EDITORIAL

People are using the World Wide Web (Web) to locate, seek, gather, and share information in increasingly complex ways. We now find it possible to interact with the Web from any number of devices, in varied settings, and for seemingly infinite types of tasks. In order to develop the next generation of Web interfaces, services, and systems, we must have a deep understanding of what people may do in these varied contexts. This requires observation beyond traditional methods such as interviews or usability lab experiments which typically employ small numbers of participants and convenient samples.

The availability of large transaction logs from both search engines and Websites and an increasing number of Web-based tools for lab studies has produced a wealth of research over the past ten or more years on search patterns for public Web search engines, vendor database products, institutional search services, browsing patterns, and click through. Much of the early research dealt with describing observed regularities in querying and browsing behaviors, both generally and specifically with searching. Findings of these studies have revealed -- with notable regularity across both search platforms and Websites -- that users do not engage in lengthy sessions (Jansen, Spink, & Saracevic, 2000), submit brief queries (Jansen & Spink, 2005), have defined but general goals (Choo, Betlor, & Turnbull, 1998; Kellar, Watters, & Shepherd, 2007) and do not browse extensively (Catledge & Pitkow, 1995). With many studies now having uncovered the presence of search and browsing regularities, what's next for Web usage research?

Basic analytics of user activity are easily available to Website owners but typically provide a shallow understanding of user satisfaction, usability, and usefulness. Current approaches to Web interaction studies do yield useful insights and solutions, but they often fail to produce comparable data and results among studies, not to mention standard definitions for even describing some Web-based activities. This inconsistency causes a huge duplication of effort among researchers from re-invented data collection instruments, perpetual development of data analysis applications, and even observation of redundant activities. We contend that building a toolbox of agreed upon methods, data collection tools, and analysis code will advance the state of Web user activity research. This will also allow a consistent progression for discovering how people use the Web, be it over time, by tasks, or via evolving interfaces and devices. This special issue of Web log tools and techniques is a first step in this endeavor.

What we propose, and what motivates this special issue, is a new generation of large-scale, empirical, naturalistic studies, with the necessary techniques and tools. These studies are indispensable to discover what people actually do when using the Web in a variety of environments, with their preferred applications and devices, and for real-world tasks of their own choosing. Much of this work will, by necessity, be outside of the lab in naturalistic settings. These naturalistic studies can focus on a increasingly nuanced set of factors related to Web usage, including trust and privacy (Hawkey & Inkpen, 2006). In order to conduct this kind of focused, experiential research, we require accurate, unobtrusive data collection methods and replicable, verifiable analysis techniques. With these methods and techniques, we need the tools to conduct such studies (Edmonds, 2003).

One goal of these methods is to yield more specific, contextual insights into user behavior while interacting with the Web. Furthermore, it is important to gather very large and often unprecedented behavioral datasets that provide more precise topical and statistical certainty upon analysis than has been achievable in prior investigations of information system use. It may very well be possible that the immense volumes of *quantitative* data collection possible with Web trace logging may in fact begin to yield actual *qualitative* insights into people's interaction with Web information and specifically the growing body of user generated content. Therefore, many researchers are turning to tools and techniques to capture and collate sequential activity and system-oriented data that extend beyond performance and page view reporting to organizing and understanding complex, multi-step user interactions.

This special issue presents several innovative research methods, data collection tools and case studies that address challenging issues and solutions for designing contextual studies, collecting and analyzing data, and interpreting and visualizing study results. In the first article in this issue, Muresan describes and discusses a formal but general methodology that integrates the conceptual design of the user interaction for interactive systems with the analysis of the interaction logs. In the second article, Velayathan and Yamada investigate the use of logged client-side browsing behavior to automatically predict user interest in a Web page.

In article three, Edmonds, White, Morris, and Drucker present a technique for the in-page logging of interaction events that aims to help interaction designers make more informed design decisions based on how users are interacting with their systems. Finally, Garofalakis, Giannakoudi, and Sakkopoulos discuss the *Ontology-oRiented usaGe ANalysis* (ORGAN) system, which aims to enhance and ease log analysis by using semantic knowledge. The system offers standard statistical analysis of Web usage logs while considering the Website's underlying semantics.

The articles presented in this issue represent a toolkit of possible techniques and uses that can be combined and customized to address a wide variety of research efforts beyond their intended uses. We encourage you to adapt and apply these ideas and tools for your own research and help advance the state of Web interaction understanding.

References

Catledge, L.D., & Pitkow, J.E. (1995). Characterizing browsing strategies in the World Wide Web. Computer Network and ISDN Systems, 27(1), 1065-1073.

Choo, C., Betlor, B., & Turnbull, D. (1998). In A Behavioral Model of Information Seeking on the Web: Preliminary Results of a Study of How Managers and IT Specialists Use the Web (pp. 290-302). Paper presented at the the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA. ASIS.

Edmonds, A. (2003). Uzilla: a new tool for web usability testing. Behavior Research Methods, Instrumentation and Computers, 35(2), 194-201.

Hawkey, K., & Inkpen, K.M. (2006, 22-26 May). In Examining the Content and Privacy of Web Browsing Incidental Information (pp. 123-132). Paper presented at the World Wide Web 2006, Edinburgh, UK.

Jansen, B.J., & Spink, A. (2005). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing & Management, 42(1), 248-263.

Jansen, B.J., Spink, A., & Saracevic, T. (2000). Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. Information Processing & Management, 36(2), 207-227.

Kellar, M., Watters, C., & Shepherd, M. (2007). A Field Study Characterizing Web-based Information Seeking Tasks. Journal of the American Society for Information Science and Technology, 58(7), 999-1018.

JWE Guest Editors

- Andy Edmonds, Windows Live Search, Microsoft Corporation, andyed@gmail.com
- Kirstie Hawkey, Faculty of Computer Science, Dalhousie University
  6050 University Ave., Halifax, NS, B3H 1W5
  hawkey@cs.dal.ca
- Bernard J. Jansen, College of Information Sciences and Technology, The Pennsylvania State University, University Park, Pennsylvania 16802, jjansen@acm.org
- Melanie Kellar, Faculty of Computer Science, Dalhousie University
  6050 University Ave., Halifax, NS, B3H 1W5
  melanie@cs.dal.ca
- Don Turnbull, University of Texas at Austin, School of Information, donturn@ischool.utexas.edu