mobile Web accesses are recorded when users are in familiar, stationary settings like at home and at work [2]. Cue & Roto [5] discovered a similar trend emerging in a series of studies they carried out between 2004-2007. That is mobile Web access is becoming a more stationary activity. These findings point to the changing pace of the mobile Web. Location-dependency isn't the only factor to consider when designed mobile services. With more and more mobile users connecting to online content while engaging in their everyday lives, we need to focus on how we can build innovative services that integrate seamlessly into their world.

2.2 Social interactions are key
Mobile phones have always been deemed as intimate, personal communications devices. They tend to be owned by one

individual and do not tend to be shared. Despite this trait, recent studies show that there is a social, shared aspect to consider in mobile environments. For example, two studies of mobile information needs have highlighted that conversations have a significant impact on the types of information needs that arise while mobile and how users choose to address those needs [3, 16]. The same is true for mobile Internet behaviours. For example, Church & Oliver have shown that in > 65% of cases, mobile search was conducted in the presence of other people [2]. Likewise, a recent study of local mobile search has shown that in 63% of cases, mobile searches took place within a social context and were discussed with someone else in the group [18].

While research on the social context of mobile search and tools to facilitate collaboration in mobile search have been limited to date [10, 11], the same is not true for general Web search [1, 12, 14, 15, 20]. Going forward we believe there will be a need to support social, collaborative online experiences in mobile environments.

2.3 Curiosity & boredom are important motivators

Although research has shown that mobile Web access is motivated mainly by awareness [17], curiosity and diversion also account for a significant proportion of mobile Internet motivations [2]. These motivations relate to the users desire to kill time, to alleviate boredom and to find out something about an unfamiliar topic (normally encountered by chance).

Searching the Internet has traditionally been viewed as driven by a specific information need in which search is considered successful if the information the user is looking for is found in a minimal amount of time. However, in casual search scenarios finding the right answer to a given query and finding that answer as quickly as possible may not be the main goals [19]. In fact, in casual search settings, the search may be considered successful even if the information the user is looking for is not found. In casual search scenarios people may browse the Web to pass time while they are idle, *e.g.* waiting for the bus. The information need may be vague or even nonexistent. Therefore, the measure of success of a casual search process is typically based on the level of user enjoyment during the search activity and/or on how long the user has been entertained for. Given that recent research in the mobile search space highlights that more and more users access content to kill time, to eliminate boredom, to satisfy their curiosity, we believe there is more opportunity to support casual search scenarios in mobile settings.

3. UNDERSTANDING THE SOCIAL CONTEXT OF MOBILE SEARCH

In this section we briefly outline results of a survey we conducted to understand more about social mobile search behavior. Survey participants were asked to recall their most recent social mobile search experience, i.e. a search conducted in a co-located group, to address a shared information need, and answer a series of questions. The questions we asked included: what they searched for, their information need, their motivation, who they were with, their relationship(s) to the people present, where they were located, what they were doing before and after the search activity, if and how they shared the search results, and if the search had any effect on their future plans.

193 participants were recruited from internal and external mailing lists, online social networks and discussion forums. All participants had to own an Internet-enabled mobile phone and must perform mobile web searches at least a few times per month. Participants ranged in age between 18-61 (average: 31, SD: 6.9). Responses were provided by 134 men (69.4%) and 59 women (30.6%) and users came from a diverse range of backgrounds, e.g. IT, engineering, sales, telecommunications, education and customer service. The majority of our participants were residents of Spain (68%) and respondents primarily used Android (40.4%) handsets to perform their searches. Finally we found that the majority of participants (87%) stated that they used mobile search in social settings at least once per week, with 54.9% of participants using it at least once a day.

Three key findings from this survey that are relevant to this position paper are as follows: (1) curiosity and alleviating boredom was the primary motivation in social mobile search (almost 50% of responses), (2) the most popular information need related to trivia and pop culture (almost 40%) and (3) mobile users tend to share results by simply speaking aloud or sometimes showing their mobile phone screen. Rarely will users hand over their phone or share the results through electronic means.

After analyzing user comments about what would improve their social mobile search experiences many users pointed to more facilities for sharing the search results easily with their peers. Here's some examples of end-user comments: "*Being able to share information through WhatsApp or applications like that*", "*Shortcuts to send the information*", "*sharing results should be a lot easier*", "*sharing the screen between all participants*", "*Some kind of co-browsing perhaps? Phone results mesh together*".

These findings combined with insights from past research shows that searching and sharing search experiences, in a casual manner, among groups of friends represents a potentially fruitful area of future research that has been largely ignored to date. In the following section we outline what we think are important and open research questions within this new direction of mobile search.

4. DISCUSSION AND OPEN RESEARCH QUESTIONS

In this section we outline a set of open research questions to frame the challenges and opportunities of developing applications to facilitate casual, shared, social mobile share:

- What types of mobile interfaces and interactions would support or enrich the "sharing experience" during social mobile search?
- How can we enrich shared search experiences in relaxed social scenarios?
- Can we make shared mobile search experiences more entertaining for end-users?
- Will users share *more* search experiences if the sharing process was simple, quick and easy?
- Does the type of content have any impact on the sharing experience? That is, will users share differently if the content is dynamic (e.g. a mobile map) versus static (a simple web-page), or if the content is textual versus visual.
- Do users have preferences in terms of how they share contents? Do users prefer to share entire pages, snippets of pages or a "print screen" type view of the page in question?
- Would users enjoy and like the ability to re-visit shared mobile search experiences? How could shared search experiences be presented to users?
- Does time, group size or the relationships within the group impact on the sharing experience?
- Do users need to share remotely, i.e. beyond co-located

groups? How might this physical distance impact on the experience?

- What are the technological challenges in building services to support casual, shared, social mobile search?

We are currently working on an early stage prototype designed to facilitate shared social mobile search in casual settings. By designing, building and evaluating this prototype, we hope we will be able to answer some of the research questions outlined previously. In the following section we present our initial ideas to support causal search and information sharing with co-located groups of friends via their mobile phones.

5. TOWARDS SHARED MOBILE SEARCH

To illustrate our ideas we present details of an early stage mobile prototype, the design challenges we face and our plans for future evaluations of this novel mobile search service. The prototype is designed to enhance social mobile search by facilitating (1) easy group identification in co-located settings, (2) options to share a variety of search elements among groups and (3) the ability to view and reminisce about past social mobile search experiences.

The software architecture we're working on consists of two components: (1) an Android application that allows users to search and share their experiences; (2) a server that synchronizes and stores all search behaviour in a database. The server will also handle group identification and coordinate a notification facility, which will inform members of the co-located group about new "shares". In addition, the server will log all the interactions between the user and the Android application for off-line analysis of user behaviour.

As a first step we worked on a number of iterations of a paper prototype. The prototype focuses on three main components, each with its own design challenges:

5.1 Easy Creation of a Sharing Session

Information sharing on mobile phones is currently a complicated process and results of our survey reveal that this is the main reason that people do not share results with one another at present. Existing mobile browsers tend to require the user to click several times in order to finally share a web page. And this sharing is normally supported via email, SMS/MMS or social media like Facebook or Twitter. Each time a user wants to share another page, the same long sequence of clicks has to be repeated all over again. Other approaches to content sharing on mobile phones rely on Bluetooth, which is well known to be a cumbersome communication mechanism for end users. The goal of our application is to make the process of mobile Web information sharing as simple as a *single click*.

The first step to achieve this goal is to detect which phones are associated with the shared search experience/session. At present we're focusing our efforts on using (1) GPS to identify all people within a given location who have the application installed and (2) a simply way for users *identified* in step 1 to confirm or verify they are a member of a *specific group*. Given that it's likely that the use case for such an application is indoors, GPS will not provide the fine level of location granularity we require. This is the motivation for employing a second step in the group identification process. For step 2, we're investigating a number of alternative approaches to confirm association with a specific group. We'd like this process to be fun and playful, therefore we're playing with the use of accelerometers, gestures, images and video. For example, one option is to ask all users within the group and at a given location to *shake their phones* within a given time period to join a group. This fun, interactive action will involve using the accelerometer within the phone.

5.2 Easy Content Sharing

Our goal is to enable mobile users to share all Web search related content with the members of their group. Figure 2 illustrates a simple paper prototype with our main thoughts on how to approach this task. Given it's likely that users will want to share a range of content types we want to provide the users the ability to (1) share a single search result or the entire page of search results by pressing an appropriate "share" button (Figure 2 (a)), (2) an entire Web page or image result (Figure 2 (b)), as well an interactive maps and addresses (Figure 2 (c)). Each time a piece of content is shared, that content is shown as a thumbnail in a bar at the bottom of the screen (Figure 2). Pressing a thumbnail opens the respective content again. The thumbnails' bar is scrollable horizontally.

(a) (b) (c)

Figure 2. Sharing different contents.

(a) (b)

Figure 3. Visualizing past shares.

5.3 Visiting past sessions

Finally, our prototype will enable users to access their past shared social search sessions. While our survey did not reveal a large proportion of users expressing a need for revisiting past sessions, this need was expressed by a few users and it's a feature we'd like to implement and explore to see if it is in fact deemed useful by end users. A past shared social search session is any session for which the user instigated a "share" or was the recipient of a "share". We are currently playing with different forms of presenting past shared search experiences to the end user. The

first method is *by time*. Figure 3 illustrates two potential approaches to grouping shared experiences by time. We could show a small thumbnail for each past share, the name of the shared content and the name of the person who shared it (Figure 3 (a)) or a larger set of thumbnails to support a more visual UI (Figure 3 (b)). Another means of showing past shared search sessions is *by group*, that is allow users to view all shared searches carried out with or among a certain group of people or with an individual. Finally, we could show past shared search sessions *by location*, that is, allow users to view all shared searches carried out at a specific place. It's likely that the choice of interface will depend on a range of factors including personal preferences.

To date, we have developed a number of iterations of a paper-based prototype and carried out design reviews with 6 users in-house to gain feedback and insights on the interface, the interaction and the core functionality. We are currently working on implementing an Android application, however, we still have a number of technological challenges to overcome. Our plan is to deploy and evaluate the application in-the-wild, among groups of friends, to learn more about shared, social mobile search behaviours in the real world.

6. CONCLUSIONS

In this position paper we motivate the need to support casual, shared, social search experiences in the mobile space through a review of past work and an outline of key findings from a recent survey of social mobile search. We highlight a set of open research questions that we think will be important for the community going forward. Finally we illustrated our initial ideas by presenting examples of a work-in-progress mobile prototype, which is designed to support causal search and information sharing with co-located groups of friends.

7. ACKNOWLEDGMENTS

This work is funded as part of a Marie Curie Intra European Fellowship for Career Development (IEF) award held by Karen Church. Sofia Reis is currently an intern in Telefonica Research. As such this work was partly funded by Telefonica Research and by FCT/MCTES, through grant SFRH/BD/61085/2009. Note that the survey portion of the work was conducted with Antony Cousin of University of Nottingham while he was an intern at Telefonica Research in Autumn 2011.

8. REFERENCES

[1] Amershi, S. and Morris, M.R., CoSearch: a system for co-located collaborative web search. In *Proceedings of CHI'08,* ACM (2008), 1647–1656.

[2] Church, K. and Oliver, N., Understanding mobile web and mobile search use in today's dynamic mobile landscape. In *Proceedings of MobileHCI '11,* ACM (2011), 67-76.

[3] Church, K. and Smyth, B., Understanding the intent behind mobile information needs. In *Proceedings IUI '09*, ACM (2009), 247-256.

[4] Church, K., Smyth, B., Bradley, K. and Cotter, P., A large scale study of European mobile search behaviour. In *Proceedings of MobileHCI '08*, ACM (2008), 13-22.

[5] Cui, Y. and Roto, V., *How people use the web on mobile devices*. In *Proceedings of WWW'08*, ACM (2008), 905-914.

[6] Kamvar, M. and Baluja, S., A large scale study of wireless search behavior. In *Proceedings of CHI '06 ,* ACM (2006), 701-709.

[7] Kamvar, M. and Baluja, S., Query suggestions for mobile search. In *Proceeding of CHI '08*, ACM (2008), 1013-1016.

[8] Karlson, A.K., Robertson, G.G., Robbins, D.C., Czerwinski, M.P. and Smith, G.R. , FaThumb: a facet-based interface for mobile search. In *Proceedings CHI '06*, ACM (2006), 711-720.

[9] Kim, H., Kim, J. and Lee, Y., An Empirical Study of Use Contexts in the Mobile Internet, Focusing on the Usability of Information Architecture. *Information Systems Frontiers*. 7, 2 (2005), 175-186.

[10] Komaki, D., Oku, A., Arase, Y., Hara, T., Uemukai, T., Hattori, G. and Nishio, S., Content comparison functions for mobile co-located collaborative web search. *Journal of Ambient Intelligence and Humanized Computing*. (2011), 1–10.

[11] Kotani, D., Nakamura, S. and Tanaka, K., Supporting sharing of browsing information and search results in mobile collaborative searches. In *Proceedings of WISE'11*, Springer-Verlag (2011), 298-305.

[12] Morris, M.R., Lombardo, J. and Wigdor, D., WeSearch: supporting collaborative search and sensemaking on a tabletop display. In *Proceedings of CSCW'10*, ACM (2010), 401–410.

[13] Nylander, S., Lundquist, T. and Brännström, A., At home and with computer access. In *Proceedings of CHI '09*, ACM (2009), 1639-1642.

[14] Paul, S.A. and Morris, M.R. , CoSense: enhancing sensemaking for collaborative web search. In *Proceedings of CHI'09*, ACM (2009), 1771–1780.

[15] Perez, J.R. Whiting, S., and Jose, J. M., CoFox: A visual collaborative browser. In *Proceedings of CIR '11,* ACM (2011), 29-32.

[16] Sohn, T., Li, K.A., Griswold, W.G. and Hollan, J.D., A diary study of mobile information needs. In *Proceedings of CHI '08*, ACM (2008), 433-442.

[17] Taylor, C.A., Anicello, O., Somohano, S., Samuels, N., Whitaker, L. and Ramey, J.A., *A framework for understanding mobile internet motivations and behaviors*. In *Proceedings of CHI '08 Extended Abstracts*, ACM (2008), 2679-2684.

[18] Teevan, J., Karlson, A., Amini, S., Brush, A.J.B. and Krumm, J. Understanding the importance of location, time, and people in mobile local search behavior. In *Proceedings of MobileHCI '11*, ACM (2011), 77-80.

[19] Wilson, M.L. and Elsweiler, D. Casual-leisure Searching: the Exploratory Search scenarios that break our current models. *HCIR'10: 4th International Workshop on Human-Computer Interaction and Information Retrieval* (2010).

[20] Wiltse, H. and Nichols, J., PlayByPlay: collaborative web browsing for desktop and mobile devices. In *Proceedings of CHI'09*, ACM (2009), 1781–1

Out and About on Museums Night: Investigating Mobile Search Behaviour for Leisure Events

Richard Schaller
Computer Science (i8)
Uni of Erlangen-Nuremberg
richard.schaller@cs.fau.de

Morgan Harvey
Computer Science (i8)
Uni of Erlangen-Nuremberg
morgan.harvey@cs.fau.de

David Elsweiler
I:IMSK
University of Regensburg
david@elsweiler.co.uk

ABSTRACT

When search behaviour is studied in information retrieval it is nearly always studied with respect to work tasks. Recent research, however, has indicated that search tasks people perform in leisure situations can be quite different. In leisure contexts needs tend to be more hedonistic in nature and often don't require specific information to be found. Instead, information is sought that can lead to a specific emotional or physical response from the user, such as feelings of being stimulated or entertained. In this paper we investigate how people behave to meet such needs in one particular leisure context. We analyse search log data collected from a large-scale (n=391), naturalistic study of behavior with a mobile search tool designed to help people find events of interest to them at the Long Night of Museums, Munich. We examine the queries submitted, establish performance metrics and investigate how spoken queries differ from those typed via the keyboard on a mobile device. The findings provide insight into how users behave in one specific casual-leisure context and lead to several open questions for future research.

1. INTRODUCTION AND MOTIVATION

Search behaviour has traditionally been studied in the context of people completing work tasks. Despite its name, a work task need not be work-related. It is simply a sequence of activities a person has to perform in order to accomplish a goal [8]. A work task has a recognisable beginning and end, it may consist of a series of sub-tasks, and results in a meaningful product [3]. Correspondingly, the models we have of information seeking behaviour tend to assume that people look for information in response to a lack of understanding or the recognition of a gap in knowledge [2] preventing the completion of the task at hand.

Based on two investigative studies, one examining information needs in the context of television viewing and the other analysing broader information behaviour reported on twitter, Elsweiler and colleagues [7] proposed a model for what they refer to as casual leisure search, which deviates from standard work-based models. According to their model, in casual-leisure situations users seek information not in response to a knowledge gap, but with the aim of being entertained or passing time. Such needs tend to be directly related to mood, physical state or the surrounding social context. A further defining characteristic of such needs is that the informational content found by users is often less important than the feelings induced by the found content and/or the search process itself.

Beyond these two studies, very little literature explicitly focuses on information seeking behaviour in casual-leisure situations. Exceptions include studies of finding fiction [12] and non-goal oriented newspaper reading [14]. To our knowledge no other naturalistic studies of information behaviour in casual-leisure contexts exist. We believe that transactional studies, such as those that have provided a rich understanding of web search behaviour [9] would be particularly beneficial, as they would provide concrete insight into how people behave to resolve such needs. If the model proposed by Elsweiler et al. is correct and people do not care what information content is about, but rather are concerned primarily with the emotional or physical response to such content then what do queries in casual-leisure situations look like? What do people try to describe with queries and how much effort do they expend in doing this? Are queries long and descriptive and are users willing to look through lots of results to find something suitable?

