SIGIR 2009 Workshop on Understanding the User – Logging and interpreting user interactions in information search and retrieval

Georg Buscher
DFKI GmbH
georg.buscher@dfki.de

Jacek Gwizdka
Rutgers University
jacekg@rutgers.edu

Jaime Teevan
Microsoft Research
teevan@microsoft.com

Nicholas J. Belkin
Rutgers University
belkin@rutgers.edu

Ralf Bierig
Rutgers University
bierig@rci.rutgers.edu

Ludger van Elst
DFKI GmbH
elst@dfki.uni-kl.de

Joemon Jose
Glasgow University
jj@dcs.gla.ac.uk

1 Introduction

Modern information search systems can benefit greatly from using additional information about the user and the user's behavior, and research in this area is active and growing. Feedback data based on direct interaction (e.g., clicks, scrolling, etc.) as well as on user profiles/preferences has been proven valuable for personalizing the search process, e.g., from how queries are understood to how relevance is assessed. New technology has made it inexpensive and easy to collect more feedback data and more different types of data (e.g., gaze, emotional, or biometric data).

The workshop "Understanding the User – Logging and interpreting user interactions in information search and retrieval" was held in conjunction with the 32nd Annual International ACM SIGIR Conference. It focused on discussing and identifying most promising research directions with respect to logging, interpreting, integrating, and using feedback data. The workshop aimed at bringing together researchers especially from the domains of IR and human-computer interaction interested in the collection, interpretation, and application of user behavior logging for search. Ultimately, one of the main goals was to arrange a commonly shared collection of user interaction logging tools based on a variety of feedback data sources as well as best practices for their usage.

2 Structure of the Workshop

Since one of the main goals of the workshop was to gather practical information and best practices about logging tools, it was structured in a way to foster collaboration and discussion among its participants. Therefore, it was less presentation intensive (it included only 4 oral paper presentations), but contained more collaboration-supporting elements: participant introductions, poster presentations, a panel discussion, and, most importantly, group discussions.

This was also reflected in the types of possible submissions: Experience papers (4 pages) should describe experiences with acquiring, logging, interpreting and/or use of using interaction data. Demos of applications or new technology could be presented. Position statements should focus on types of

user interaction data / their interpretation / their use.

Each of those papers and demo descriptions got reviews by two members of the program committee. The program committee also judged the interestingness of each paper with regard to oral presentation (e.g., suitability to spawn discussion). The final selection of the 4 papers for oral presentation was made also with respect to the diversity of topics and approaches they covered. The accepted demos and all remaining accepted papers were selected for poster presentation.

The program of the workshop also reflected the focus on collaboration: It started with an extended participant introduction session where each participant of the workshop was asked to shortly present his or her main research interests related to the workshop's topics. A poster and demo session followed, succeeded by oral presentations of the 4 selected papers. After each paper, there was limited time for focused questions. In that way, each participant got the chance to see all workshop submissions (either as posters or presentations) and to talk to the authors, after which a panel with 3 panelists was formed based on submitted position statements. Following the panel discussion, breakout groups were formed based on common research interests and practical issues collected during the participant introduction session. The workshop ended with a summary of the achieved results and next steps to take.

Overall, reflecting on the general structure, we believe that the workshop was very successful in generating lots of focused discussions among the participants. Especially during the participant introductions, the poster sessions, and the breakout group discussions, every single participant was actively involved and very engaged (rather than only those who submitted papers).

3 Participant Diversity

The workshop drew 34 registered attendees. They mostly came from academia, but there were strong contributions from industry as well.

The participant introductions in the beginning of the workshop revealed a positively surprising broad range of expertise, experiences, and interests. In Table 1, we give an overview of the range of scenarios focused on by the different attendees. Table 2 shows topics the participants were most interested in.

