

Making the Web More User-Friendly

By [Rob Knies](#)

April 22, 2009 2:00 PM PT

Search logs can provide immense value to researchers. Those examining search queries can look at raw data, view trends, and formulate hypotheses that form the basis for further study. And with the hundreds of millions of searches performed by major search engines each day, the data set of search logs is astoundingly rich and growing richer by the second.

But protecting the privacy of the data is critical, and that reality is acknowledged by Aleksandra Korolova of Stanford University and [Krishnaram Kenthapadi](#), [Nina Mishra](#), and [Alexandros Ntoulas](#) of Microsoft Research's Internet Services Research Center (ISRC) [Search Labs](#). In fact, they're working assiduously to secure such information.

In a paper entitled [Releasing Search Queries and Clicks Privately](#), submitted to the 18th International World Wide Web Conference (WWW), to be held April 20-24 in Madrid, Spain, the researchers argue that queries, clicks, and associated, perturbed counts—the data that interests the research community—can be published in a manner that rigorously preserves privacy.

The paper, nominated for the conference's Best Paper Award, is an example of how Microsoft Research is committed to working openly with the industry and academia to drive state-of-the-art technology to produce a more accessible, searchable, user-friendly Internet. But it's hardly the only one.

Microsoft Research, a silver sponsor for this year's event, has been an active leader, contributor, and participant for many years in WWW, a global event that draws key researchers, innovators, decision-makers, technologists, businesses, and standards bodies working to shape the Web.

Of 104 peer-reviewed papers accepted for WWW 2009, 16—15 percent—were written wholly or in part by Microsoft Research, which has more papers in the conference than any other organization. Four of Microsoft Research's six labs worldwide—those in Beijing; Cambridge, U.K.; Redmond; and Mountain View, Calif.—are represented in the papers to be presented.

That level of participation reflects Microsoft Research's intention of advancing the state of the art of the next generation of Internet technologies, finding answers to the Internet's greatest challenges, and extending the boundaries of the Internet.

The Microsoft Research papers accepted by WWW 2009 fall into six areas, each listed with a representative submission:

- **Data Mining:** The aforementioned, best-paper nomination by Korolova, Kenthapadi, Mishra, and Ntoulas not only shows how query logs can be published while maintaining privacy, but also examines the opposite side of the issue: Are data that can be safely published of real value?
- **Internet Monetization:** In [How Much Can Behavioral Targeting Help Online Advertising?](#) Jun Yan, Ning Liu, [Gang Wang](#), and [Zheng Chen](#) of [Microsoft Research Asia](#), along with Wen Zhang of the University of Science and Technology of China and Yun Jiang of Shanghai Jian Tong University, examine the ads click-through log from a commercial search engine over a week and draw conclusions, in the first empirical study in academia on behavioral targeting in online advertising.
- **Rich Media:** [Tag Ranking](#)—written by Dong Liu of the Harbin Institute of Technology; [Xian-Sheng Hua](#), [Linjun Yang](#), and Meng Wang of Microsoft Research Asia; and [Hong-Jiang Zhang](#) of the Microsoft Research [Advanced Technology Center](#)—proposes a scheme by which users' tags of shared images are automatically ranked according their relevance to the image content.
- **Search:** In the absence of explicit knowledge of user intent, search engines are prone to substitute diversity for relevance in presenting results. In [An Axiomatic](#)

Related Links

- [Krishnaram Kenthapadi](#)
- [Nina Mishra](#)
- [Alexandros Ntoulas](#)
- [Search Labs](#)
- [Releasing Search Queries and Clicks Privately](#)
- [How Much Can Behavioral Targeting Help Online Advertising?](#)
- [Gang Wang](#)
- [Zheng Chen](#)
- [Microsoft Research Asia](#)
- [Tag Ranking](#)
- [Xian-Sheng Hua](#)
- [Linjun Yang](#)
- [Hong-Jiang Zhang](#)
- [Advanced Technology Center](#)
- [An Axiomatic Approach to Result Diversification](#)
- [Sreenivas Gollapudi](#)
- [Behavioral Profiles for Advanced Email Features](#)
- [Thomas Karagiannis](#)
- [Milan Vojnovič](#)
- [Microsoft Research Cambridge](#)
- [Mining Interesting Locations and Travel Sequences from GPS Trajectories for Mobile Users](#)
- [Yu Zheng](#)
- [Xing Xie](#)
- [Wei-Ying Ma](#)
- [A Game Based Approach to Assign Geographical Relevance to Web Images](#)
- [Click Chain Model in Web Search](#)
- [Exploiting Web Search Engines to Search Structured Information Sources](#)
- [Exploiting Web Search to Generate Synonyms for Entities](#)
- [Incorporating Site-Level Knowledge to Extract Structured Data from Web Forums](#)
- [Learning Consensus Opinion: Mining Data from a Labeling Game](#)

Research ISRC Search Labs and Aneesh Sharma of Stanford describe a set of natural axioms that a diversification system is expected to satisfy and demonstrate that no diversification function can satisfy all the axioms simultaneously.

- **Social Networks and Web 2.0:** *Behavioral Profiles for Advanced Email Features*, by Thomas Karagiannis and Milan Vojnovič of Microsoft Research Cambridge, examines e-mail usage patterns in a large-scale enterprise over a three-month period to learn what replies depend on and how friends of friends augment the e-mail experience. Their conclusions provide significant insights into informed design of advanced e-mail features.
- **User Interfaces and Mobile Web:** In *Mining Interesting Locations and Travel Sequences from GPS Trajectories for Mobile Users*, by Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma of Microsoft Research Asia, the authors aim to identify interesting locations and classical travel sequences in a given region. Such information can help users understand surrounding locations and would enable travel recommendations.