In this paper we describe a study designed to answer these kinds of questions. We report analyses of interaction logs for a search system supporting one specific leisure situation - the Long Night of Munich Museums, 2011. While we do not claim that the logs are representative of all casual-leisure search behaviour, they do provide an insight into how users behave in one specific casual-leisure context and a situation where the user has a high-level, hedonistic goal. Our findings represent a good starting point from which to investigate search behaviour more generally in casual-leisure situations.

2. DISTRIBUTED EVENTS

A distributed event is a collection of single events occurring at approximately the same time and conforming to the same general theme. One such event is the Long Night of Munich Museums (Lange Nacht der Münchner Museen), an annual cultural event organised in the city of Munich, Germany[1]. In addition to a diverse range of small and large museums, other cultural venues, such as the Hofbräuhaus and the botanical garden open their doors during one evening in October. Many venues organise special activities and exhibitions not otherwise available.

Visitors to the Long Night include both locals and tourists and represent a broad range of age groups and social backgrounds. In 2011 an estimated 20,000 people visited a total of 176 events at 91 distinct locations, including exhibitions, galleries and interactive events. Events take place all over the city, mostly in the city centre, but some, such as the Mu-

[1]The event is organised by Münchner Kultur GmbH (http://www.muenchner.de/museumsnacht/)

seum of the MTU Aero Engines and the Potato Museum, are located in suburbs. Special bus tours are set up to transport visitors between events.

From interviews (n=25) we conducted with people attending the evening we know that on average each visitor attends 4 events meaning that approximately 80,000 visits took place in 2011. The standard way to discover events on offer is to use the booklet that is distributed for free by the organisers and contains descriptions of all events in the order they lie along the bus tours. This booklet is necessarily large (110 A6 pages) and can be difficult to navigate.

Only a few of our interviewees reported having specific events they would like to visit. Instead, most described having the same kinds of high-level, hedonistic needs as reported in the literature [6, 15]. i.e. "to have a pleasant evening", "to enjoy time with friends", "to extend or diversify their general knowledge" etc. We will report on the interview results in detail in a future publication, but the findings seem to substantiate Elsweiler et al.'s model.

Here we want to establish how visitors to the Long Night of Museums query a search system to address these kinds of needs. We also want to know how successful they are, and identify noteworthy behaviours, problems and any potential solutions. The long-term goals of our work are to learn about behaviour in order to understand how to build better search tools and to augment existing theoretical models of casual-leisure search. We present the results of initial analyses that lead to more detailed future research questions.

3. SYSTEM

An Android app was developed to help visitors of the Long Night find events of interest to them personally. Once they have found and indicated the events they would most like to visit, the system can create a time plan for the evening, taking into account constraints such as start and end times of events, time to travel between events and public transport routes and schedules. If the user chooses more events than would fit into the available time[2], then the system tries to maximise the number of scheduled events by leaving out those that require long travel time. It is also possible for the user to manually customise the plans by adding, removing and re-ordering events to be visited. Based on the created plan, the application can lead the user between chosen events using a map display and textual instructions. Figure 1 provides some screenshots of the app[3].

The user has four ways to find events he would like to visit, namely he can: Browse events by bus route; browse events by event type (e.g. exhibitions, guided tours, interactive event, etc.); submit free-text queries, which search over the names and descriptions of the events; receive recommendations based on a pre-defined profile and collaborative filtering algorithm built into the app.

In this paper, in line with the research aims as outlined above, we focus on the way the search features were used. The search functionality was implemented in Lucene[4] and documents were represented by titles and descriptions from the Long Night booklet. Based on interviews conducted, we expected visitors to search for topics or for other high level needs not accessible for a full text search. Therefore we extended Lucene to perform a search based on topics. In a first step the event descriptions and titles were tokenised and stemmed. To match topically similar words we then map every token to one or more topic groups (these groups are taken from [4]). This way terms such as "dinner" and "food" are mapped to the same groups, thus event descriptions containing one of these words could be found by the other. To speed up interaction with the system, queries were submitted after each typed character (search-as-you-type). The presented result list contains the name and nearest bus stop for each of the retrieved events.

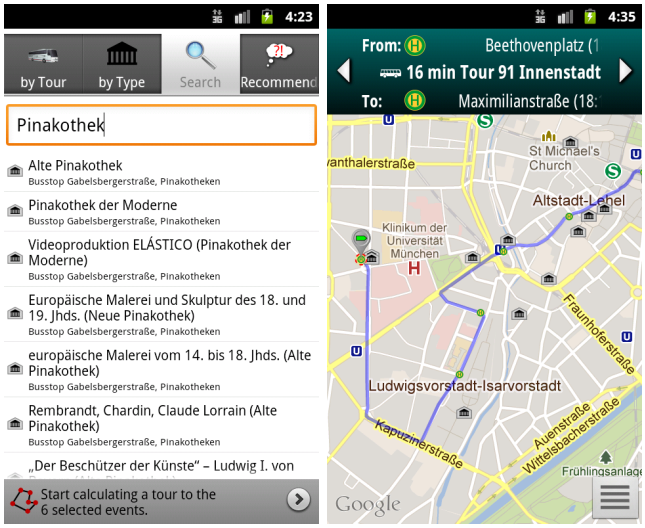


Figure 1: The search screen with a query (left) and the map screen with the planned route (right)

4. METHOD

We examined user search behaviour by recording user interactions with our app at the 2011 Long Night. The app was available for download from the Android Market and advertised on the official Long Night of Museums web page. In total the application was downloaded approximately 500 times and 391 users allowed us to record their interaction data. We recorded all interactions with the application including submitted queries, result click-throughs, all interactions with browsing and recommendation interfaces, tours generated, modifications to tours, as well as all ratings submitted for events. Users interacted on average for 45.26 minutes[5] with the system (median 19.31). 80.1% of users interacted for more than 5; 38.4% for more than 30.

A short questionnaire provided us with demographic information. 51% of the app users were first-time visitors to the Long Night of Museums, 22% were second-time visitors and 27% had attended more than twice previously. 4% of users were 17 years of age or younger, 39% were between 18 and 29, 30% 30-39, 18% 40-49, 8% 50-59 and 1% above 60 years old. These demographics are very similar to those reported by event organisers for previous Long Nights [1] suggesting that our sample of users should reflect well the visitors as a whole. Comparing both age distributions with Fisher's exact test reveals a p-value of 0.29; thus it is highly

[2]most events are open between 7pm and 2am

[3]a video demo of the application can be found on YouTube (http://www.youtube.com/watch?v=woVjpivxtMc)

[4]Lucene version 3.1. (http://lucene.apache.org)

[5]discounting times where no user interaction was recorded for more than 15 seconds

unlikely that the counts are drawn from different underlying distributions.

Since queries were submitted after every typed character, it was necessary to pre-process the recorded queries to establish those that the users actually intended to submit. For example, if the user wanted to search for "food", the system logged "f", "fo", "foo", as well as "food". Furthermore, should the user wish to submit a new query, then he must first remove the old search terms from the search box, resulting again in all prefixes but this time in decreasing length.

Automatically extracting the intended query proved difficult due to spelling errors and automatic correction. We therefore manually judged queries to be intended or not. 3 assessors separately annotated all of the approx. 10,000 queries logged as being either intended or not-intended. A high inter-assessor agreement was found (Fleiss' kappa = 0.872, 86.2% of queries which were labeled by at least 1 assessor were also labelled by at least one other assessor). This process resulted in a final list of 801 search queries, which is used in the following analyses.

5. QUERY CHARACTERISTICS

Overall the search queries were short, having a mean length of 1.21 terms ($\sigma = 0.52$) and 8.9 characters ($\sigma = 5.31$). These values are much shorter than those reported for similar mobile-like devices for web search. [10] report lengths of 2.3 terms for older mobile phones and new research suggests even longer queries (2.9 terms and 18.25 characters) for modern phones similar to those used in our study [11].

It was very apparent while analysing the queries that many represented searches for named entities, in particular the names of specific museums. Again 3 human assessors were asked to assign queries into categories: specific event name, not a specific event name or indeterminate. The third category was necessary as some queries were short and it was not possible to definitively claim that the term referred to a specific event. For example "deutsches" is likely to be a reference to the "deutsches Museum" but it is not possible to say for certain. For 87.3% of all queries at least two of the assessors were able to agree on one of the three categories (Fleiss Kappa of 0.43).

59.4% of the agreed on queries were marked as clearly named entities and 34.6% that might be named entities. Only 6.0% were labeled as non named entity searches. These remaining searches were often queries for non-museum locations, e.g. 18.2% of these are names of bus stops.

Notably absent from the logs were queries describing topical content of events e.g. "art history", "engineering", "modern art", etc. There were also no queries referring to properties of events e.g. "interactive", "talks", "discussions" and no evidence of high-level, hedonistic qualities an event might bring about e.g. "fun", "exciting", "entertainment", etc.

In line with previous query analysis papers, we analysed the diversity of submitted queries. The cleaned query set contained 417 unique queries. As expected the distribution looks rather Zipf-like with the top 2 queries being "deutsches" and "deutsches Museum". The top 50 unique queries amount to 43.1% of all queries, the top 10 amount to 16.6% and the most common search term was used in 2.5% of all searches. The entropy of the unique search terms is 2.44 bits. The queries submitted were therefore far less diverse than web search queries on desktop or mobile devices. This can be partially explained by the fact that our collection is much smaller and much more specific than the web. Another explanation for the more homogenous queries is the fact that most queries are event names which are usually only one or two words long. This reduces the possibilities for searching for these names when compared with the possibilities to express interest, constraints or needs in general.

In summary, our main observation is that the queries submitted to the search system did not reflect the information needs described in the pre-study interviews. It seems as if the users did not use the search engine to discover new events, but rather used the feature to filter to events they already knew existed. Reflecting this, our queries have similar properties to those reported for known-item searches in web, email and desktop search, which have also been shown to be very short and contain a high percentage of named-entities [5, 13].

6. QUERY PERFORMANCE

We wanted to understand how successful queries were. With this in mind we defined three success metrics based on the user's interaction with search results. The first refers to whether the user selected a returned result to read a detailed description of the event. This metric is our equivalent to click-through data. 58.4% of all searches resulted in a click-through with an average of 0.73 clicks per query ($\sigma = 0.93$) and 5.95 results on average ($\sigma = 9.10$). We didn't consider good abandonment since the result list contains no information beyond name and nearest bus stop.

Two further, more explicit, definitions of success were if the user marked a returned event as a candidate for tour inclusion (38.0% of all searches) or the user added the event to an preexisting tour (15.6% of all searches). These searches were performed at different stages of application use. Reflecting this we derived a general success metric: in 59.7% of all searches at least one of these three actions was performed. Of the remaining 40.3% unsuccessful queries 59.8% were using a search term which resulted in an empty result list, in most cases a miss-spelled or only partial written named entity. The huge number of spelling errors underlines the need for fuzzy search methods in this application context.

As the queries that were submitted were very short, we wanted to investigate if the length of the query had any impact on the success of the search. Searches defined as successful were on average longer with a mean of 1.26 terms ($\sigma = 0.57$) compared to unsuccessful searches with a mean of 1.13 terms ($\sigma = 0.42$); a highly significant difference ($p \ll 0.01$). Likewise the number of characters per query was significantly ($p \ll 0.01$) longer with the successful searches having on average 9.90 characters ($\sigma = 5.42$) and the unsuccessful searches having just 7.47 characters ($\sigma = 4.80$). We implemented a search-as-you-type system which searches for whole words, however the evidence suggests that users used the system as a means to filter to events they already knew about. Therefore while entering the search term the result list is empty till you entered the complete word. This might have led users to the conclusion that their queries will be unsuccessful and abandon the search early. This would be one explanation for the shorter query length in unsuccessful searches.

7. TYPED VS SPOKEN QUERIES

An additional feature our app offers is the possibility to submit spoken queries. Rather than typing search terms

in using the keyboard, the user speaks the query into the phone. The system uses Google Speech Recognition to identify the query terms and the user selects the queries based on a list. This is familiar to android users as it is a standard feature for web search on Android phones. We wanted to establish how this feature was used, if queries submitted in this way differed from typed queries and whether there was a notable difference in performance between spoken and typed queries.

In total 22 app users submitted 68 spoken queries, which equates to 8.5% of all search queries. Of these 6 users used it more than three times. When comparing the length of the search queries we discovered that voice searches tend to be considerably longer than typed searches: 1.8 ($\sigma = 0.65$) vs. 1.2 ($\sigma = 0.46$) terms and 14.9 ($\sigma = 8.1$) vs. 8.4 ($\sigma = 4.6$) characters. Both comparisons[6] are significant ($p \ll 0.01$). It seems it is easier to create long queries with the voice interface than typing. The success rate is also significantly higher: 75% success for speech queries compared to 58.3% (p-value[7]: 0.01) success for typed queries.

It could be that the complicated input method when typing combined with the expectation of a filtering system might have tempted people to give up early, whereas spoken queries are always full words. This would explain the ratio of empty result list where 11.8% of the voice searches have an empty result list compared to 25.2% of non-voice searches; a difference which is significant (p-value[7]: 0.013). In summary, there is evidence to suggest that voice search can be an effective tool for entering search queries on a mobile device in leisure situations. There are, however, issues such as background noise and user self-consciousness that may explain why only a limited set of users used this functionality.

8. DISCUSSION AND CONCLUSIONS

In this paper we analysed the query behaviour of users in a specific casual-leisure situation: a mobile application to assist users at a distributed event. It was apparent when analysing the queries that there was a mismatch between the queries people submitted to the search system and what we anticipated based on the needs reported in the interviews. The overwhelming majority of queries were partial or complete event names, where the user was trying to filter to a specific event. There were very few queries relating to topics that the user may be interested in e.g. "art", "history", etc. Furthermore there were no references to descriptors of events that people noted they wanted in interviews e.g. "interactive", "talks", "discussions". Likewise there was no evidence of the high-level, hedonistic qualities an event might bring about e.g. "fun", "entertainment", etc.

This poses the question: why are people using the search system in this way? Are people conditioned to do so, i.e. do they have a preconceived notion about how search engines work and only use the system in ways that reflects this? Or is it because the app has other features, such as browsing by tour or genre that might be better suited for tasks other than known-item search? To answer these questions we are currently analysing the log data for the other features of the system. A comparison with other casual-leisure search would also complement our understanding of this issue. Are there similar trends for search on YouTube, Wikipedia or the web?

[6]Wilcoxon sign rank test

[7]Two-Tailed Test of Population Proportion

Our analysis of query performance showed that a high number of spelling mistakes were made. We wonder if this is caused by environmental factors, e.g. typing on a bumpy bus or if it is caused by a high number of named entities, the spelling of which people are not familiar? Further research would be needed to differentiate between the two, however a fuzzy search feature would certainly help people who struggle with the query input. A grep-style search would further reduce this problem since users would only need to enter a few characters as opposed to whole terms. In the comparison of spoken vs. typed queries we have seen that although not used much it provides a more successful way of querying the system.

We also believe that voice-queries deserve further research. The reason behind the decision for typing or speaking a query is difficult to analyse based on the logged data. Perhaps users are shy of speaking to their smartphone in the public. Further studies would be necessary to gain a proper insight into this behaviour. The information obtained from this early study points to a number of potential avenues for further research. One plan we have is to look at different usage patterns with the system and see how they correlate with the outcomes of the evening e.g. number of events visited, the ratings of visit events, the geographical coverage of the user etc. This would provide insight into how the features of our system support casual-leisure needs.

Acknowledgments This work was supported by the Embedded Systems Initiative (http://www.esi-anwendungszentrum.de).

9. REFERENCES

[1] *Die Lange Nacht der Musik Besucherbefragung 2010.* Münchner Kultur GmbH, 2010.

[2] N. J. Belkin, R. N. Oddy, and H. M. Brooks. ASK for information retrieval: Part I. Background and theory. *Journal of Documentation*, 38(2):61–71, 1982.

[3] K. Byström. *Task complexity, information types and information sources. Examination of relationships.* PhD thesis, University of Tampere, Dep. of Inf. Studies, 1999.

[4] F. Dornseiff. *Der deutsche Wortschatz nach Sachgruppen.* DeGruyter, Berlin, New York, 2004.

[5] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins. Stuff I've seen: a system for personal information retrieval and re-use. SIGIR '03, pages 72–79, NY, 2003. ACM.

[6] D. Elsweiler, S. Mandl, and B. Kirkegaard Lunn. Understanding casual-leisure information needs: a diary study in the context of television viewing. IIiX '10, pages 25–34, NY, 2010. ACM.

[7] D. Elsweiler, M. L. Wilson, and B. Kirkegaard Lunn. *New Directions in Information Behaviour*, chapter Understanding Casual-leisure Information Behaviour. Emerald Pub., 2011.

[8] P. Hansen. User interface design for IR interaction. a task-oriented approach. In *CoLIS 3*, pages 191–205, 1999.

[9] B. J. Jansen and A. Spink. How are we searching the world wide web?: a comparison of nine search engine transaction logs. *IPM*, 42(1):248–263, 2006.

[10] M. Kamvar and S. Baluja. A large scale study of wireless search behavior: Google mobile search. In *CHI 2006*, 2006.

[11] M. Kamvar, M. Kellar, R. Patel, and Y. Xu. Computers and iphones and mobile phones, oh my!: a logs-based comparison of search users on different devices. WWW '09, pages 801–810, NY, 2009. ACM.

[12] C. S. Ross. Finding without seeking: The information encounter in the context of reading for pleasure. *IPM*, 35(6):783–799, 1999.

[13] J. Teevan, E. Adar, R. Jones, and M. A. S. Potts. Information re-retrieval: repeat queries in Yahoo's logs. SIGIR '07, pages 151–158, NY, 2007. ACM.

[14] E. Toms. Understanding and facilitating the browsing of electronic text. *J. of Human-Comp. Studies*, 52:423–452, 2000.

[15] M. L. Wilson and D. Elsweiler. Casual-leisure Searching: the Exploratory Search scenarios that break our current models. HCIR '10, Aug 2010. New Brunswick, NJ.