4 Paper, Poster and Demo Presentations

In this section, we group and briefly list the papers that have been accepted for the workshop. Overall, 11 experience papers and 4 demos were accepted which are arranged into 5 topical groups below. Four papers (one from 4 of the 5 groups) were selected for oral presentation. (More information about the program can be found on the Web: http://uiir-2009.dfki.de/index.php/program.) The workshop proceedings can be found in [1].

Logging tools / frameworks

- Oral presentation by Ralf Bierig, Jacek Gwizdka and Michael Cole: *A User-Centered Experiment and Logging Framework for Interactive Information Retrieval*. They presented a framework for multidimensional (interaction) data logging that can be used to conduct interactive IR experiments.
- Demo by Claus-Peter Klas and Matthias Hemmje. *Catching the User - User Context through Live Logging in DAFFODIL*. This demo presented an interactive IR experimentation framework that can be used to log events during a search session such as querying, browsing, storing, and modifying contents on several levels.
- Demo by Robert Capra. *HCI Browser: A Tool for Studying Web Search Behavior*. This demo showed a browser extension that contains the most important functionalities needed when

conducting a browser-based user study, such as logging browser-specific events and presenting questionnaires to the user before and after an experiment.

Table 1: Scenarios workshop participants focused on with respect to logging and using (implicit) user interaction data

Types of information interacted with	**Types of (implicit) interaction data**
• Information visualizations / search interfaces • Web text documents • Personal information (emails, files on desktop) • Notes/annotations in documents • Music • Images • Structured or semi-structured data (e.g., medical information) • Physical content (pictures, books)	• Queries • Clicks, URL visits o Identification of interaction patterns, e.g., repeat actions (repeat queries, repeat URL visits) • Notes/annotations • Changes made by author in document • Eye movements • Biometric feedback: EEG, galvanic skin response (GSR), facial expressions

Uses of implicit interaction data

- Modeling the user
  - Identification of domain knowledge / expertise
  - Better expression of interests
  - Emotion detection (frustration, stress)
  - Identification of good / bad experiences
- Personalization / contextualization
  - Improving relevance
  - Proactive information delivery
- Introspection / reflection (e.g., analyzing what makes a good searcher)
- Finding better ways to display retrieved information

Table 2: Topics of interest

Topics focused on in the above scenarios

- Tools for processing low-level logs (e.g., eye tracking, EEG, ...)
- Ways to combine implicit and explicit feedback data (frameworks)
- Ways (tools) to record context (current task, etc.)
- Sharing of logging tools and log data sets (collection of tools, data formats, etc.)
- Uses for implicit data:
  - Improving information experiences in the aggregate
  - Personalizing information experiences
  - Social sciences: Reflecting on people in the aggregate
  - Introspection: Reflecting on self or individual
- Validity of collected data (collected in the wilds vs. in a user study; dependence on used collection tools)
- Privacy issues

Demo by Stephen Dignum, Yunhyong Kim, Udo Kruschwitz, Dawei Song, Maria Fasli and Anne De Roeck. *Using Domain Models for Context-Rich User Logging*. The demo presented an interface where users can explore a domain using structured representations thereof. The authors propose using the explored paths of the domain model as contextual feedback.

Analyzing user behavior logs

- Oral Presentation by Robert Capra, Bill Kules, Matt Banta and Tito Sierra. *Faceted Search for Library Catalogs: Developing Grounded Tasks and Analyzing Eye-Tracking Data*. The authors aim at examining how faceted search interfaces are used in a digital library. They conducted an eye tracking user study and discuss challenges and approaches for analyzing gaze data.
- Poster by Hitomi Saito, Hitoshi Terai, Yuka Egusa, Masao Takaku, Makiko Miwa and Noriko Kando. *How Task Types and User Experiences Affect Information-Seeking Behavior on the Web: Using Eye-tracking and Client-side Search Logs*. They used screen-capture logs and eye tracking to identify differences in search behavior according to task type and search experience.
- Poster by Maristella Agosti, Franco Crivellari and Giorgio Maria Di Nunzio. *Evaluation of Digital Library Services Using Complementary Logs*. The authors argue that analyzing query logs alone is not sufficient to study user behavior. Rather, analyzing a larger variety of behavior logs (beyond query logs) and combining them leads to more accurate results.

