

# The Past, Present and Future of Web Information Retrieval (Invited Tutorial)

Monika Henzinger  
Google Inc.  
2400 Bayshore Parkway  
Mountain View, CA 94043  
U.S.A.  
monika@google.com

## ABSTRACT

Web search engines have emerged as one of the central applications on the Internet. In fact, search has become one of the most important activities that people engage in on the the Internet. Even beyond becoming the number one source of information, a growing number of businesses are depending on web search engines for customer acquisition.

The first generation of web search engines used text-only retrieval techniques. Google revolutionized the field by deploying the PageRank technology - an eigenvector-based analysis of the hyperlink structure - to analyze the web in order to produce relevant results. Moving forward, our goal is to achieve a better understanding of a page with a view towards producing even more relevant results.

An exciting new form of search for the future is query-free search: While a user performs her daily tasks, searches are automatically performed to supply her with information that is relevant to her activity. We present one type of query-free search, namely query-free news search: While a user watches TV news the system finds in real-time web pages that are relevant to the news stories.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*PODS 2004* June 14-16, 2004, Paris, France.

Copyright 2004 ACM 1-58113-858-X/04/06 ... \$5.00.