The Information Needs of Mobile Searchers: A Framework

Tyler Tate
TwigKit
Cambridge, UK
tyler@twigkit.com

Tony Russell-Rose
UXLabs
London, UK
tgr@uxlabs.co.uk

ABSTRACT
The growing use of Internet-connected mobile devices demands that we reconsider search user interface design in light of the context and information needs specific to mobile users. In this paper the authors present a framework of mobile information needs, juxtaposing *search motives*—casual, lookup, learn, and investigate—with *search types*—informational, geographic, personal information management, and transactional.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: Search process;
H.3.5 [**Online Information Services**]: Web-based services

General Terms
Design, Human Factors, Theory.

Keywords
Search, information retrieval, information needs, user experience, HCI, mobile, design principles.

1. INTRODUCTION
We live in a post-desktop era. In the UK alone, 45% of Internet users used a mobile phone to connect to the Internet in 2011 [7], and Morgan Stanley predicts that by 2014 there will be more mobile Internet users than desktop Internet users globally [6]. Not only are more people connecting with mobile devices, but they're also consuming more and more data. Mobile data usage more than doubled every year between 2008 and 2011, and is predicted to grow from 0.6 exabytes per month in 2011 to 6.3 EB/month in 2015 [3]. The numbers are impressive, but all it really takes is a quick glance at the people around us to recognize that mobile Internet is pervasive.

Yet the practice of designing search experiences for mobile users is still in its infancy. The challenge is much more sophisticated than simply reworking existing user interfaces to fit on the smaller screens of mobile devices, which would be to ignore the vast situational differences between desktop and mobile search. Mobile search user interfaces must be based on an understanding of the contextual factors specific to the mobile user.

Chief among those contextual factors are the information needs that give rise to mobile search activities in the first place. In this paper we propose a framework for describing the diverse range of information needs observed in mobile users. Of particular relevance to the *Search 4 Fun!* workshop is our inclusion of the *casual* category alongside traditional classifications of information needs.

2. TWO DIMENSIONS OF INFORMATION NEEDS
Mobile information needs can be assed by two criteria: search motive and search type.

2.1 Search Motive
The *search motive* describes the sophistication of the information need, along with the degree of higher-level thinking it involves and the time commitment required to satisfy it (see Figure 1). The lookup, learn, and investigate elements of motive shown below are derived from Gary Marchionini's work on exploratory search [5], while the casual element has been more recently studied by Max Wilson and David Elsweiler [9]:

- **Casual.** Undirected/semi-directed activities with a hedonistic rather than task-driven purpose.
- **Lookup.** "Known item" searching.
- **Learn.** Iterative information gathering that requires moderate interpretation and judgment.
- **Investigate.** Long-term research and planning that demands significant high-level thinking.

While lookup, learn, and investigate are informational in nature, casual activities are more experientially and hedonistically motivated, "frequently associated with very under-defined or absent information needs" [9]. Though it may be possible to describe some casual activities in terms of other motives (e.g. casual information needs that share qualities of lookup or investigation), we believe that differentiating casual from the other three motives provides both clarity and legitimization.

2.2 Search Type
The *search type*, on the other hand, is concerned with the genre of information being sought (see Figure 2). Broder is often cited for recognizing the informational and transactional nature of many needs [1], while the geographic and personal information management goals identified by Church and Smyth are especially significant for mobile users [2]:

- **Informational.** Information about a topic.
- **Geographic.** Points of interest or directions between locations.
- **Personal Information Management.** Private information not publicly available.
- **Transactional.** Action-oriented rather than informational goals.

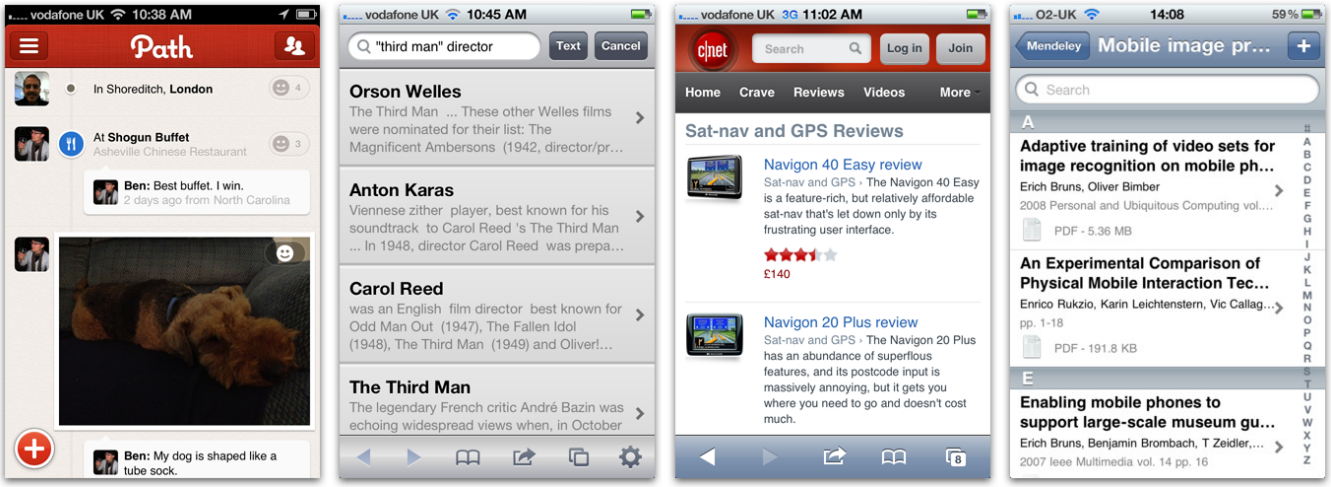


Figure 1: Path's notification screen, Wikibot's search results, product reviews on CNET, and Mendeley's personalized library of academic papers represent the casual, lookup, learn, and investigate motives, respectively.

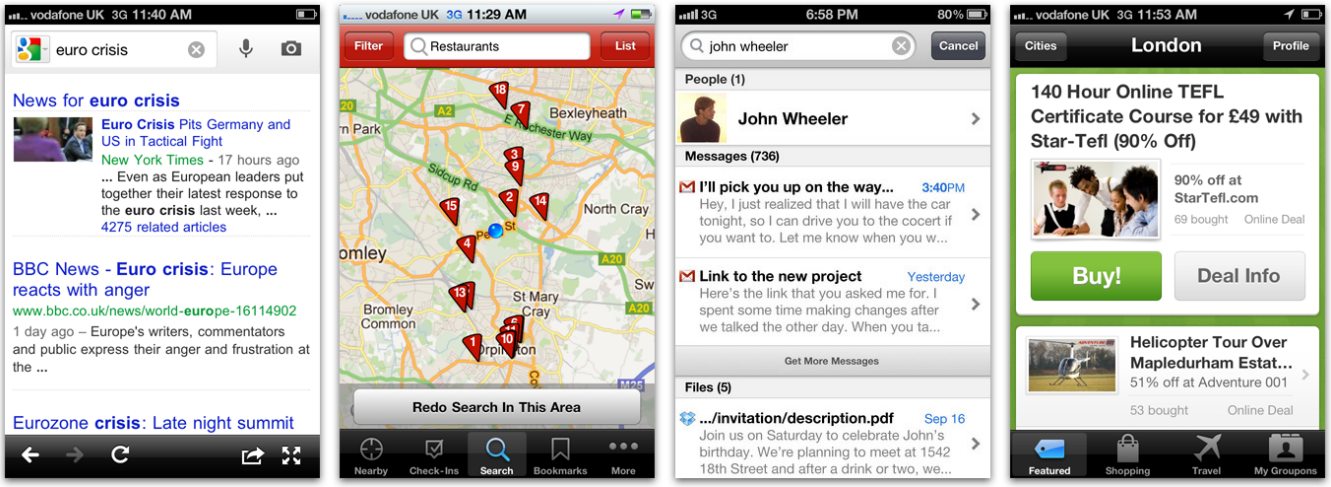


Figure 2: Google Search, Yelp, Greplin, and Groupon demonstrate the informational, geographic, personal information management, and transactional types, respectively.

3. A MATRIX OF MOBILE INFORMATION NEEDS

While the dimensions of motive and type provide a framework, they don't tell us about the information needs themselves. Fortunately, Sohn *et al*. [8] and Church and Smyth [2] have each conducted diary studies in which smartphone-equipped adults spread across the globe were instructed to record every information need that arose over a period of weeks. In addition, Cui and Roto [4] have performed a contextual inquiry study of mobile Web usage. This research enables us to construct a matrix of mobile information needs based on the motive and type dimensions (see Table 1).

The majority of the information needs in the matrix were explicitly identified in the diary studies, though we added a few of our own in order to fully populate the framework. Below are examples of each information need, with quotation marks denoting statements recorded in the original diary studies.

3.1 Informational

- **Window Shopping.** I don't know what I want. Show me stuff.
- **Trivia.** "What did Bob Marley die of, and when?"
- **Information Gathering.** "How to tie correct knots in rope?"
- **Research.** What is Keynesian economics and is it sustainable?

3.2 Geographic

- **Friend Check-ins.** "Where are Sam and Trevor?"
- **Directions.** "Directions to Sammy's Pizza"
- **Local Points of Interest.** "Where is the nearest library or bookstore?"
- **Travel Planning.** Flights, accommodations, and sights for my trip to Italy.

Table 1: A matrix of mobile information needs

	Casual	Lookup	Learn	Investigate
Informational	Window Shopping	Trivia	Information Gathering	Research
Geographic	Friend Check-ins	Directions	Local Points of Interest	Travel Planning
Personal Information Management	Checking Notifications	Checking Calendar	Situation Analysis	Lifestyle Planning
Transactional	Acting on Notifications	Price Comparison	Online Shopping	Product Monitoring

3.3 Personal Information Management

- **Checking Notifications.** "Email update for work"
- **Checking Calendar.** "Is there an open date on my family calendar?"
- **Situation Analysis.** "What is my insurance coverage for CAT scans?"
- **Lifestyle Planning.** What should my New Year's resolutions be this year?

3.4 Transactional

- **Act on Notifications.** Mark as read, delete, respond to, etc.
- **Price Comparison.** "How much does the Pantech phone cost on AT&T.com?"
- **Online Shopping.** I want to buy a watch as a gift. But which one?
- **Product Monitoring.** I know the make and model of used car I want. Alert me when new ones are listed.

4. DISCUSSION

This framework of mobile information needs originated out of an attempt to synthesize top-down HCIR concepts with bottom-up empirical data. We hope that future investigations of mobile behavior will use this framework as a conceptual point of reference when both constructing their studies and analyzing the results, which will would undoubtedly bring about iterative improvement to the framework.

While the specific information needs that we have identified are unique to the mobile context, the dimensions of search motive and search type are themselves generic. We envision future studies applying this same framework to desktop information needs, as well as comparing and contrasting desktop vs. mobile information needs.

5. CONCLUSION

In this paper we have proposed a framework of mobile information needs in order to inform the design of mobile search user interfaces.

6. REFERENCES

[1] Broder, A. 2002. *A taxonomy of web search*. SIGIR Forum, Fall 2002, Vol. 36, No. 2

[2] Church, K. and Smyth, B. 2009. *Understanding the intent behind mobile information needs.* IUI'09, February 8 - 11, 2009, Sanibel Island, Florida, USA. Copyright 2009 ACM 978-1-60558-331-0/09/02

[3] Cisco. 2011. *Cisco visual networking index: global mobile data traffic forecast update, 2010–2015*.

[4] Cui, Y., & Roto, V. 2008. *How people use the web on mobile devices. WWW 2008,* April 21–25, 2008, Beijing, China. ACM 978-1-60558-085-2/08/04.

[5] Marchionini, G. 2006. *Exploratory search: from finding to understanding*. In Commun. ACM 49 (2006), no. 4, 41–46.

[6] Morgan Stanley: Meeker, M., Devitt, S., Wu, L. 2010. *Internet trends*.

[7] Office for National Statistics 2011. *Internet access - households and individuals, 2011.*

[8] Sohn, T., Li, K., Griswold, W., Hollan, J. 2008. *A diary study of mobile information needs*. CHI 2008, April 5–10, 2008, Florence, Italy. Copyright 2008 ACM 978-1-60558-011-1/08/04

[9] Wilson, M.L. and Elsweiler, D. 2010. *Casual-leisure searching: the exploratory search scenarios that break our current models*. In Proc. HCIR'10*,* New Brunswick, NJ, USA, 28- 31. 2010.

Role of Emotion in Information Retrieval for Entertainment (Position Paper)

Yashar Moshfeghi
School of Computing Science
University of Glasgow
Glasgow, UK
yashar@dcs.gla.ac.uk

Joemon M. Jose
School of Computing Science
University of Glasgow
Glasgow, UK
Joemon.Jose@glasgow.ac.uk

ABSTRACT

The main objective of Information Retrieval (IR) systems is to satisfy searchers' needs. A great deal of research has been conducted in the past to attempt to achieve a better insight into searchers' needs and the factors that can potentially influence the success of an Information Retrieval and Seeking (IR&S) process. One of the factors which has been considered is searchers' emotion. It has been shown in previous research that emotion plays an important role in the success of an IR&S process which has the purpose of satisfying an information need. However, these previous studies do not give a sufficiently prominent position to emotion in IR, since they limit the role of emotion to a secondary factor, by assuming that a lack of knowledge (the need for information) is the primary factor (the motivation of the search). In this paper, we propose to treat emotion as the principal factor in entertainment-based IR&S process, and therefore one that ought to be considered by the retrieval algorithms.

Categories and Subject Descriptors: H.3.3 Information Storage and Retrieval - *Information Search and Retrieval* - *Information Filtering*

General Terms: Theory

Keywords: Entertainment, Search, Information Retrieval, Information Science, Emotion

1. INTRODUCTION

The idea that *IR systems help searchers to overcome their information need (IN)* is a leitmotif since the early days of IR: the main task is to locate documents containing information *relevant* to such needs. Within this view, a searcher is considered as an agent that interacts with an IR system with the intention of seeking information [3]. The *information* can be defined as facts, propositions, and concepts, as well as evaluative judgements such as opinion [6].

In this paper, we argue that standard and dominant view doesn't sufficiently consider all the possible aspects of searchers' needs. Information Science (IS) researchers have argued about the existence of needs other than IN, and discussed their roles in the cognitive aspects of human beings and in IR&S behaviour. Examples include Wilson's interrelation between physiological, affective and information needs in IR&S behaviour [6], Kuhlthau's uncertainty principle [3]; these studies have investigated the role of affective and cognitive experience of a searcher in an information seeking process model.

Although these views better capture the searchers' mind compared to the traditional view, their accounting for the role of emotion is limited to its relation with cognition in the process of satisfying an IN in an IR&S behaviour, e.g., Kuhlthau's [3] model. Therefore, emotion plays a marginal role in these views in their modelling of needs. For example, in an IR&S scenario, where searchers' task is to find documents that are topically relevant to a given query (e.g., Iraq War), the emotion that they experience during the completion of this task influences their performance and satisfaction. Other examples are those of Arapakis et al. [1] and Lopatovska [4] that investigated the use of facial expressions and peripheral physiological signals as implicit indicators of topical relevance.

Others, e.g., Wilson [6], consider a more autonomous role for affect and define *affective need* as an independent need which can motivate an IR&S behaviour. For example, gathering information to satisfy affective needs, such as the need for security, for achievement, or for dominance [6]. However, there is no operationalisation of this *affective need* suitable for use in real IR systems.

In general, the current landscape of the role of emotion in IR&S behaviour is incomplete. Moshfeghi [5] argued that people use computers for individual as well as social purposes, such as entertainment, dating, getting to know people, finding 'friends', gaming, etc., which strongly indicates that users try to satisfy needs other than information ones. The study conducted by Elsweiler et al. [2] also supported this claim. The current views of emotion in IR/IS do not sufficiently explain these types of activities accurately, even though it is clear that users search for emotionally-rich documents from the Internet to satisfy these needs.

The pervasiveness of emotionally-rich content on the web, such as movies, music, images, news, blogs, customer review, Facebook comments and Twitter, highlights the demand for such contents, and, indirectly, their role in satisfying searchers' needs. Therefore, it is important to under-

stand the IR&S behaviour backed up by an entertainment aspect. The position of this paper is that emotion is a primary motivation (either directly or indirectly) behind an entertainment-based IR&S behaviour.

The rest of the paper is organised as follows: Section 2 discusses Kuhlthau's [3] model, followed by our approach in Section 3 and discussion and conclusion in Section 4.

2. EMOTION IN IR/IS

There are many theories and models that attempt to explain the information seeking behaviour. Kuhlthau's *information seeking process* model is one of the first and most popular models to investigate the affective along with cognitive and physical aspects of a searcher in an information seeking process. She proposes that people's feelings, thoughts and actions interact within their information seeking process. Kuhlthau's information seeking process model describes the searchers' common patterns of seeking meaning from information, to extend their knowledge state on a complex problem or topic which has a discrete beginning and ending [3]. The fundamental principle behind Kuhlthau's information seeking process is the *uncertainty principle* [3]. This refers to the existence of a cognitive state which causes feelings of anxiety and lack of confidence. Feelings of doubt, anxiety and frustration are in association with vague and unclear thoughts. The model shows that during a typical information seeking process, the thoughts of a searcher become clear and consequently their confidence increases and their feeling of doubt, anxiety and frustration decrease.

Although this model is an important step towards understanding the role of emotion in IR/IS, it does not encompass many important aspects of emotion in IR. Kuhlthau considers emotion/affect as a factor influencing the information seeking process, rather than a need in itself. Moreover, Kuhlthau's model is limited by making uncertainty central, i.e., as driving the seeking process while we argue that positive or negative emotion states, high or low arousal level, such as stress or boredom respectively, could also motivate users to engage in an information seeking behaviour. Therefore, a key limitation lies in the fact that the affective side of searchers is interpreted as only being a secondary motivational source for information need. In this paper, we consider emotion as a separate need. This is explored further in next section.