- [Learning to Tag](#)
- [StatSnowball: a Statistical Approach to Extracting Entity Relationships](#)
- [Towards Context-Aware Search by Learning a Very Large Variable Length Hidden Markov Model from Search Sites](#)
- [Understand User's Query Intent with Wikipedia](#)

The complete list of Microsoft Research papers accepted for WWW 2009:

[A Game Based Approach to Assign Geographical Relevance to Web Images](#), by Yuki Arase, Osaka University; Xing Xie, Microsoft Research Asia; Manni Duan, University of Science and Technology of China; Takahiro Hara, Osaka University; and Shojiro Nishio, Osaka University.

[An Axiomatic Approach to Result Diversification](#), by Sreenivas Gollapudi, Microsoft Research Internet Services Research Center Search Labs; and Aneesh Sharma, Stanford University.

[Behavioral Profiles for Advanced Email Features](#), by Thomas Karagiannis, Microsoft Research Cambridge; and Milan Vojnovic, Microsoft Research Cambridge.

[Click Chain Model in Web Search](#), by Fan Guo, Carnegie Mellon University; Chao Liu, Microsoft Research Redmond; Tom Minka, Microsoft Research Cambridge; Yi-Min Wang, Microsoft Research Redmond; and Christos Faloutsos, Carnegie Mellon University.

[Exploiting Web Search Engines to Search Structured Information Sources](#), by Sanjay Agrawal, Microsoft Research Redmond; Kaushik Chakrabarti, Microsoft Research Redmond; Surajit Chaudhuri, Microsoft Research Redmond; Venkatesh Ganti, Microsoft Research Redmond; Arnd König, Microsoft Research Redmond; and Dong Xin, Microsoft Research Redmond.

[Exploiting Web Search to Generate Synonyms for Entities](#), by Surajit Chaudhuri, Microsoft Research Redmond; Venkatesh Ganti, Microsoft Research Redmond; and Dong Xin, Microsoft Research Redmond.

[How Much Can Behavioral Targeting Help Online Advertising?](#) by Jun Yan, Microsoft Research Asia; Ning Liu, Microsoft Research Asia; Gang Wang, Microsoft Research Asia; Wen Zhang, University of Science and Technology of China; Yun Jiang, Shanghai Jian Tong University; and Zheng Chen, Microsoft Research Asia.

[Incorporating Site-Level Knowledge to Extract Structured Data from Web Forums](#), by Jiang-Ming Yang, Microsoft Research Asia; Rui Cai, Microsoft Research Asia; Yida Wang, Chinese Academy of Sciences; Jun Zhu, Tsinghua University; Lei Zhang, Microsoft Research Asia; and Wei-Ying Ma, Microsoft Research Asia.

[Learning Consensus Opinion: Mining Data from a Labeling Game](#), by Paul Bennett, Microsoft Research Redmond; Max Chickering, Microsoft Live Labs; and Anton Mityagin, Microsoft Live Labs.

[Learning to Tag](#), by Lei Wu, Microsoft Research Asia; Linjun Yang, Microsoft Research Asia; Nenghai Yu, University of Science and Technology of China; and Xian-Sheng Hua, Microsoft Research Asia.

[Mining Interesting Locations and Travel Sequences from GPS Trajectories for Mobile Users](#), by Yu Zheng, Microsoft Research Asia; Lizhu Zhang, Microsoft Research Asia; Xing Xie, Microsoft Research Asia; and Wei-Ying Ma, Microsoft Research Asia.

[Releasing Search Queries and Clicks Privately](#), by Aleksandra Korolova, Stanford University; Krishnaram Kenthapadi, Microsoft Research Internet Services Research

Center Search Labs; and Alexandros Ntoulas, Microsoft Research Internet Services Research Center Search Labs.

[StatSnowball: a Statistical Approach to Extracting Entity Relationships](#), by Jun Zhu, Tsinghua University; Zaiqing Nie, Microsoft Research Asia; Xiaojiang Liu, Microsoft Research Asia; Bo Zhang, Microsoft Research Asia; and Ji-Rong Wen, Microsoft Research Asia.

[Tag Ranking](#), by Dong Liu, Harbin Institute of Technology; Xian-Sheng Hua, Microsoft Research Asia; Linjun Yang, Microsoft Research Asia; Meng Wang, Microsoft Research Asia; and Hong-Jiang Zhang, Microsoft Research Advanced Technology Center.

[Towards Context-Aware Search by Learning a Very Large Variable Length Hidden Markov Model from Search Sites](#), by Huanhuan Cao, University of Science and Technology of China; Daxin Jiang, Microsoft Research Asia; Jian Pei, Simon Fraser University; Enhong Chen, University of Science and Technology of China; and Hang Li, Microsoft Research Asia.

[Understand User's Query Intent with Wikipedia](#), by Jian Hu, Microsoft Research Asia; Gang Wang, Microsoft Research Asia; Fred Lochovsky, The Hong Kong University of Science and Technology; and Zheng Chen, Microsoft Research Asia.

 Share

 E-mail this

 Print

[Back to top](#) 