Analyzing query logs in the aggregate

- Poster by Laura Granka. *Inferring the Public Agenda from Implicit Query Data*. The author presents an approach how to apply query log analysis to create indicators of political interest. As an example, poll ratings of presidential candidates are approximated by query log analysis.
- Poster by Suzan Verberne, Max Hinne, Maarten van der Heijden, Eva D'hondt, Wessel Kraaij and Theo van der Weide. *Annotating URLs with query terms: What factors predict reliable annotations?* The authors try to determine factors that predict the quality of URL annotations from query terms found in query logs.

Interpreting interaction feedback for an improved immediate/aggregated search/browsing experience

- Oral presentation by Mark Cramer, Mike Wertheim and David Hardtke: *Demonstration of Improved Search Result Relevancy Using Real-Time Implicit Relevance Feedback*. The paper reports about Surf Canyon, an existing browser plugin that interprets users' browsing behaviors for immediate improved ranking of results from commercial search engines. They show that incorporating user behavior can drastically improve overall result relevancy in the wild.
- Poster by Rui Li, Evelyn Rozanski and Anne Haake. *Framework of a Real-Time Adaptive Hypermedia System*. The authors present an adaptive hypermedia system that makes use of both browsing behavior and eye movement data of a user while interacting with the system. They use this information to automatically re-arrange information for more suitable user presentation.
- Poster by Max Van Kleek, David Karger and mc Schraefel. *Watching Through the Web: Building Personal Activity and Context-Aware Interfaces using Web Activity Streams*. They use user activity logs from Web-based information to build more personalized activity-sensitive information tools. They particularly focus on activity-based organization of user-created notes.
- Demo by Xuanhui Wang and ChengXiang Zhai. *Massive Implicit Feedback: Organizing Search Logs into Topic Maps for Collaborative Surfing*. In this demo, search and browsing logs from Web searchers are organized into topic maps so that users can follow the footprints from searchers who had similar information needs before.

Behavior-based evaluation measures

- Oral presentation by Emine Yilmaz, Milad Shokouhi, Nick Craswell and Stephen Robertson. *Incorporating user behavior information in IR evaluation*. The authors introduce a new user-centric measure (Expected Browsing Utility, EBU) for information retrieval evaluation which is reconciled with click log information from search engines.
- Poster by Tereza Iofciu, Nick Craswell and Milad Shokouhi. *Evaluating the impact of snippet highlighting in search*. The authors present the idea of highlighting important terms in search result snippets for helping the user to quickly identify whether a result matches the own query interpretation. They use speed and accuracy of clicks to evaluate the effect of highlighting.

5 Panel Discussion

The panel was moderated by Jacek Gwizdka; Claus-Peter Klas, Jeremy Pickens, and Xiaojun Yuan were panelists. Panelists first presented their views in short presentations and discussion followed.

Claus-Peter Klas talked about supporting users within a search task by logging activities between the user and the system. He emphasized the notion of an information dialog that supports the cognitive abilities of users not only for a query at hand, but rather for the complete search process. Information being most important to personalize the user's search experience would be the task context, the environment (both virtually concerning the computing device worked with as well as the physical location). This kind of information about the user should not only be applied for, personalization, and thus to increase the efficiency and effectiveness of search processes. It should also be used to teach information competence for search. To reach those goals, careful, but massive long-term logging will be needed as well as a repository of shared tools and data collections.

Jeremy Pickens focused on the problem of inferring information about cognitive processes during search from interaction logs. He pointed out the problem that interaction logs only provide information about physical actions and do not directly tell anything about the quality of the overall user experience which is especially true for sparse interaction logs like page transition logs. In order to proceed with inferring information about the user experience, he introduced ideas by Max Wilson and himself. For example, classifying meaningful pattern sequences in log files might help to indentify user intentions during specific time frames. Further, triangulating between different kinds of logs might lead to better estimations of user experience.