3. APPROACH

The goal of this section is to argue that emotion should be considered as the *primary factor* in entertainment-based IR&S behaviour: emotion can be considered as an individual need which can motivate searchers to engage in an IR&S process. The secondary factor of emotion refers to the fact that emotion (in relation to cognition) influences every aspect of the searchers' IR&S behaviour, and can thus influence the success or failure of an IR&S process. First, we will elaborate on emotion as a secondary factor in IR&S process.

As discussed in Section 2, the secondary nature of emotion in IR&S scenarios has been investigated for a long time [3]. The results of such investigations show that (i) participants experience a burst of negative feelings due to uncertainty associated with vague thoughts, leading them to recognise that they have an information need; and that (ii) there is a positive correlation between a successful information seeking process and a decrease in these negative feelings [3]. From this point of view, not only is emotion a factor that exists throughout an IR&S process which aims to meet an IN, but also it can be considered as a need: the need to change negative feelings caused by uncertainty during the initiation phase (e.g. feelings of doubt, anxiety and frustration) to feelings of satisfaction and comfort.

When the emotion need of the searcher is to diminish the negative feelings associated with a lack of knowledge (i.e., an IN), the emotion need would be satisfied if the IN associated with it is resolved. However, in an entertainment-based IR&S process, the emotion need of the searcher is not associated with a particular IN, and is an autonomous need by itself. An example of such needs are the scenarios where the searchers are stressed and look at some clips that could help to relieve their stress, e.g., when searchers are seeking for funny clips in YouTube. Of course, one way of finding these clips is by looking at the popular (most viewed/highly recommended) videos. In such scenario there is no particular information need to be resolved, but only an emotion need.

From the above, we can now argue that emotion in an entertainment-based IR&S process acts as a primary factor, i.e. as an autonomous and important need.

4. CONCLUSIONS

In this paper, we explained the role of emotion in entertainment-based IR&S behaviour. We explained that in the normative view of IR/IS, the focus is on the satisfaction of searchers' IN. Although the role of emotion is acknowledged as a factor influencing the whole IR&S behaviour, its role was limited to the study of its influence on the process of satisfying an IN. However, emotion can be a source of motivation on its own for a searcher to engage in an IR&S process. Such scenarios have not been considered in the IR/IS community, and this motivated the definition of the emotion need concept. We argued that there are emotion needs that can motivate searchers to engage in IR&S behaviour which strictly speaking does not have an IN. The pervasiveness of the use of IR applications for the purpose of entertainment and the existence of emotionally-rich data on the web provides evidence that some information seeking behaviour can be categorised under other strategies than information need that can lead to better satisfaction of the searchers' needs. Given all these evidences, the conclusion of this paper is that emotion act as a primary factor behind entertainment-based IR&S behaviours. Finally, there is not much research about entertainment-based IR&S processes. This is due to the limitations associated with it, such as lack of datasets, evaluation methodology, metrics and procedure. An attempt to solve such limitations is a possible direction for future work.

5. REFERENCES

[1] I. Arapakis, Y. Moshfeghi, H. Joho, R. Ren, D. Hannah, and J. M. Jose. Enriching user profiling with affective features for the improvement of a multimodal recommender system. In *CIVR*, 2009.

[2] D. Elsweiler, S. Mandl, and B. Kirkegaard Lunn. Understanding casual-leisure information needs: a diary study in the context of television viewing. In *IIiX '10*, pages 25–34, 2010.

[3] C. C. Kuhlthau. A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4):339–355, 1993.

[4] I. Lopatovska. Emotional correlates of information retrieval behaviors. In *WACI'11*, pages 1 –7, april 2011.

[5] Y. Mosheghi. *Role of Emotino in Information Retrieval*. PhD thesis, University of Glasgow, 2012.

[6] T. A. Wilson. On user studies and information needs. *Journal of Documentation*, 37(1):3–15, 1993.

Searching Wikipedia: learning the why, the how, and the role played by emotion

Hanna Knäusl
Department of Information Science
University of Regensburg
93040 Regensburg
hanna.knaeusl@sprachlit.uni-regensburg.de

ABSTRACT

Searching Wikipedia has been the focus of study for an increasing number of information retrieval publications. In recent years different IR tasks have used Wikipedia as a basis for evaluating algorithms and interfaces for various types of search tasks, including Question Answering, Exploratory Search, Entity Search and Structured Document retrieval. Despite being associated with these well-defined task types, little is known about why people *actually* search wikipedia, what they try to find, how and why they try to find it or the criteria they use to define success. We argue that the way wikipedia content is generated influences the way it is used, including search behaviour. We are particularly interested in learning about affective aspects of search, which have been suggested to be an important motivating factor in wikipedia search behaviour, particularly in leisure scenarios. In this position paper we motivate the investigation of wikipedia search behaviour in the wild and present our ideas on the best way to study this behaviour.

1. INTRODUCTION AND MOTIVATION

Wikipedia[1] is a free online encyclopedia, which due to its open source design and community-based editing policy has become one of the largest reference works of all time. The large volume of information, the breadth of topics covered and open-access nature of the collection has made Wikipedia a natural target of study within the Information Retrieval research community. Wikipedia is now used as the document collection for several retrieval evaluation efforts at CLEF [4] and INEX [3] and has formed the basis of evaluations in several IR domains including:

- Question answering, e.g. [4], which attempts to provide answers to questions such as "How fast can a Cheetah run?", sometimes supplementing answers with additional relevant snippets that might be helpful to the user.
- Entity search, e.g. [2], which assumes the user has an information need that could be solved by with a list of entities that satisfy some properties. A query might, for example, indicate the type of entities to be retrieved (e.g., "castle") and distinctive features (e.g., "German", "medieval").
- Structured retrieval e.g. [3], which aims to retrieve relevant parts of documents in a collection in response to given information need.
- Exploratory search e.g. [5], whereby the user has a poorly defined information need, little knowledge of the topic of interest or is unfamiliar with the search space.

Each of these examples are associated with well-defined tasks or situations. However, it is unclear how reflective these tasks are of real-life wikipedia search behaviour. Are these the most appropriate tasks to be investigating? Are we evaluating these tasks appropriately? Are there more pressing aspects that we, as a research community, should be investigating?

As a starting point to answering these questions, in the following section, we briefly review research that informs on wikipedia search behaviour in naturalistic situations.

2. SEARCHING WIKIPEDIA

The main source of knowledge of wikipedia search behaviour comes from transaction log analyses. Sakai and Nogami [6], for example, logged user interaction with a wikipedia search interface, designed to encourage exploration and development of information needs. They discovered that information needs tend to progress and develop in small steps, usually within query type. For example, users tended to browse pages from person to person or from place to place etc. The implicit structure of wikipedia most likely encourages this behavior

Fissaha and de Rijke [1] also used log analyses to learn about wikipedia searches, distinguishing between "directed" and "undirected" searches by analysing the phrasing of queries. They [also] discovered that a large percentage of searches were undirected and exploratory in nature.

Log-based investigations such as these have the advantage of collecting large quantities of data from naturalistic situations. However, they are limited in that they say nothing about the intention of the user, his experience, or the outcome of the search. For example, the work of Wilson and Elsweiler [7] asserts that many searches will not be motivated by information needs per se, but purely by the user

[1] http://www.wikipedia.org

having an interest in a topic. In their work, they found example search tasks that were motivated by the desire to achieving a particular mood, emotional or physical state or by the presence or need of someone else in the social context. In such cases, the support the user would need from the system and the criteria that should be used to evaluate system performance would be very different to those currently featured in information retrieval research.

We believe that the way wikipedia is constructed, i.e., collaboratively by a subset of the users, the large collection size and broad topic range, linked structure, as well as multimedia prominence of multimedia content will mean that wikipedia will be used for leisure-time tasks. People are motivated to create / edit wikipedia pages as it mirrors their interests. This may not always be positive.

For example, Wilson and Elsweiler [7] describe one study participant reporting frustration that he has again wasted a lot of time aimlessly browsing ebay. This negative outcome - realised through a negative emotion - would not be considered in any current IR methodology.

In the following section we outline our thoughts on what we believe to be a more suitable study design to learn about wikipedia search tasks. We would like to use the workshop as a platform for discussion to improve on this design.

3. LEARNING ABOUT BEHAVIOUR WITH A LOG / DIARY HYBRID

We need to design a study that helps us learn about the the user's motivation for searching, his behaviour in response to this motivation, his satisfaction with the experience as well as his emotional response to the experience.

To investigate these aspects we propose combining the log based approaches scholars have used previously with user diaries. Diary Studies offer the ability to capture factual data, in a natural setting, without the distracting influence of an observer. They also offer the chance to question the user regarding his motivation to search, as well as the search process and feelings and emotions experienced during the search process.

Diary studies also have limitations. These include difficulties in maintaining participant dedication levels throughout the period of study and getting the participants to remember that situations of interest should be recorded. These negative aspects can be offset, however, through careful study design. For example, since Wikipedia is digital and accessed within a web browser, it makes sense to use a digital diary that can also be filled out in a web-browser session, perhaps as a pop up. We plan to build an extension to the Firefox web-browser that detects when a wikipedia page is accessed and if a certain time threshold has elapsed since the last diary entry, the user will be asked to record details about his information need and the motivating situation surround the search. The extension will also record interactions with wikipedia (e.g. pages viewed, search queries submitted etc.), allowing analyses similar to those published previously to be complemented by the diary study data.

To limit the irritation that filling out such a form would cause and to minimise distraction to the search process we plan only to ask two short questions at that time point. The user will be asked to give a brief description of what they are looking for and why. This will be enough information to remind them of the situation at a later time point when we ask more detailed questions regarding the experience, success of the task, how the feelings realized and the factors that influenced these. This data will be elicited through a mixture of fixed and free-form questions.

We plan to triangulate the data collected from the various aspects of our study to create a rich understanding of user needs and behaviour. For example, we plan to look at the content of visited pages; the topic and the kind of media used etc. and look to see how this relates to how participants describe their experiences. We want to see, what affects user behaviour, e.g. does the link structure or the way information is presented, certain content influence behaviour or emotions experienced. The different sources of data we will collect will help us to learn about these complicated behavioural aspects.

4. CONCLUSIONS

So what will we learn from the study and why is it important? The most important point is to find out what makes the users happy; what do they need, how do they behave to achieve these needs and emotional aspects are involved when Wikipedia is searched? An understanding of these issues will inform us on the kind of functionality a wikipedia search tool should offer. Do users want to browse to related topics? Do they like a wide range of possible interesting information or just quirky look up pieces of information as and when they are needed? The proposed study would offer the chance to answer these questions by providing naturalistic data, as well as additional comments from the participants of interest.

5. REFERENCES

[1] S. F. Adafre and M. de Rijke. Exploratory search in wikipedia. In *Proceedings SIGIR 2006 workshop on Evaluating Exploratory Search Systems*, 2006.

[2] G. Demartini, C. Firan, T. Iofciu, R. Krestel, and W. Nejdl. Why finding entities in wikipedia is difficult, sometimes. *Information Retrieval*, 13:534–567, 2010. 10.1007/s10791-010-9135-7.

[3] INEX. Initiative for the evaluation of xml retrieval, 2006. url: http://inex.is.informatik.uni - duisburg.de/2006/.

[4] V. Jijkoun and M. de Rijke. Overview of WiQA 2006. In A. Nardi, C. Peters, and J. Vicedo, editors, *Working Notes CLEF 2006*, September 2006.

[5] B. Kules and R. Capra. Designing exploratory search tasks for user studies of information seeking support systems. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, JCDL '09, pages 419–420, New York, NY, USA, 2009. ACM.

[6] T. Sakai and K. Nogami. Serendipitous search via wikipedia: a query log analysis. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '09, pages 780–781, New York, NY, USA, 2009. ACM.

[7] M. L. Wilson and D. Elsweiler. Casual-leisure searching: the exploratory search scenarios that break our current models. In *4th International Workshop on Human-Computer Interaction and Information Retrieval*, Aug 2010. New Brunswick, NJ.

Rushed or Relaxed? – How the Situation on the Road Influences the Driver's Preferences for Music Tracks

Linas Baltrunas
Telefonica Research,
Plaza de E. Lluchi Martin 5,
Barcelona, Spain
Linas@tid.es

Bernd Ludwig
University of Regensburg,
Universitätsstraße 31,
Regensburg, Germany
bernd.ludwig@ur.de

Francesco Ricci
Free University of Bolzano,
Piazza Domenicani 3,
Bolzano, Italy
fricci@unibz.it

ABSTRACT

In context-aware recommender systems, the dependency of the user's ratings on factors that describe important aspects of the recommendation context is used to provide more relevant recommendations.

Individual users may be influenced differently by the same set of contextual factors. By understanding this kind of dependency between the user's ratings (evaluations) and context, it is possible to identify user profiles and use them to predict precisely the user ratings for items to be recommended. In this paper, we present our methodology to identify user profiles in a corpus of ratings for music tracks. These ratings were collected in a user study, which simulated typical situations that occur while driving a car. We present the findings derived from the data, and argue that it is feasible to distinguish different typologies of users from the ratings they give to music tracks in specific contexts.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Information Filtering*

Keywords

Recommender Systems, Context-based Reasoning, Collaborative Filtering

1. INTRODUCTION

Recommender systems predict user ratings for items on the basis of previous ratings for similar items or similar users [5]. As users may rate the same item differently depending on the situation in which they will experience or use the item, context-aware recommender systems [4, 6, 3, 1] have become a popular research focus. The main idea is to model context as a set of variables (contextual factors) each of which can take one of a finite set of discrete values (contextual value). The user ratings are stochastically dependent on the contextual values.

For a recommender system, there is a major implication from this observation. If we can assess such an influence for individual users we are able to better personalize recommendations. Beyond this, it may even be possible to group users influenced in a similar way by certain contextual conditions. This knowledge could lead to an improved prediction of ratings for items not previously rated by the user.

With this in mind, it seems worth understanding the influence of context on user ratings. In previous work [2], we reported on a collection of ratings data for music tracks while users experienced different stereotypical situations while driving a car. In this report, we focus on the analysis of this data with respect to the aims discussed above. Whether or not a particular aspect of context is important for predicting user ratings, is dependent on the user to whom the recommendations are targeted. Our data suggest that different users have different perceptions of their surroundings and that these perceptions may influence musical preferences. Our data reveal that people assign different ratings to the same music track in different contexts and in many cases these differences are statistically significant.

Our paper is structured as follows: In the next section we briefly present our data. Next, we introduce the mathematical tools we use to analyze the influence of context on user ratings. In sections to follow, we present evidence that context can provoke a change the music genres preferences of the user. In the final section, we discuss whether or not the influence of the context on ratings can even be observed for individual users, and conclude the paper with a discussion of the results and outline our plans for future work.

2. DATA CORPUS AND CONTEXT MODEL

As described in [2], we collected two independent data samples. In these experiments, driving situations were simulated with descriptions on a website. In the first experiment, we intended to capture the influence of context on the active and conscious decision of a user to listen a tracks of a certain genre if at the same time he was exposed to a certain contextual factor. For this purpose, users were asked to focus on one context factor at a time and rate the *influence* of this context factor on their *decision to listen* to a track of a randomly proposed genre on a three-level scale (`POSITIVE`, `NEGATIVE`, or `NONE`). In this way, the decision making process in this experiment was modeled as an active modification of the user's attitude towards a genre. Over a period of three weeks, we acquired 2436 ratings from 59 users (Users were recruited via email-lists and social networks). This study was considered a pilot, and in order to avoid the sparse data

Context Factor	$MI_Y(X,Y)$
`sleepiness`	0.169766732
`traffic conditions`	0.034971332
`weather`	0.027759496
`driving style`	0.025347564
`road type`	0.022788139
`natural phenomena`	0.015574021
`mood`	0.013993043
`landscape`	0.010431354

Figure 1: Mutual Information between Influence of Context on Ratings and Context Factors

problem a small number of tracks for each genre were proposed. 95 ratings were collected per contextual factor.

For our model of context, we relied on cognitive task analyses of car driving and considered three different kinds of a driver's perceptions and actions as potentially relevant:

Context Factor	Possible Values
driving style	relaxed driving, sport driving
road type	city, highway, serpentine
landscape	coast line, country side, mountains/hills, urban
sleepiness	awake, sleepy
traffic conditions	free road, many cars, traffic jam
mood	active, happy, lazy, sad
weather	cloudy, snowing, sunny, rainy
natural phenomena	day time, morning, night, afternoon

Situations where more than one passenger was present were beyond the scope of our research.

For the second sample, we collected tracks with ratings on a five star scale. The sample consists of 955 ratings ignoring any context factor and 2865 ratings taking one contextual condition into account. The ratings were given by 66 different users (including many who had participated in the first study). 69 to 167 ratings were collected per contextual factor depending on the assumed relevance for the experiment (see Figure 1 and the discussion in Sect. 3).

3. RELEVANCE OF CONTEXT FACTORS

When analyzing the dependency between contextual factors and ratings we could not make any modeling assumptions regarding the nature of the dependency. The same holds for inter-factor dependencies. Therefore, parametric models for the dependency such as linear regression are not appropriate. Instead, we had to find a non-parametric model. In information theory, the concept of mutual information of two random variables is known exactly for this purpose: it provides means to quantify the mutual dependence of two random variables.

In our case, we can apply mutual information to quantitatively assess the difference in the average ratings for music ignoring any influence of context compared to the average rating taking single contextual factors into account. More formally, we define a random variable X for the event that users assign one of the ratings 1, 2, 3, 4, or 5 to a genre (in the first sample) or to a track (in the second sample).

