Spam: It’s Not Just for Inboxes Anymore
pp. 28-34
Zoltán Gyöngyi and Hector Garcia-Molina

When an innocent search for a local auto-body repair shop lands someone on a page that promotes great mortgage rates, inexpensive prescription drugs, legal assistance, or pain-free cosmetic surgery, that person has been Web spammed. This practice undermines the trustworthiness people have come to expect from search engines. Google, Yahoo, and MSN have built their reputation on providing reliable, unbiased, trustworthy references. Web spammers exploit this trust, a practice that has far-reaching technical, economic, and social implications. Search engine companies will continue to fight spammers, but in the long run, the best solution to the ongoing battle will be to make spamming ineffective.

Using Web Search Engines to Find and Refind Information
pp. 36-42
Robert G. Capra III and Manuel A. Pérez-Quiñones

To better inform the design of next-generation Web search tools, researchers have begun to examine users’ finding and refining behaviors as well as the limitations of existing search technologies. As part of this effort, Virginia Tech’s Center for Human-Computer Interaction is exploring how users find and refine online data and the factors that affect these processes.

Synthesizing results from one of their laboratory studies with related work, the authors have developed a search engine use model based on how frequently users perform a given task as well as how familiar they are with that task.

Intelligent Search Agents Using Web-Driven Natural Language Explanatory Dialogs
pp. 44-52
Anita Ferreira and John Atkinson

Although traditional keyword-based information search systems can provide a first step for use in the overall search process, the challenge is to perform tasks more accurately and intelligently, applying a user’s knowledge to improve search capabilities.

The authors explore the generation of natural-language interactive dialogs for bibliographical searching on the Web to improve the searching and filtering process by minimizing user exchanges. Their approach focuses on enhancing the information-searching paradigm with both a computational linguistics model and a more suitable search agent.

Searching Association Networks for Nurturers
pp. 54-61
Bharath Kumar Mohan

Search the Web involves more than sifting through one huge graph of pages and hyperlinks. Specific association networks have emerged that serve domain-specific queries better by exploiting the principles and patterns that apply there.

A technique that searches these association networks and finds nurturers—early adopters—in them can be particularly effective. Finding nurturers can improve Web search, especially when answering sticky queries—persistent searches for which a user expects frequent fresh updates. Studying the evolution of association networks offers insights that can be used for developing new forms of information retrieval and for improving searches.

Search on the Semantic Web
pp. 62-69
Li Ding, Tim Finin, Anupam Joshi, Yun Peng, Rong Pan, and Pavan Reddiwarai

Most knowledge on the Web is presented as natural-language text with occasional pictures and graphics. Although it is convenient for human users to read and view, this format limits the indexing capabilities of state-of-the-art search engines because they cannot infer meaning. Thus, users share a significant burden in terms of constructing search queries intelligently.

Even with increased use of XML-encoded information, computers still must use application-dependent semantics to process the tags and literal symbols.

Developers must experiment with how much and where a Semantic Web search engine should reason over the contents of documents and queries.

CPE: A Parallel Library for Financial Engineering Applications
pp. 70-77

To effectively utilize parallel computing, developers need software that reduces the complexity of the process as well as tools to support integration of parallel and desktop machines. The Clustertech parallel environment (CPE) is a C++ library that facilitates development of large-scale parallel applications.

This environment offers unprecedented efficiency in developing financial engineering applications and porting them to a wide range of parallel-cluster and symmetric-multiprocessing machines. CPE’s object-oriented approach facilitates information hiding as well as code reuse.