Xiaojun Yuan focused on the specific problem of using interaction logging data to inform the design of information visualization systems. She presented a user study which aimed to compare search performance using logged user interactions and dependent on individual domain knowledge.

6 Results of Breakout Groups

Three breakout groups were formed based on the interests of the participants. All groups had lively and very engaged discussions. In the end, all breakout groups reported about their results back to all participants of the workshop. Here, we briefly give an overview of the discussed topics.

Group 1 primarily discussed ideas and steps to share logging tools and frameworks as well as log data itself in the community. They decided that the best way of sharing frameworks and data would be to create a repository on the Web. The repository should contain a wiki with sections about logging tools and software, log data (as well as log meta-data), guides and how-to manuals, and a list of experts and research projects in the area. Furthermore, a mailing list should be created informing about updates of the wiki and about general topics of interest. The wiki will be created at http://culpool.dei.unipd.it.

Group 2 focused on issues concerning the interpretation of log data. The participants had experience with a great variety of user interaction logs, e.g., not only with mouse- and keyboard

based interaction with the computer, but also with eye tracking, skin-conductance feedback, and pressure-sensitive mice or seat cushions. They primarily created a list of software and tools that have proven useful for logging interaction data. (The list will be posted on the wiki.) Further, they discussed issues and problems encountered by the participants concerning the interpretation of eye tracking data and biometric feedback, i.e., time synchronization issues between logs of different modalities, problems arising from different log granularities, and challenges with respect to finding meaningful log patterns.

Group 3 dealt with behavior modeling. Since this is a very hard task with numerous facets, the participants concentrated on a more principled approach. They started with observing / logging user behavior and then posed the question what factors and theories might explain those behaviors. Then, from a cognitive perspective an answer should be found why the identified factors influence or yield the observed behaviors, whereas from an engineering perspective it should be tried to predict what they will do. This approach was then exemplified in several scenarios.

7 Conclusions

Over the course of the workshop, we have seen a great variety of types of logged user interactions, of methods how they are interpreted, and how this information is used and applied. Concerning the latter point, how log data is used and applied, we have seen an especially great variety: from personalization purposes, over a more informed visual design of search systems, to teaching users how to search more effectively.

However, the basis for all those different kinds of applications is the same: logged interaction data between a user and a system. There are basic kinds of interaction data, e.g., based on explicit events from the user while browsing the Web, such as clicks and page transitions as well as mouse movements and scrolling. More advanced and more implicit interaction data logging becomes more and more popular, e.g., based on eye tracking, skin conductance, and EEG. During the workshop, we identified common needs and problems with respect to logging interaction data. They reached from extracting the focused data from different software applications to merging interaction data streams from different sources. Here, we clearly see a need for a common basis of tools and frameworks shared within the community so that individual researchers don't have to re-invent the wheel over and over again.

As a consequence, the workshop participants agreed on setting up a wiki (http://culpool.dei.unipd.it) containing a collection of tools and frameworks that have been proven useful for interaction data logging and merging. In addition to the wiki, a mailing list will be set up informing about changes in the wiki and about topics of general interest to the community addressed by this workshop.

Acknowledgements

We would like to thank ACM and SIGIR for hosting this workshop as well as the SIGIR workshop committee and especially its chair Diane Kelly for their very helpful feedback. We are further very thankful to the authors, the members of our program committee, and all participants. They helped to form a very lively, spirited, highly interesting, and successful workshop.

References

[1] Belkin, N. J.; Bierig, R.; Buscher, G.; Elst, L. v.; Gwizdka, J.; Jose, J. & Teevan, J. (ed.): Proc. of the Workshop on Understanding the User – Logging and Interpreting User Interactions in Information Search and Retrieval (UIIR-2009), July 23, 2009, Boston, MA, USA. CEUR-WS.org, 2009, vol. 512.