Secondly, we define another random variable Y for the event that one of the context factors holds in the current situation. Mutual information (MI) between X and Y is then defined as:

$$MI(X,Y) = \sum_{y \in Y} \sum_{x \in X} P(x,y) \cdot \log \frac{P(x,y)}{P(x) \cdot P(y)}$$

MI can be normalized to the interval $[-1;1]$ by computing its value relative to the entropy of Y:

$$MI_Y(X,Y) = \frac{MI(X,Y)}{-\sum_{y \in Y} P(y) \cdot \log P(y)}$$

For X we have 2436 ratings (see Section 2 above). For each of the context factors, we collected 95 ratings. Figure 1 gives a numeric overview of the average ratings in the second data set and the impact of the single context factors on the average rating.

The results indicate that users are influenced heavily by variable driving conditions such as their own physical condition (sleepiness) and external factors such as traffic and weather. Personal factors, such as their mood, and factor not directly related to the car driving task, such as the landscape in which users are traveling, are of minor impact.

In the next step of our analysis, we wanted to understand whether the influence of context depends on the user preference for a music track. We hypothesized that if the user more strongly likes or dislike a track then his rating can be significantly influenced by contextual factors. In order to analyze this hypothesis we grouped the data into 5 partitions for each of the 5 possible ratings a user could assign to a track. I.e. the partition 1 ("the tracks disliked without considering context") contains all tracks rated with 1 (while different context factors were activated), and partition 5 ("the highly preferred tracks") contains the tracks rated with 5 in any context. Again, the influence of the context factors can be computed by measuring the mutual information and therefore the dependence between the random variable "a track is rated r without considering context" ($r \in \{1,2,3,4,5\}$) and the random variable "context factor c is active while a track is rated r". Figure 2 shows the results of this experiment. A first look at the numbers gives the impression that the mutual information is generally higher than in the experiment documented in Figure 1. To test this in a statistically sound way, we compared the mutual information values for each partition to those shown in Figure 1 using a t-test. The results are given in the last column. With the exception of partition 3 which groups the tracks that users did rate neutrally, for each partition the difference is statistically significant (the dot stands for $\alpha = 0.5$, $**$ for $\alpha = 0.01$, $***$ for $\alpha = 0.001$). These findings suggest that when users have strong positive or negative opinions for certain tracks, the conditions they experience while driving a car can influence more their ratings for these tracks.

We also analyzed the influence of context on the preferences for certain music genres. For this purpose, we analyzed the data coming from the first study (see above). We formalized the user responses (`POSITIVE`, `NEGATIVE`, or `NONE`) as a random variable I. Given this variable, the genre G and the activated context factor C given, we can estimate the probability distribution $P(I|G,C)$ from the first data set and compare it to the distribution $P(I|G)$ which does not take any context into account. For our purposes, it is again interesting to compute the mutual information for the above random variables $(C|G)$ and $(I|G)$. The following table presents the top-3 results for all combinations of genres and context factors:

	Partition				
Context Factor	1	2	3	4	5
driving style	0.145373959	0.048822968	0.18469473	0.035874718	0.028085475
landscape	0.039462852	0.025682432	0.05470132	0.042950347	0.038938108
mood	0.017266963	0.029724906	0.052830753	0.046422692	0.093026607
natural phenomena	0.022655695	0.053228548	0.084777547	0.024086852	0.082907254
road type	0.062203817	0.027293531	0.040344565	0.073388508	0.143056622
sleepiness	0.136737517	0.17566705	0.053153867	0.396715694	0.31060986
traffic conditions	0.036059416	0.121036344	0.124320839	0.032237073	0.139863842
weather	0.089973183	0.064745768	0.03265592	0.019943082	0.053972648
Level of Significance	.	∗∗		.	∗∗

Figure 2: Mutual Information between Influence of Context on Ratings (POSITIVE, NEGATIVE, or NONE) and Context Factors Given a Certain Rating *(key: '.': $\alpha = 0.5$. $*, **$: $\alpha = 0.01$)*

Blues	driving style	0.324193188
	road type	0.216609802
	sleepiness	0.144555483
Classics	driving style	0.77439747
	sleepiness	0.209061123
	weather	0.090901095
Country	sleepiness	0.469360938
	driving style	0.363527911
	weather	0.185619311
Disco	mood	0.177643232
	weather	0.17086365
	sleepiness	0.147782999
Hip Hop	traffic conditions	0.192705142
	mood	0.151120854
	sleepiness	0.105843345
Jazz	sleepiness	0.168519565
	road type	0.127974728
	weather	0.106333439
Metal	driving style	0.462220717
	weather	0.264904662
	sleepiness	0.196577939
Pop	sleepiness	0.418648658
	driving style	0.344360938
	road type	0.268688459
Reggae	sleepiness	0.549730059
	driving style	0.382254696
	traffic conditions	0.321430505
Rock	traffic conditions	0.238140493
	sleepiness	0.224814184
	driving style	0.132856064

From these results, we can learn two lessons. First, within a given genre, the mutual information is very high only for some factors. Evidently, these have a strong influence on the user ratings. This outcome was not obvious before the experiment as the user preferences could have been stronger than the influence of the driving situation. However, some of these factors influence the ratings for (almost) all genres. We may conclude that they are strongly related to the cognitive and emotional state of a driver and therefore constitute important features of recommending music in car.

Second, as the influence of context is evident, we may conclude that even users with strong preferences for certain tracks may change their opinion if they experience their driving situation intensively enough.

4. INDIVIDUAL USER TYPES

We now investigate the influence of context on individual users. We analyze the user ratings of the four users who gave most of the ratings in our second data collection phase (see above). We show that different contextual factors can influence different users in different ways. In the following tables, *Mean with context* (MCY) is the average rating of a user for all items rated under the assumption that the given contextual factor holds. *Mean without context* (MCN) is the average (of all users) rating for the same items without considering context. Differences in these averages are compared using a t-test in order to assess whether a contextual factor actually influences the user's ratings in a significant way. We indicate the statistical significance of the difference between MCY and MCN with the p-value of the t-test.

We note that a recommender system can exploit the results of our data analysis when building a prediction model that integrates the average rating of many users for an item, a personalized component for a particular user, and a component for the context (see [2] for details).

User 1: Preferences above Average.

As can be seen in column MCN in Table 3b, this user, on average, rated the tracks in the data base higher than the others. The comparison with MCN of all users (see Table 3a) suggests that for this user many of the tracks were perceived very positively in driving situations demanding the driver's attention. In fact, driving on a highway, on a serpentine or mountain road leads to an increase of the average rating (compared to MCN for all users). On the other hand, situations that can be perceived as negative (e.g. traffic jam) provoke a decrease of the user ratings. This observation similarly holds for some other factors: *lots of cars*, a situation quite similar to *traffic jam*, or driving in *morning* time. Interestingly, *sport driving* – which stands for a consciously sportive style of driving – has negative influence on the average ratings of this user. Hence we hypothesize that the user is affected negatively by the tracks (mainly pop music) in situations that are likely to produce stress.

User 2: Preferences around Average with Positive Tendency towards Tracks.

In this example the user has a personal average rating similar to the other users. This phenomenon is not an ef-

Factor	MCN	MCY	Tendency	α
highway	2.498429	3.521739	↑	* * *
traffic jam	2.498429	1.647059	↓	*, *
city	2.498429	3.800000	↑	* *
serpentine	2.498429	3.529412	↑	* *
sport driving	2.498429	1.705882	↓	* *
lots of cars	2.498429	1.894737	↓	* *
coast line	2.498429	3.500000	↑	*
mountains/hills	2.498429	3.307692	↑	.
active	2.498429	1.866667	↓	.
country side	2.498429	3.272727	↑	.

(a) MCN of all Users versus MCY for User 1

Factor	MCN	MCY	Tendency	α
traffic jam	3.077586	1.647059	↓	* * *
lots of cars	3.077586	1.894737	↓	* * *
sport driving	3.077586	1.705882	↓	* * *
active	3.077586	1.866667	↓	* *
morning	3.077586	2.000000	↓	* *
city	3.077586	3.800000	↑	*

(b) MCN versus MCY of User 1

Figure 3: Profile of User 1. Only those factors with statistical significance are shown.

Factor	MCN	MCY	Tendency	α
happy	2.498429	1.444444	↓	* *
serpentine	2.498429	1.709677	↓	* *
urban	2.498429	1.760000	↓	*
awake	2.498429	3.642857	↑	*
country side	2.498429	1.807692	↓	*
sad	2.498429	1.846154	↓	*
afternoon	2.498429	2.000000	↓	.
relaxed driving	2.498429	2.025641	↓	.

(a) MCN of all Users versus MCY of User 2

Factor	MCN	MCY	Tendency	α
happy	2.432692	1.444444	↓	* *
serpentine	2.432692	1.709677	↓	*
awake	2.432692	3.642857	↑	*
urban	2.432692	1.760000	↓	*
country side	2.432692	1.807692	↓	.
sad	2.432692	1.846154	↓	.

(b) MCN versus MCY of User 2

Figure 4: Profile of User 2. Only those factors with statistical significance are shown.

fect of any context. The sign of the significant differences between MCN and MCY in Table 4a indicate that this user likes the tracks in the corpus when he feels *awake*. Being *sad*, he would never like to listen to the tracks. In general, for this user the traffic situation (differently from user 1) seems to play a minor role. Many significant differences in his ratings can be found comparing his MCY with his non-contextualized ratings (own MCN) as well as with the rating of all the users (MCN), for personal factors such as the mood and the perception of the surrounding landscape.

User 3: Preferences slightly below or on Average with Negative Tendency towards the Tracks.

In this user profile, the factors provoking significant differences between MCN and MCY (see Table 5a) are mostly personal ones or factors that indirectly influence personal attitudes or the cognitive load of the driver (i.e. road type).

As many of the tracks used for our data collection were pop songs, and on average the user assigns low ratings, we can conclude that he has a strong dislike for this kind of music. This impression is strengthened by the observation that negative emotions (such as sad) lead to even worse ratings for tracks than on average for this user.

User 4: Preferences below Average.

In this user profile, there are several highly significant differences between the MCN of all users and MCY (see Table 6a). In every case, the tendency is negative indicating that there are almost no situations in which tracks from the data set should be recommended to such a user. Probably this user does not like the tracks in the corpus, or he even does not like to listen to music at all while driving. The significance level of the difference between the personal MCN and MCY (see Table 6b), here is slightly smaller than in the previous comparison. Moreover, there is one personal factor (awake) under which the user rated significantly higher. But, as there are many factors with almost identical ratings to the already low non-contextualized ratings, in most situations the items should not be recommended to this user. From this observation, we can assume that as this user dislikes tracks very strongly, it is hard to find context factors that may change his attitude.

5. CONCLUSIONS AND FUTURE WORK

We have presented a non-parametric approach to assess the impact of a set of contextual factors on the user ratings. Our findings from the analysis of two data collections suggest that the perceptions and experiences during the execution of a task influence user preferences even for non-crucial items such as music tracks to be played in a car.

5.1 Influence of Context

We found empirical evidence that the driving situation indeed influences the driver's preferences for music. The influence of context may even be strong enough to modify the preference of a user for his favorite tracks.

The findings also suggest that the cognitive load of the driver, his emotional, mental, and physical state, and current traffic conditions influence his preferences.

These findings are surely affected by the set of tracks used in the study. We used this set as the reported experiments were developed within an industrial project, and the tracks were provided by the media platform of the industrial partner. It is an interesting task to collect data for other set of tracks – in a wider set of types of tracks or with a different specialization – and repeat the analysis.

Factor	MCN	MCY	Tendency	α
sad	2.498429	1.333333	↓	∗∗
day time	2.498429	1.666667	↓	∗∗
active	2.498429	1.769231	↓	∗
serpentine	2.498429	1.714286	↓	∗
coast line	2.498429	2.000000	↓	.

(a) MCN of all Users versus MCY of User 3

Factor	MCN	MCY	Tendency	α
sad	2.329787	1.333333	↓	∗∗
day time	2.329787	1.666667	↓	∗
active	2.329787	1.769231	↓	.

(b) MCN versus MCY of User 3

Figure 5: Profile of User 3. Only those factors with statistical significance are shown.

Factor	MCN	MCY	Tendency	α
day time	2.498429	1.166667	↓	∗∗∗
afternoon	2.498429	1.666667	↓	∗∗
highway	2.498429	1.700000	↓	∗
urban	2.498429	1.769231	↓	∗
morning	2.498429	1.714286	↓	.
mountains/hills	2.498429	1.714286	↓	.
country side	2.498429	1.700000	↓	.

(a) MCN of all Users versus MCY of User 4

Factor	MCN	MCY	Tendency	α
day time	2.175676	1.166667	↓	∗∗∗
awake	2.175676	3.222222	↑	.
afternoon	2.175676	1.666667	↓	.

(b) MCN versus MCY of User 4

Figure 6: Profile of User 4. Only those factors with statistical significance are shown.

5.2 Critical Discussion of the Study Design

It is important to note the constraints and conditions of our study design. First of all, in the web survey, we created fictive situations that the subject should imagine. Hence, the test persons may have overestimated the relevance of the contextual factors on their music preferences. Hence, a different study where users are actually facing certain contextual conditions is in order. But before performing that evaluation, our study clearly indicates that users perceive context as important and influential, and different users, with different music preferences, have completely different perceptions. To assess this result quantitatively, the web survey and the described methods represent a simple way to collect and analyze data. In fact, we exploited our results in the implementation of a real music recommender system and player [2]. Besides, it is also important to note that during our study users rated the music tracks just after listening to them. This is not always the case in many recommender systems (e.g. MovieLens or Netflix), where often the ratings are provided long after the user experienced the items.

5.3 Consequences for Future Work

Currently, we are preparing a new study with an improved experimental setup: we are merging our prototype with another application that allows to log onboard data in a car. We will equip cars of test persons with this tool and collect data in real driving situations. The logged data will allow us to detect the values of certain contextual factors from onboard information about the car and its navigation system. Furthermore, we will be able to combine this data with feedback from the users (e.g., which of the recommended tracks are played or skipped). From such a new collection of data, gained in a naturalistic setting, we will validate the findings of our simulation study.

6. REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 217 – 250. Springer, 2011.

[2] L. Baltrunas, M. Kaminskas, B. Ludwig, O. Moling, F. Ricci, A. Aydin, K.-H. Luke, , and R. Schwaiger. Incarmusic: Context-aware music recommendations in a car. In *(to appear) Proceedings of the 12th International Conference on Electronic Commerce and Web Technologies*, 2011.

[3] L. Baltrunas, M. Kaminskas, F. Ricci, L. Rokach, B. Shapira, and K.-H. Luke. Best usage context prediction for music tracks. In *2nd Workshop on Context-Aware Recommender Systems*, 2010.

[4] A. Chen. Context-aware collaborative filtering system: Predicting the user's preference in the ubiquitous computing environment. In T. Strang and C. Linnhoff-Popien, editors, *Location- and Context-Awareness*, volume 3479 of *Lecture Notes in Computer Science*, pages 244–253. Springer Berlin / Heidelberg, 2005.

[5] Y. Koren and R. Bell. Advances in collaborative filtering. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*. Springer, 2011.

[6] G.-E. Yap, A.-H. Tan, and H.-H. Pang. Discovering causal dependencies in mobile context-aware recommenders. In *MDM 06: Proceedings of the 7th International Conference on Mobile Data Management*, page 4, Washington, DC, USA, 2006. IEEE Computer Society.

Serendipitous Browsing: Stumbling through Wikipedia

Claudia Hauff and Geert-Jan Houben
Web Information Systems
Delft University of Technology
Delft, the Netherlands
{c.hauff,g.j.p.m.houben}@tudelft.nl

ABSTRACT

While in the early years of the Web, searching for information and keeping in touch used to be the two main reasons for 'going online', today we turn to the Web in many different situations, including when we look for entertainment to pass the time or relax. A popular tool to facilitate the users' desire for entertainment is StumbleUpon, which allows users to "stumble" through the Web one (semi-random) page at a time. Interestingly to us, many StumbleUpon users appreciate being served Wikipedia articles, which are informative pieces of text that educate the reader about a particular concept. The leisure activity of stumbling can thus also incorporate a learning experience. Since life-long learning is an important characteristic of knowledge economies, it is crucial to understand the interplay between these two - at first sight - opposing forces. We hypothesize that a greater understanding of what makes certain Wikipedia articles more attractive to the serendipitously browsing user than others, will enable us to develop adaptations that expose a greater amount of Wikipedia articles to the leisure seeking user.

Categories and Subject Descriptors: H.3.3 Information Storage and Retrieval: Information Search and Retrieval
General Terms: Human Factors, Experimentation
Keywords: free-choice learning, educational leisure, serendipitous browsing

1. INTRODUCTION

In the early years of the Web, searching for information and keeping in touch used to be the two main reasons for 'going online'. Today, we rely on the Web in increasingly diverse situations including shopping, consultations and learning. While these examples are all directed towards a particular goal the user has, we also turn to the Web at times when we simply want to be entertained to pass the time or relax. The possibilities for entertaining yourself on the Web are manifold, one can play games, listen to music, watch movies or simply browse through the Web in the hope of finding entertaining pages. Due to the sheer size of the Web though, random browsing is not effective for discovering pages that may b interesting to the individual user. For this reason, a number of services have become popular that recommend web pages to users based on their interests. One popular tool to facilitate the users' desire for entertainment by serendipitous browsing is StumbleUpon[1] (SU), which allows users to "stumble" through the Web one (semi-random) page at a time. Interestingly to us, many SU users appreciate being shown Wikipedia[2] articles, which are informative pieces of text that educate the reader about a particular concept. The leisure activity of stumbling thus can also incorporate a learning experience, which might contribute to the development of novel ideas and lead to creative insights. Since life-long learning is an important characteristic of knowledge economies, it is crucial to understand the interplay between these two seemingly opposing forces (entertainment vs. learning). We hypothesize that a greater understanding of what makes certain Wikipedia articles more attractive to the serendipitously browsing user than others, will enable us to develop adaptations that expose a greater amount of Wikipedia articles to the leisure seeking user.

In this position paper we make an argument for the importance of this task. We draw from a number of insights gained in museum studies [11] where the question of how learning can be facilitated in leisure settings (the museum visit) has been investigated for many years. While we do not consider the SU pages to be similar to museum objects, we do find a number of parallels.

A first experiment on the stumbled Wikipedia pages revealed that, just as in museums not all objects are equally attractive to visitors, not all articles are interesting to the average StumbleUpon user. In fact, only a very small number of Wikipedia articles gather a large number of views by SU users, most articles are rarely viewed. While we have no answer yet to the question of how to automatically classify articles according to their attractiveness to the serendipitously browsing user, we have developed a number of hypotheses which are outlined in Section 3.2.

If we assume for a moment that we are indeed able to develop such an approach, a number of application scenarios can be envisioned:

- A qualitative study of the features that play a role in to trickling the interest of users who do not have an information need, will enable Wikipedia contributors to write their articles in a way that is more accessible to such users.
- Wikipedia is available in many different languages and such a prediction method would allow us to bootstrap a recommender like StumbleUpon in different languages by adding an initial set of interesting, high quality pages before the critical mass of users is reached.

[1] http://www.stumbleupon.com/
[2] http://www.wikipedia.org/

- Outliers (articles with many 'Likes' but a low probability of being attractive) can be manually investigated to reduce spam. Or conversely, undiscovered articles are obtained and can be injected into the index.
- The passages that trigger the surprise or the attractiveness of an article can be identified and highlighted to the browsing user. This may help to keep those serendipitously browsing users engaged that initially only quickly scan the article.
- E-learning applications can also benefit, as articles which are interesting to the casual reader can be found this way.

The rest of the paper is organized as follows: related work is presented in Section 2, followed by a preliminary analysis of stumbled Wikipedia pages (Section 3) and the conclusiosn (Section 4).

2. RELATED WORK

For this work, we draw inspirations from two areas. On the one hand we consider research into so-called *educational leisure settings* and *free-choice learning* which is a multidisciplinary field that includes aspects from sociology, psychology and education. On the other hand, our work is also strongly related to serendipity.

Education leisure settings can be found in a wide range of institutions including museums [12], national parks, zoos, science centers [5], etc. As the name suggests, these institutions serve two purposes: to educate the public as well as to provide an entertaining experience to the visitors. Education leisure settings can be characterized by a number of commonalities with respect to the visitors and their learning experience [9, 10, 11]: (i) the visitors gain direct experience, (ii) they decide what and whether at all to learn, (iii) the learning process is guided by their interests, (iv) learning is influenced by the visitors' social interactions and (iv) the visitors are a highly diverse group, with different educational backgrounds and prior knowledge. Since learning in this setting is voluntary, the visitors' motivation plays an important role: why did they come?

Serendipity, the act of encountering information nuggets unexpectedly, has mostly been investigated in the context of education [3] and work-related discoveries after serendipitious moments. One of the works outside of this realm is [6] where tools were developed to help people reminisce in their own digital collections. In goal-directed Web search the potential for serendipitous encounters has also been recently investigated [2], while [1] offers an insightful discussion of serendipity and how it is used, exploited and induced in computer science.

Finally we note that different aspects of Wikipedia articles have also been investigated in the past, though not from a perspective of serendipitously browsing users. For instance, in [7] it was found that the writing style distinguishes so-called featured articles in Wikipedia[3] from unfeatured articles. Classifying Wikipedia articles according to their quality, as defined by Wikipedia contributors, was also investigated in [13], where network motifs and graph patterns in the editor-article graph were exploited.

3. STUMBLEUPON

[3]Featured Wikipedia articles are of particularly high quality and chosen by Wikipedia editors.

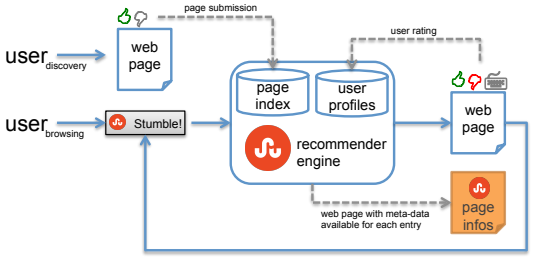


Figure 1: A StumbleUpon user can contribute Web pages he likes to the index and he can "stumble" pages that are in the SU index according to his interests. One page at a time is shown; the user can provide feedback in terms of like and dislike.

The usage of StumbleUpon is depicted in Figure 1. A user "stumbles" pages with a simple click of the 'Stumble!' button in his browser toolbar. In response, the user is presented with a random page from the Web, biased according to his user profile or his friends' 'Likes'. The simplicity of the system protects the user from information overload [8, 4], a user has only two choices when faced with a stumbled page: either to start reading or to continue stumbling. Users can also contribute pages to the SU index: whenever a SU user discover a web page that is not yet in the index and that he likes, he can add it by means of the 'Like' button. Finally, for each page in the SU index, there is a SU page which contains meta-data, including the number of users who viewed/liked the page, the category the user who discovered the page placed it in and the comments users left about the page.

3.1 Wikipedia Articles in StumbleUpon

In all experiments we report here, we utilize the English Wikipedia dump *enwiki-20111007* from October 2011. In a pre-processing step, we selected all Wikipedia articles that are neither redirects to other articles, nor new articles or explicit disambiguation pages and have a length of at least 500 characters (to remove stubs). In total, $3,552,059$ articles remained.

In order to determine the popularity of Wikipedia articles in StumbleUpon, we randomly selected half of these Wikipedia articles and queried the StumbleUpon API for their number of views by SU users. Since SU is a recommendation engine, we can safely assume that the highly viewed pages are also highly popular and liked. We note, that the number of 'Likes' a page has received is not accessible through the StumbleUpon API. The information is accessible though at the SU meta-data page, which we manually checked for the results reported in Table 1.

Among the evaluated $1,776,029$ articles, we found $267,958$ (15.13%) of them to be contained in the SU index. In our initial investigation, we also considered French and German Wikipedia which are two of the largest non-English Wikipedia repositories. However, we only found a very limited number of their articles in the SU index (in both cases less than 1%) and thus did not consider them further. Thus, an application scenario as proposed in the introduction (to bootstrap a recommender for a new language) is highly desirable.

Let us now focus on those articles that were submitted

by Stumblers to the index. Figure 2 shows a scatter plot of the number of views versus the number of Wikipedia articles in the index. As can be expected, most articles have very few views (the median number of views is 10) while a small number of articles have gathered more than half a million views.

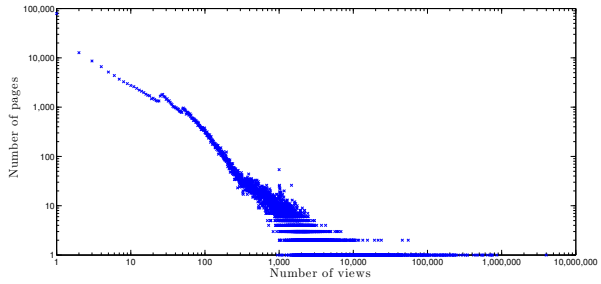


Figure 2: Log-log scatter plot of the number of views versus the number of articles in the SU index.

To give an impression of the type of articles that have gathered few or many views, Table 1 contains the ten most viewed Wikipedia articles in our data set as well as ten random examples of articles that were viewed one hundred times. We chose these two settings as they represent two extremes: on the one hand, articles that were viewed and also liked by a large number of people and on the other hand articles, that were shown a number of times but less well received by the SU users.

It should also be noted that the SU category *Bizarre & Oddities*, which dominates the list of the ten most viewed articles is not as prevalent when considering a larger set of articles. In fact, the top 100 viewed articles in our data set belong to 59 different SU categories: *Bizarre & Oddities* occurs 12 times, followed by the *Writing* category (5 times) and a number of categories with three occurrences, including *Arts*, *Science* and *Linguistics*. Only one of the top 100 articles was a so-called featured article (indicating that previous work on featured article prediction, e.g. [7], might not be applicable here), while seven were semi-protected articles due to previous vandalism activities. Notable is also the fact that 12 out of the 100 articles are of the form *List_of_X* where $X = \{algorithms, legendary\_creatures, band\_name\_etymologies\}$ to name three examples.

While for a human reader it is usually not difficult to quickly judge whether an article is potentially interesting to him or not, it is a challenge to derive a method that automatically classifies articles accordingly. What exactly makes one article more interesting to the general public than another? In order to get get a first understanding of what users think about the most viewed articles and possibly also why they like them, we analysed the comments that were posted on the SU info page for each of the ten most viewed Wikipedia articles. This analysis is very cursory, as compared to the number of views, very few users actually comment on an article, as commenting distracts from the 'stumbling' experience. For example, the article *Wrap_rage* with 0.86 million views and forty-thousand likes has a 41 comments. In total, we analysed 479 comments and identified four broad categories:

(A) Comments expressing surprise
- "There's a name for this?"
- "I'd never heard of this before (go StumbleUpon!). Very cool."

(B) Comments expressing admiration, sadness, sorrow, etc.
- "That's so sad"
- "No one should go through life afraid to take a walk."
- "don't know what to say actually.."

(C) Comments about the usefulness of the knowledge
- "Simple, but helpful for designers."
- "An exceptional list of colours and their code, invaluable to graphic designers, webmasters etc."

(D) Comments expressing negative sentiments towards the article
- "Fake."
- "Why stumble everyday wikipedia articles?"

3.2 Working Hypotheses

Based on the preliminary qualitative insights gained, we developed three intuitions that we believe will enable us to predict to what a Wikipedia article is likely to be beneficial to the average SU user.

Intuition A. Articles that contain unexpected nuggets of information can be identified by considering how semantically related the article is to the other articles it contains links to. For instance, the *List_of_unusual_deaths* Wikipedia article has, among others, outgoing links to the following diverse articles: *Common_fig*, *Malvasia* (wine), *Eddystone_Lighthouse*, *Hawaii*, and *Chimney*. We hypothesize that finding such seemingly unrelated articles can be used as a measure of the likelihood of the article being of interest.

Intuition B. Articles that evoke emotional feelings can be discovered through a form of sentiment analysis. Although Wikipedia articles are written in a neutral style, some topics are bound to evoke emotions and those emotional topics can be identified.

Intuition C. Articles that contain useful knowledge may be identified indirectly, when considering their Talk pages, the amount of discussions that are ongoing and the style of the discussions. Articles about practically useful information are not likely to be emotionally charged, unlike discussions for instance about politicians, religious topics, etc.

We emphasize, that these are hypotheses that need to be verified in future work.

4. CONCLUSIONS

In this position paper we have proposed to investigate what makes certain Wikipedia articles interesting to users who are browsing the Web without a goal in order to pass the time or relax. Since such articles are education to some degree, the leisure activity of browsing (stumbling) can thus also incorporate a learning experience. Since life-long learning is an important characteristic of knowledge economies, it is crucial to understand the interplay between these two

Most viewed articles	#Views	#Likes	SU Category	Date	Example articles viewed 100 times	SU Category
List_of_unusual_deaths	3.99M	0.423M	Bizarre/Oddities	12/2004	Biblioscape	Software
Flying_Spaghetti_Monster	1.39M	0.121M	Satire	08/2005	Edge_of_chaos	Chaos/Complexity
Wrap_rage	0.86M	0.040M	Bizarre/Oddities	01/2008	Gottfried_Wilhelm_Leibniz_Prize	Biology
Shigeru_Miyamoto	0.75M	0.019M	Video Games	10/2003	Mario_Buda	Crime
Benjaman_Kyle	0.74M	0.051M	Bizarre/Oddities	12/2008	Proto-Indo-European_language	Linguistics
One_red_paperclip	0.72M	0.070M	Bizarre/Oddities	09/2006	Cisco_Adler	Alternative Rock
List_of_colors	0.70M	0.066M	Arts	01/2005	Biofeedback	Psychology
Do_not_stand_at_my_grave_and_weep	0.64M	0.132M	Poetry	10/2007	Ovipositor	Sexual Health
Fuel_cell	0.56M	0.009M	Science	06/2005	Concealer	Beauty
Raymond_Robinson_(Green_Man)	0.54M	0.036M	Bizarre/Oddities	05/2008	Winklepickers	Fashion

Table 1: A list of Wikipedia articles that are contained in the SU index. For the most viewed articles, shown are also the number of views and likes in million, the category in StumbleUpon the page was assigned to by the user who discovered the page and the date (month/year) at which the page was discovered.

forces. We argue that a greater understanding of features are indicative of an article's attractiveness to the average user (stumbler) will enable us to develop adaptations that expose a greater amount of Wikipedia articles to the leisure seeking user.

5. REFERENCES

[1] P. André, m. schraefel, J. Teevan, and S. T. Dumais. Discovery is never by chance: designing for (un)serendipity. In *C&C '09*, pages 305–314, 2009.

[2] P. André, J. Teevan, and S. T. Dumais. From x-rays to silly putty via uranus: serendipity and its role in web search. In *CHI '09*, pages 2033–2036, 2009.

[3] L. Björneborn. Design dimensions enabling divergent behaviour across physical, digital, and social library interfaces. In *Persuasive Technology*, volume 6137, pages 143–149. 2010.

[4] D. Bollen, B. P. Knijnenburg, M. C. Willemsen, and M. Graus. Understanding choice overload in recommender systems. In *RecSys '10*, pages 63–70, 2010.

[5] J. H. Falk and M. Storksdieck. Science learning in a leisure setting. *Journal of Research in Science Teaching*, 47(2), 2010.

[6] J. Helmes, K. O'Hara, N. Vilar, and A. Taylor. Meerkat and tuba: Design alternatives for randomness, surprise and serendipity in reminiscing. In *Human-Computer Interaction - INTERACT 2011*, volume 6947, pages 376–391. 2011.

[7] N. Lipka and B. Stein. Identifying featured articles in wikipedia: writing style matters. In *WWW '10*, 2010, pages 1147–1148.

[8] A. Oulasvirta, J. P. Hukkinen, and B. Schwartz. When more is less: the paradox of choice in search engine use. In *SIGIR '09*, pages 516–523, 2009.

[9] J. Packer. Learning for fun: The unique contribution of educational leisure experiences. *Curator: The Museum Journal*, 49(3):329–344, 2006.

[10] J. Packer. Beyond learning: Exploring visitors' perceptions of the value and benefits of museum experiences. *Curator: The Museum Journal*, 51(1):33–54, 2008.

[11] J. Packer and R. Ballantyne. Motivational factors and the visitor experience: A comparison of three sites. *Curator: The Museum Journal*, 45(3):183–198, 2002.

[12] J. M. Packer. *Motivational factors and the experience of learning in educational leisure settings*. PhD thesis, Queensland University of Technology, 2004.

[13] G. Wu, M. Harrigan, and P. Cunningham. Characterizing wikipedia pages using edit network motif profiles. In *SMUC '11*, pages 45–52, 2011.

A Diary Study of Information Needs Produced in Casual-Leisure Reading Situations

Max L. Wilson
Future Interaction Technology Lab
Swansea University, UK
m.l.wilson@swansea.ac.uk

Basmah Alhodaithi
Future Interaction Technology Lab
Swansea University, UK
basmah.alhodaithi@gmail.com

Michael Hurst
Department of Information Science
Loughborough University, UK
m.a.hurst@lboro.ac.uk

ABSTRACT
Both information seeking and leisurely activities are commonplace in people's daily lives, but very little is know about searching behaviours outside of the work context. To study such leisurely information needs and subsequent searching, a diary study was performed, focusing on the context of casual-leisure reading. The week-long diary study with 24 participants was performed by a team of six graduate students. Reading was often both an act of casual searching, as well as a motivator for subsequent searching episodes, and around half were hedonistically or emotionally motivated. Casual searching often began with topical or personal interests, but did not always involve information needs. The findings confirm prior literature on casual search, while providing new insights into these less-critical and experience-driven episodes of searching, for fun.

General Terms
Experimentation, Human Factors, Theory.

Keywords
Casual-leisure, Reading, Information Seeking

1. INTRODUCTION
Although there has been decades of research into Information Seeking and Information Retrieval, very little has focused on the casual searching experiences of people outside of work. Research by Harris and Dewdney in 1994 indicated that 95% of 3,100 surveyed information seeking studies had focused on work-driven tasks [8]. Yet Pew Research found that searching simply for fun, and often for no particular reason, is one of the most popular online pastimes and counts for a significant portion of internet traffic [17]. Elsweiler et al suggest that casual, leisurely searching situations differ significantly to work or project driven tasks in that they produce search experiences that often begin without a given information need. Further, their investigations indicated that actually finding relevant information is typically less important than having fun [5]. Such scenarios involve passing time and relaxing, can be driven by the need to recover from a bad day, or to have fun with other people. Casual searching includes scenarios such as window shopping, browsing eBay, and delving into Wikipedia. To further investigate such casual-leisure searching experiences in more detail, this paper describes a diary study of searching for fun, performed in the context of casual reading.

2. RELATED WORK
The study of searching behaviour has long been embedded in the history of library and information science, where searching is presumed to be a goal-oriented research activity. This is highlighted by the common definition that Information Seeking is focused on the resolution of an *information need* [12] or *knowledge gap* [1]. Further, the common approach to describing tasks for empirical research, is named a 'Work Task' [2]. Despite implying work-oriented scenarios, Work Tasks are described as including non-work personal tasks too, but these tasks are still typically goal and need-driven scenarios. Examples include studies of everyday-life information seeking [18] and information encountering [6], which relate to non-work contexts, but can still be quite serious.

To understand non-work leisure time better, Stebbins introduced a taxonomy containing three levels: serious-leisure, project-leisure, and casual-leisure [22]. Serious leisure typically covers activities relating to committed hobbies, or volunteering outside of work [9]. Project-leisure relates to extended but temporal efforts like buying a car, planning a holiday, or researching family histories [3]. These goal- and need-driven leisure scenarios could be easily captured in Work Tasks. The third level, casual-leisure, relates to activities often involved in play and relaxation, such as watching television [4] or searching online [23], and much more. Based on their prior work, Elsweiler et al proposed a model of casual-leisure information behaviour [5] that highlighted some key differences between casual scenarios and Work Tasks. First, these scenarios were often driven by hedonistic needs, rather than information needs. Consequently, searching often began with ephemeral or absent information needs. Further, success in meeting their hedonistic needs, did not necessarily involve successfully finding information and results. Hedonistic needs include factors such as affect, novelty, social relationships, and enjoyment [10], where O'Brien, for example, studied their importance in online shopping experiences [14].

Many have also studied reading as a casual or pleasurable activity. Early work by Pjetersen converted observed book-finding behaviour into a naturalistic library-style search interface [16], helping people to browse in different modes. In 1980, Spiller found that 46% of library loans (n=500) were based upon browsing and 54% on known authors [21]. During a much smaller (n=12) qualitative study in 2011, however, Ooi and Liew saw participants often only using the library to retrieve books that they had already selected in everyday life [15]. Further, along with the introduction of e-readers and tablet devices, the nature of reading in casual episodes is changing. Research continues to highlight that increasing numbers of people perform their reading online or through digital mediums [11, 20].

3. DIARY STUDY

The main goal of this study was to investigate the information seeking behaviours performed in the context of casual-leisure reading. Prior work by Ross found that people who read for pleasure often encounter new information, without having an existing related information need [19]. Here, six researchers, as part of their post-graduate studies, coordinated a diary study of casual-leisure information behaviour. The methodology used was similar to the diary study performed by Elsweiler et al [4], which studied information needs produced while watching television. In total 24 participants took part in the diary study for one week. Participants were recruited by the six researchers using snowball-sampling; participants were primarily young adults in their 20s.

Participants were given a small portable physical diary, so that it could be used in both digital and physical contexts; an example is shown in Figure 1. Participants were asked to fill out one entry page per information need or searching episode that was initiated during a period of reading undertaken for self-motivated pleasurable reasons. To support continued participation, the participants were managed by one of the six researchers. Each participant had regular contact with their researcher, including but not limited to: an initial interview, an informal interim discussion, and a final debriefing interview.

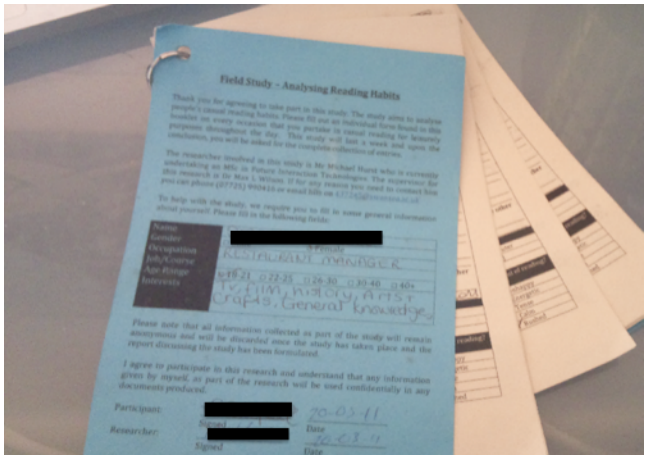


Figure 1: An example diary; a bound set of A5 card.

The diary consisted of a mix of open and closed questions. After logging the time and date, participants were asked to indicate the type of material they were reading and their environment, such as home, work, library, coffee shop, etc. Participants were then asked to describe a) what they wanted to search for, and b) why they wanted to search. Participants were then asked to identify how they then performed the search, if at all.

3.1 Analysis

Although some summative information was collected about the nature of the reading scenario, a Grounded Theory analysis [7] was performed to systematically extract key elements from the information needs and information seeking described in the open text fields. The six researchers individually transcribed their diaries and initially coded them for key points. As a group, and in collaboration with the supervising author, these codes were discussed, analysed, and configured into affinity diagrams, using post-it notes and a whiteboard. These codes, and the relationships captured in the affinity diagrams, were discussed, referring back to example diary entries, until they stabilized and all researchers were in agreement. Entries that challenged the evolving definitions and affinity diagrams were frequently considered during this process. The six researchers then returned to their diary entries to re-examine them in the context of the final codes.

4. RESULTS

Over the course of the week, most participants recorded around 1 or 2 diary entries per day, producing around 120 usable entries in total. To provide an overview, approximately 20% of reading was performed with physical paper objects (books, newspapers, and magazines), with the remaining being split between e-readers and mobile devices (around 30%) and laptops and PCs (50%). Reading content included: News (around 45%), email (20%), magazines (15%), and fiction (10%). In terms of physical surroundings, around 40% of entries were produced in work contexts, with the remaining performed in home environments. Figure 2 shows the model developed from the analysis, which is described further below.

1. *Reading Motivations*
    a. *Hedonistic or Emotional*
    b. *General knowledge interests*
        i. *Interest driven*
        ii. *Carer*
        iii. *In-the-know*
        iv. *Decision*
2. *Searching Motivations*
    a. *Information need*
    b. *Personal scoping*
    c. *General topical*
    d. *Decision-making*
3. *Search focus*
    a. *Factual information*
    b. *Background information*
    c. *Object related information*
4. *Source of Information*
    a. *Paper sources*
    b. *Social networks*
    c. *Expert sites*
    d. *Generic sites*

Figure 2: The developed coding scheme.

4.1 Reading motivations

Reading material can be considered a source of information itself. Consequently, our study observed reading as being both the act of casual searching, and as a source motivating separate casual search episodes. This section focuses on the former, where casual reading is itself sometimes an act of casual search.

Although around 50% of casual reading episodes were driven by hedonistic or emotional needs, around 50% were driven by the participants' general knowledge interests. Examples of hedonistic or emotional motivations included "to pass time", "to help cope with things", and "to relax after my day". Although following knowledge interests could also be seen as a pleasurable pastime, the knowledge-driven entries also occasionally broached the concepts of 'project leisure', such as reading about possible holiday destinations, and 'serious leisure', such as reading around a hobby domain. The majority of the knowledge-drive situations described by participants, however, were casual episodes relating to a project-leisure interest, rather than active periods of research or work. One participant, for example, was reading about a neighbourhood area as they were soon to be "moving into a new house".

While the hedonistic and emotional scenarios were pretty uniform in motivation, we further classified the casual knowledge-driven reading scenarios into four types: Interest driven, Carer, In-the-know, and Decision-oriented. Interest driven were those casual bouts of reading relating to a hobby or temporary interest.

Examples included "information about buying a car abroad" and "information on fixing my PC". For a participant who was a "new fan of J.K. Rowling's novel series", they were "reading about the latest Harry potter sequel", which was due to be delivered.

Carers were those that were reading information that has personal or emotional relevance. Carers often read news, for example, about zones with natural disasters, or places and events relating to their childhood, or to distant friends. One participant cited choosing to read "more information on tsunamis", while another had a personal interested in the unrest in the Bahrain.

In-the-know readers were those that casually monitored general knowledge information sources, including news, to be aware of current events and new technology. Example diary entries included a participant who "read about the 2011 budget meeting in today's paper" in order to get "updates on current budget meetings". Another participant said "I wanted to know what was happening while I was asleep". In-the-know readers often recorded more frequent small reading sessions, than extended periods like those with hedonistic or emotional motivations.

Finally, decision makers were those that read up on interest areas related to things like casual purchases, such as new movie releases or new cameras. In another example, a participant wrote that they were reading "reviews of the movie 'Inception'", because they were "planning for a movie at the weekend".

4.2 Motivations for Searching for fun

The casual reading, recorded in our diary study, often created separate episodes of casual searching. These episodes were driven by encountering information that created an Anomalous State of Knowledge [1], but did not always relate to a direct information need. Some ASKs also led to additional smaller bouts of casual-interest reading, rather than searching. The four identified key motivations for additional searching or reading, were: information need, personal scoping, general topical, and decision-making.

Information need examples included those that identified a clear piece of information they would like to know in order to continue reading. These specific information needs often consisted of dictionary definitions, such as one participant who was looking for "the meaning of the word 'oakum'" because they did not know what it meant.

Personal scoping motivations related to participants who encountered information that was somehow related to their history or personal life. The participant interested in the Bahrain also provides a good example here. Personal scoping examples also often led to searching behaviour within one's own information, such as email or media collections, or within social networks. Typically, personal scoping was aimed at establishing, or remembering, the connection they had with the information they had just encountered.

General topical searching was motivated by discovering something of novel interest, and often initiated casual learning without a specific information need. One participant, another example of a Carer, wanted to "know more about children with dementia" after they "read [an] article in [the] newspaper about a 9yr girl with this disease".

Finally, decision-makers were those searching when motivated by the need to make a new decision. Often relating to a topical interest, such decision-making motivations included deciding if an activity was something they would want to do, or to learn more about in future casual reading. One participant said that they wanted to "check the weather for the weekend" in order to make some plans.

4.3 Focus of information sought

The information that people sought in these casual scenarios could be largely broken into three types: factual information, background/overview information, and object related information. Factual information, of course, related to specific information needs, and were often represented by factual content, such as dates, prices, locations, etc. One participant was searching for "yesterday's lottery results". Background and overview information was typically sought in general topical situations and interest-driven reading, such as "wales football information". Finally, object related information pertained to places, people, and events with one participant suggesting they were "searching for more about Mississippi". Such information was often sought by caring readers, or personal-scoping searchers.

4.4 Sources of information

The diary study also asked participants to describe how they sought information during episodes of casual searching, motivated by their casual reading. Perhaps correlating with the large percentage of our participants who read using digital devices, much of the information was sought online. Figure 3 highlights that some participants sought their information using additional physical paper resources, often including those who performed additional topical interest reading. Of those that used the internet to search, many consulted their social network, especially those establishing personal scope with the information. The remainder typically referred to news sources and Wikipedia articles, or generally searching the web for related pages. Several participants described themselves as searching for websites with authority on a topic, such as one participant who went to the UK government website for "…census information. To find out the deadlines".

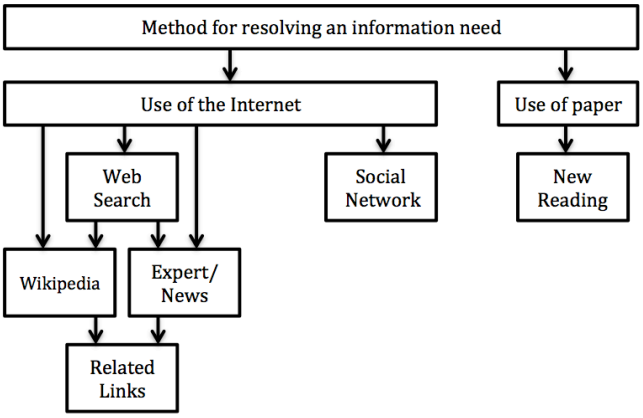


Figure 3: Methods used for casual searching.

5. DISCUSSION

This research has continued the recent interest in investigating casual searching behaviour that people undertake for fun. We aimed to further investigate the findings of researchers like Elsweiler et al [5], and the model of casual-leisure searching behaviour they produced. In line with their model, our study found that around half of the casual reading episodes were motivated by hedonistic or emotional needs, rather than information needs. For those that engaged in searching behaviour, some did aim to find specific information, either facts or information connecting what they had found to their own lives, while others began additional reading or topical browsing without

a given information need. This finding, however, highlights that although Elsweiler et al's model separated information and hedonistically driven motivations, these episodes are often intertwined and highly connected. Further, our work contributed additional insights into variables created by person- and situation-types, both of which have an affect on the interplay between informational and emotional motivations. While these findings are novel, future work should focus on fully understanding these conditions; some notions, for example, are closely related to elements of McQuails Mass Communication Theory [13].

Unfortunately, the design of the study meant that we did not capture information about whether people succeeded in finding information. Future work could help to validate these latter phases of Elsweiler et al's model, by focusing on the success, failure, and importance of casual searches.

5.1 Limitations

Although the study covered 24 participants over the space of a week, and gathered over 120 casual searching episodes, there are some potential limitations in the methodology that should be acknowledged. First and foremost, the study was performed by five masters and one PhD student, each in the first few months of their postgraduate study. Consequently, this was their first field study and they were learning the techniques by performing them; their individual skills varied. Further, each researcher produced their own paper diaries, which also introduced some slight variations in content. Despite the fact that execution of the study may have been less rigorous than many diary studies, the results did reveal several findings that both confirmed elements of other research and revealed new insights into casual-leisure searching.

6. CONCLUSIONS

This paper has described a diary study that investigated searching for fun, in the context of casual reading. Research has shown that such activities make up a significant portion of internet traffic, while remaining largely under-studied. Our findings provided further evidence for previously proposed models of casual searching, including the significance of hedonistic and emotional, rather than information-driven, motivations. Further, we have shown that many of these activities relate to areas of interest and personal scope, rather than being specifically related to an information need. Finally, much of the casual leisure searching was for decision-making, but in regards to pleasurable hedonistic activities and purchases. Combined with previous research in this area, our findings contribute to the developing understanding of these less-critical, experience-driven, often-hedonistic episodes of searching, for fun.

7. ACKNOWLEDGMENTS

Thanks both to the participants, and the remaining researchers who helped to run the study: Tashi Rapten Bhutia, Mohammed Taheri, Daniel Williams, and Tim Crawford. Also thank you to the reviewers for their valuable comments.

8. REFERENCES

[1] Belkin, N.J., Oddy, R.N. and Brooks, H.M., Ask for information retrieval: parts I and II. *Journal of Documentation*, *38*(2/3), 61-71, 145-164. 1982.

[2] Borlund, P., Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, *56*(1), 71-90. 2000.

[3] Butterworth, R., Information seeking and retrieval as a leisure activity. In *DL-CUBA'06*, 29-32. 2006

[4] Elsweiler, D., Mandl, S. and Lunn, B.K., Understanding casual-leisure information needs: a diary study in the context of television viewing. In *IIiX'10*, 25-34. 2010

[5] Elsweiler, D., Wilson, M.L. and Kirkegaard Lunn, B. Understanding Casual-leisure Information Behaviour. in Spink, A. ed. *Future Directions in Information Behaviour*, Springer, 2011 (to appear).

[6] Erdelez, S., Information encountering: a conceptual framework for accidental information discovery. In *ISIC'97*, 412-421. 1997

[7] Glaser, B.G. and Strauss, A.L., The discovery of grounded theory. *The British Journal of Sociology*, *20*(2), 1967.

[8] Harris, R.M. and Dewdney, P. Barriers to information: How formal help systems fail battered women. Greenwood Press Westport, CT, 1994.

[9] Hartel, J., The serious leisure frontier in library and information science: hobby domains. *Knowledge organization*, *30*(3-4), 228-238. 2003.

[10] Hassenzahl, M., The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, *13*(4), 481-499. 2001.

[11] Liu, Z., Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation*, *61*(6), 700-712. 2005.

[12] Marchionini, G. Information Seeking in Electronic Environments. Cambridge University Press, 1995.

[13] McQuail, D. McQuail's mass communication theory. Sage Publications Ltd, 2000.

[14] O'Brien, H.L., The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers*, *22*(5), 344-352. 2010.

[15] Ooi, K. and Liew, C.L., Selecting fiction as part of everyday life information seeking. *Journal of Documentation*, *67*(5), 748-772. 2011.

[16] Pejtersen, A.M. The Book House: Modelling User'Needs and Search Strategies a Basis for System Design. Ris√∏ National Laboratory, 1989.

[17] Purcell, K. Search and email still top the list of most popular online activities *Pew Internet & American Life Project*.2011.

[18] Savolainen, R., Everyday Life Information Seeking: approaching information seeking in the context of. *Library & Information Science Research*, *17*(3), 259-294. 1995.

[19] Sheldrick Ross, C., Finding without seeking: The information encounter in the context of reading for pleasure. *Information Processing & Management*, *35*(6), 783-799. 1999.

[20] Smith, R. and Young, N.J., Giving Pleasure Its Due: Collection Promotion and Readers' Advisory in Academic Libraries. *The Journal of Academic Librarianship*, *34*(6), 520-526. 2008.

[21] Spiller, D., The provision of fiction for public libraries. *Journal of Librarianship and Information Science*, *12*(4), 238. 1980.

[22] Stebbins, R.A., Leisure and Its Relationship to Library and: Information Science: Bridging the Gap. *Library trends*, *57*(4), 618-631. 2009.

[23] Wilson, M.L. and Elsweiler, D., Casual-leisure Searching: the Exploratory Search scenarios that break our current models. In *HCIR'10*, 28-31. 2010

In Search of a Good Novel Examining Results Matter

Suvi Oksanen
School of Information Sciences
University of Tampere
33014 University of Tampere, Finland
Suvi.Oksanen@uta.fi

Pertti Vakkari
School of Information Sciences
University of Tampere
33014 University of Tampere, Finland
Pertti.Vakkari@uta.fi

ABSTRACT
We studied how an enriched public library catalogue is used to access novels. 58 users searched for interesting novels to read in a simulated situation where they had only a vague idea of what they would like to read. Data consist of search logs, pre and post search questionnaires and observations. Results show, that investing effort on examining results improves search success, i.e. finding interesting novels, whereas effort in querying has no bearing on it. In designing systems for fiction retrieval, enriching result presentation with detailed book information would benefit users.

Categories and Subject Descriptors
H.3.7. [**Digital Libraries**]: User Issues

General Terms
Human Factors

Keywords
Fiction Retrieval, Novels, Readers, Public Libraries, Search Tactics, Search Effort, Search Success

1. INTRODUCTION
Reading novels is a popular leisure time interest. Fiction was read at least once a year by 50 % of Americans in 2008 [10] and by 80 % of Finns in 2010 [13]. Public libraries are major channels of getting access to novels [9]. Studies on the outcomes of public libraries show that the major benefit derived from their use is the pleasure of reading fiction [6, 15]. Despite this fact, there has not been much interest in studying and developing systems for fiction retrieval since the 1980s [1]. The effort in developing search systems has been focused on retrieving non-fiction [2, 4].

Traditionally library catalogs have supported accessing novels if the reader knows the name of the author or the title of the novel. It is know that about half of the fiction borrowed is found by browsing, half by known item search [14]. This indicates a need to develop systems supporting other fiction search tactics than known item search. There are signs of enriching public library catalogs to include features supporting fiction retrieval like extended book descriptions or indexing [1, 12]. However, the utility of these tools for accessing novels is not studied. Our aim is to analyze how tools provided by an enriched public library catalogue are used to access interesting novels to read.

Presented at Searching4Fun workshop at ECIR2012, Barcelona, Spain.

2. RELATED RESEARCH
Next we introduce studies on how readers access fiction in libraries and on evaluation of fiction search systems. The literature in this field is scarce [1]. In [8], Pejtersen summarizes her seminal works in fiction retrieval. As far as we know, there have been no published studies on fiction searching in commercial sites like Amazon. The discussion in [3] hints also to that.

Goodall [5] differentiates two stages in the book search process in the library. Readers identify first attributes in the books, which trigger their interest, and after that focus on attributes, which generate the decision to borrow the book. In the filtering stage, external attributes of the book like its cover or title are perceived as important, whereas in the selection stage, internal attributes of the book like text on the back of the cover or passages of the text in the book are considered as useful. Ross [11] has made a roughly similar distinction based on interviewing 194 committed readers. She distinguished between the clues in the book and elements in the book as indicators of an interesting book.

Pejtersen [8] has defined three major tactics for accessing fiction, which match to our research goals. Analytical search strategy is used when readers wish to find novels about some topic like the Second World War. Search by analogy is generated when readers want something similar to novel X, e.g. a novel they had previously read. Browsing strategy is applied in situations when readers have only a vague idea of what they would like to read. They are simply browsing for finding a good novel.

Based on observing user-librarian negotiations for finding fiction, Pejtersen [8] has designed a fiction search system called the Book House. It consisted of facets representing various attributes of novels as perceived by library users. These facets were access points to novels. The evaluation showed that the system was useful and pleasurable to use [8] All the available system functionalities were used and the fiction classification system fully accepted. The users found it useful in finding novels.

3. RESEARCH DESIGN
The aim of this study is to analyze how an online catalog in a public library is used for finding novels to read. We focused on a situation when the readers have only a vague idea of what they would like to read. This corresponds to the browsing strategy in Pejtersen [8]. In addition to known item search, browsing is the second major strategy for accessing fiction [8, 14]. Conceptually, browsing includes also similarity search and category search, because in these search modes the reader does not know exactly

what she wants. Browsing may lead to similarity search and category search. Therefore, we chose browsing as the search mode in our study. The specific research questions are:

- What kind of search moves were used for accessing novels?
- Was there an association between moves and search success?

PIKI library system serves several municipalities in Tampere region in Finland. It includes a database containing metadata about the books in the networked libraries, and an interface to interact with that information and search books. The metadata for fiction contains typical bibliographic information added with keywords from the fiction thesaurus "Kaunokki" [12] and tags assigned by users. The metadata includes also images of book covers, recommendations by users and librarians, and availability information. The object of a default search is the whole database. Search results are ranked by relevance, but they can be ordered also alphabetically by author or title, and by publication year. Search results can be limited by category, i.e. fiction vs. non-fiction, by the type of material like book, video, etc., by keyword, by language, or by library. Clicking the book title on the result list reveals the metadata of the book with availability information.

In addition to author, title, free term or keyword search, users may start from recommendation pages. They include various lists of books and recommendations by users and librarians. Users can also search for similar books based on keywords.

For the study 58 participants were recruited in May 2011 from three public libraries of various sizes in PIKI area. Of the study subjects, 26 were recruited in a big main library, 22 in a medium sized main library and 10 in a small branch library. 36 were females and 22 males. Their age varied between 14 and 70 years, the average age being 34 years. They were relatively highly educated, 39 % had a university degree, and 23 % had a high school education, and the rest had a lower education. They read on average 24 novels per year ranging from 0 to 120 novels.

The search task was as follows: You are in a library in a situation when you do not have a clear idea of what you would like to read. Please use the PIKI catalog to search for a novel of interest to you, which you would like to read. Do not search for a particular author or novel, although you may use this as a point of departure for your search. Thus, we simulated a typical browsing situation [5, 11] when readers have only a vague idea of what they would like to read [8]. The search was ended when an interesting novel was found, or when the searcher gave up the search task as unsuccessful.

The search screen was recoded. The researcher observed the search sessions and made notes. The searchers filled in a pre-search questionnaire eliciting demographic information, information about reading orientation, the use of the library and search tactics for books in the library. After the search they filled in a post-search questionnaire including a pattern of questions for assessing various features of PIKI interface, ranking of the novel found and open questions concerning the criteria of selecting the novel and the difficulty of the search task.

Search moves were observed from the recordings of search screen. 29 move types were identified. A move is an identified use of a system feature like a keyword search, an author search, inspecting result list, limiting it, or exploring book metadata. The number of the moves varied from 2 to 21. The distribution of moves was very scattered. The four most common moves were book clicks (20.4 %), result list (20.2 %), free text search (8.2 %) and category limitation (6.5 %). The proportion of all other 25 moves varied between 4.8 % and 0.2 %. Therefore, for the economy of analysis we collapsed similar move categories like field search (by publication date, library, language, category, material) or limiting result list (by keyword, language, etc). We also recorded the time used for the search.

The indicator of the success of search was an interesting novel found. The searchers rated the novel in a three-point scale from one to three (least to most interesting). If the searcher could not find an interesting novel, the scoring was zero.

4. RESULTS

When starting a search, readers could select either a quick search, an advanced search or a recommendation page as their point of departure. Quick search consists of a search box with a drop down menu suggesting a keyword with information about its type like author when keying in search terms. In an advanced search it is possible to formulate a query by selecting several fields to search. Recommendation pages include various lists of books and recommendations with links.

Advanced search was the most popular search mode (72.4 %) followed by quick search (19 %) and recommendations (17.5 %) (table 1). Readers made on average 7.9 moves when attempting to find a good novel. Of these moves on average 3 were advanced searches, 0.4 quick searches and 0.5 recommendation moves. Users retrieved on average 1.6 result lists, and limited these result lists 0.6 times. On the result lists they clicked 1.6 books, but read only 0.2 book descriptions containing more than bibliographic data. The average interest score of the book accepted was 2.4. The average search time was 215 seconds.

Table 1. Basic statistics of the main study variables (n=58)

Variable	Mean	Stddev	Min	Max	% using
Quick search	0.4	1.1	0	6	19.0
Advanced search	3.0	2.9	0	12	72.4
Result list	1.6	1.4	0	6	86.2
Result list limit.	0.6	1.3	0	6	27.6
Book clicks	1.6	1.3	0	7	95.1
Book description	0.2	0.6	0	3	10.3
Recommendation	0.5	1.3	0	7	17.5
All moves	7.9	4.3	2	21	100
Book scores	2.4	0.9	0	3	100
Search time	215	118	76	593	100

As table 2 indicates, the most popular search tactic was field search (63.8 %) followed by free term search (44.8 %). Known item search and keyword search were equally popular.

An average search was relatively short consisting of about eight moves and lasting about 3.5 minutes. A typical search consisted of advanced searches including mostly field searches or searches with terms from controlled or free text vocabulary. Searchers seldom limited the result list, but immediately assessed novels by

examining bibliographic book information. They explored very seldom more detailed book descriptions for assessing novels' value. The searches can be considered as successful. Only five searchers out of 55 could not find a novel, which they considered as interesting. Evaluation scores in three cases were missing. Thus, 50 searchers had a successful result, i.e. a novel rated at least with value one. Of the searchers only one rated the novel with value one, nineteen with value two, and the rest thirty with value three. Thus, about 55 % of the searchers retrieved a novel with the highest interest rank.

Table 2. Basic statistics of the search tactics variables (n=58)

Search Variable	Mean	Stddev	Min	Max	% using
Known item	0.6	1.1	0	5	32.8
Free term	0.8	1.3	0	6	44.8
Keyword	05	1.0	0	5	32.8
Field search	1.4	1.4	0	5	63.8

We were curious to know whether the search process variables were associated to the success of search measured by the interest rate of the novel found. We analyzed the association between search moves and search success by calculating Pearson correlation coefficients. The results indicated that none of the search process variables in tables 1 and 2 excluding the result list was significantly associated with the perceived value of the novel. The number of result lists visited correlated significantly with the success (r=28; p=04). Thus, it seems that search success was not associated with the search moves or their combinations used excluding the number of visits in the result list.

Success was neither associated with search effort measured as time used in searching (r=-.14; p=.31) or the total number of moves (r=.23; p=.10). However, we observed that effort invested in exploring the search results and in querying were significantly associated with the search success. Correlation between the time invested on an average move and the interest rating of a novel found was -.45 (p=001) (table 3). Thus, quick shifts from move to move predict finding an interesting novel. The correlations show also that the greater proportion of the moves devoted to looking at the result list (r=.34; p=013) or examining novels in detail found on the result list (r=.31; p=.022), the more likely searchers found an interesting novel. Deviating from this finding, the proportion of quick and advanced searches of all moves was negatively associated with the ratings of the novels selected (r=-.27; p=.045). Thus, the greater the proportion of quick or advanced search moves of all moves, the less interesting novels were found.

In all, these findings hint, that search formulation variables, i.e. querying, were not associated with finding an interesting novel to read, and their great proportion of all moves contributed to an unsuccessful search result. The proportion of moves devoted to exploring result lists and book information, however, helped searchers to find interesting novels. Thus, the more swiftly the searchers proceeded from move to move, but the more effort they invested in exploring results list and book information, and the less effort in search formulation moves, the more interesting novels they found. The findings imply, that search formulations are less important than examination of search results as conditions for finding an interesting novel to read.

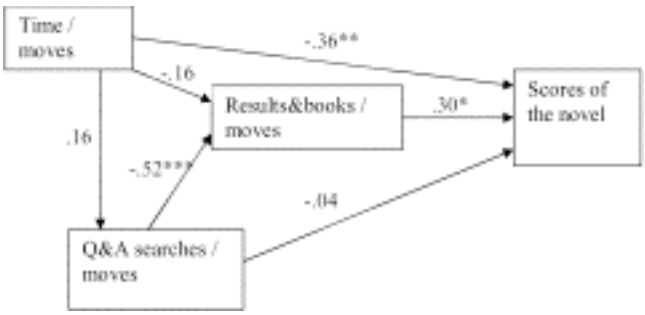
Table 3. Correlations between the average time per move, search effort and the interest grade of a novel (n=58)

Variables	Book scores	Time/ moves	Results/ moves	Results, book/mo
Time/moves	-.45**			
Results/mo	.34*	-.19		
Results, book/moves	.31*	-.24	.70***	
Q&A searches/mo	-.27*	.16	-.03	-.54***

Legend: *= p<.05; **=p<.01; ***=p<.001

The previous correlation analyses suggest that the following variables were significantly associated with search success: the average time per move, result lists per move, results and book information per move, quick and advanced searches per move. We use these variables for predicting search success, i.e. the rating of the novel found. Because the two variables measuring the proportion of result list exploration of all moves were conceptually correlated, we removed the variable measuring only visits in result lists, and kept that one which included also exploring book information. The latter one reflects more validly the effort put in exploring the search results.

The model building aims at analyzing the direct and intermediated effects of each independent variable to dependent variable. The model indicates the relative effect of each variable to other variables, i.e. it indicates the effects other variables controlled [7]. Path analysis was used for testing the model. In the path analysis standard regression coefficient are used [7]. The model (figure 1) was significant (F=7.14; p=000) indicating a good fit with the data. The multiple correlation (R) of the model was .548, and adjusted R squared .258. Thus, the model explains about 26 % of the variance in the scores of the novels.



Legend: * = p<.05; ** = p<.01; ***=p<.001 (n=58)

Figure 1. A path model for predicting the scores of the novel

The path analysis indicates that time used per move has a significant direct effect on the scores of the novel found (beta=-.36). Also the proportion of search result exploration of all moves has a significant effect on novel scores (beta=.30), whereas the proportion of quick and advanced searches of all moves has no effect on the interest rating of the novel (beta=-.04). The average time per move has a significant effect neither on the proportion of results exploration (beta=.-.16) nor on quick and advanced searches of all searches (beta=.16). Interestingly, the proportion of quick and advanced searches has a very large significant effect on the variation in the proportion of result exploration (beta=-.52).

In all, the model indicates, that the less time the searcher used per move, the more interesting novels were found. The average time used per move did not have a significant influence on the proportion of moves devoted either to search formulation or the results exploration. Although these beta coefficients were not significant, their directions hint, that the less time used per move, the more effort was invested in examining the result lists and books information, and the less effort in search formulations. In addition, the more effort put in querying, the less effort allocated in examining results. Thus, it seems that there was a bifurcation of search strategies emphasizing either querying or result list examination. These two strategies had very different effects on finding an interesting novel. Investing effort on examining the result list and book information has a significant positive effect on finding an interesting novel, whereas emphasis on search formulations has no bearing on finding an interesting novel.

5. DISCUSSION AND CONCLUSIONS

As far we know, this is the first study since Pejtersen [8] to analyze the search tactics used by readers for accessing fiction in enriched public library catalogs. We observed how readers searched for an interesting novel in a situation where they had only a vague idea of what they would like to read [8]. We found out that the use of various moves for searching novels was scattered. The most common moves were advanced search, browsing result list and examining book information. The use of various moves was not associated with the success of the search, with finding an interesting novel. However, it turned out that the less time used per move, and the greater the proportion of moves for examining the result list and book information, the more interesting the novel found. The proportion of search formulation moves was not associated to the search success. The model build hints that readers used two alternative strategies with differing success for accessing good novels. The strategy emphasizing search formulations was not associated with finding an interesting novel, whereas the more effort invested in examining results in the search, the more interesting novel was found.

Effort invested in exploring search results instead of querying is an essential factor for finding interesting novels in a situation when readers do not have a clear idea of what they wish to read. Although readers have only a vague idea of the object of interest, they know genres, authors and titles, and have attributes of good novels in their mind [11]. They use this information when selecting books to read. It is likely that what is considered as an interesting novel varies a lot in the sense that the substitutability of novels is great in this situation. Several alternatives may do, not only one. Therefore, effort put on exploring the result list is more productive than querying in the search for good novels to read.

Our results suggest that in designing systems for fiction retrieval, it is important to enrich result list presentation. Readers need more clues about where to infer that the novel could be of interest, and also more options to be informed about the content of the novel [5, 11]. The latter include e.g. recommendations by fellow readers and librarians, texts on the back of the books and links to critics of the novels and to author information like in some electronic bookshops.

It can be supposed that the more readers know about literature, the more effectively they can identify interesting fiction [11]. In the studies to come, we analyze whether readers' literary competence is connected to fiction search process and output. Also experimental studies on evaluating new tools for supporting fiction retrieval are needed.

6. REFERENCES

[1] Adkins, D. & Bossaller, J. E. 2007. Fiction access points across computer-mediated book information sources: a comparison of online bookstores, reader advisory databases, and public library catalogs. *Library & Information Science Research* 29(3), 354-368.

[2] Case, D. 2007. *Looking for information. A survey of research on information needs, seeking and behavior*. 2nd ed. Academic Press.

[3] Buchanan, G. & McKay, D. 2011. In the bookshop: examining popular search strategies. In *Proceedings of the 11th Joint Conference on Digital Libraries*. New York: ACM Press, 269-278.

[4] Elsweiler, D., Wilson, M.L. & Lunn, B.K. 2011. Understanding casual leisure information behavior. In: Spink, A & Heinström, J.. (Eds). *Future directions in information behaviour*. Emerald, 211-241.

[5] Goodall, D. 1989. *Browsing in the public libraries*. LISU Occasional paper No 1. Loughborough: Library and Information Statistics Unit.

[6] Lance, K.C., Steffen, N.O., Logan, R., Rodney, M.J., & Kaller, S. 2001. *Counting on results: new tools for outcome-based evaluation of public libraries*. Aurora, CO: Bibliographical Center for Research. Retrieved in December 20, 2009 from http://www. lrs. org/documents/cor/CoRFin. Pdf.

[7] Pedhazur E. 1982. *Multiple regression in behavioral research*. 2nd ed. New York: Holt, Rinehart & Winston.

[8] Pejtersen, A.M. 1989. *The Book House. Modeling user's needs and search strategies as a basis for systems design*. Roskilde: Risö National Laboratory.

[9] *Perceptions of Libraries, 2010: Context and Community*. (2010). OCLC. Retrieved August 10, 2011 from http://www.oclc.org/reports/2010perceptions.htm.

[10] *Reading on the Rise* (2008). National Endowment for the Arts. Retrieved November 27, 2011 from www.nea.gov/research/Readingonrise.pdf

[11] Ross, C.S. 2001. Making choices: What readers say about choosing books to read for pleasure. *The Acquisition Librarian* 13(25): 5-21.

[12] Saarti, J. & Hypén, K. 2010. From thesaurus to ontology: the development of the *Kaunokki* Finnish fiction thesaurus. *The Indexer* 28(2): 50-58.

[13] Serola, S. & Vakkari, P. 2011. *The public library in the activities of people*. In Finnish. Publications of the Ministry of Culture and Education 2011/21. MCE: Helsinki.

[14] Spiller, D. 1980. The provision of fiction for public libraries. *Journal of Librarianship* 12(4): 238-266.

[15] Vakkari P. & Serola, S. 2012. Perceived outcomes of public libraries. *Library & Information Science Research* 34(1): 37-44. doi:10.1016/j.lisr.2011.07.005