

---

## Introduction

BETTINA FABOS

---

The Global Brand of the Year in 2003 title did not go to Coca-Cola, Nike, or Starbucks, some of the most ubiquitous commercial names in our midst. Instead, it went to Google, a highly used but lightly promoted search engine, which beat Apple for the second year in a row. The leading brand consultancy, Interbrand, had surveyed about 4,000 “branding professionals” who determined that the Google brand had made the most impact internationally (Google voted, 2004). To think of Google—the most popular searching tool on the Web—as a *brand* is important for this issue of *Library Trends* because it underscores how closely mainstream online information resources are tied to the commercial economy.

The Web has been a commercial medium since 1995, when the government quietly sold the Internet’s backbone (previously controlled by the National Science Foundation) to private enterprise. Ten years ago we saw the beginning of a tremendous push—from the Clinton administration, Bill Gates, and the computer and telecommunication industries in general—to get schools and libraries wired. The push, it turns out, was not necessarily to bring the promised “universe of knowledge” (Clinton’s words) to all young and “lifelong” learners alike. Instead, the push was a careful public relations strategy to build up a user base so that the Web could become a viable commercial advertising medium (Fabos, 2004). Indeed, the rhetoric and accompanying media campaign of the mid-1990s was successful: in just five short years, the Web (as part of the larger Internet) became a mass medium—faster than any communication medium before it.

Before 1998, search engine providers such as AltaVista and Google were some of the most popular destinations on the Web. Beyond syndicating their services to search portals like Yahoo!, however, they generated low

revenue because they were merely the stepping stones to other content-rich pages containing banner ads. That all changed in 1998, when the startup Goto.com began combining the impartial algorithmic searches from search engine providers (usually one of the top five: AltaVista, AlltheWeb, Google, Inktomi, or Teoma) with a database of advertisers, so that many searches, unbeknownst to users, became prioritized according to the highest advertising bidder. Suddenly there was money in search engines. Goto syndicated its services to all the leading search portals, with the rationale that most people search for commercial products anyway. Then the impartial search engine providers themselves began to skew their searches in favor of commercial enterprise. Except for Google, all search engine providers implemented paid inclusion practices: accepting flat fees for including a client's Web page in *every* search conducted.

In that year the Yahoo! portal, which had been syndicating Inktomi's and then Google's impartial search results, purchased Inktomi outright. Then the leading commercial search provider, Overture (formerly Goto), purchased AltaVista and AlltheWeb. And not long afterwards, Yahoo! purchased Overture, an acquisition that put three of the top search engine providers and the leading advertising index under one portal. And perhaps most significantly, Microsoft (by now regretting not getting into the search business sooner) tried to buy Google in 2003 but ended up building its own search engine provider, which was launched on the MSN site in February 2005 and will be bundled with Microsoft's next Windows operating system, "Project Longhorn," in 2006.

Search engines, once solely the online conduit for information, have taken on the contradictory role of conduit for online commerce. These days, even Google, the "ethical" search engine with the company motto "Don't Be Evil," is now focusing most of its attention on ad placement, either through its own search pages or through "contextual links" on other content pages (a practice that undermines the very integrity of its own PageRank algorithm). Indeed, the company's success in this vein is all too evident: Google sold \$1 billion of advertising in the last three months of 2004 (Markoff & Ives, 2005). Reflecting on the company's motto after Google went public in 2004, a *New York Times* editorial stated: "Such idealistic talk out of Silicon Valley, so seemingly empowering back in 1999, seems embarrassingly naïve now that the party's ended, at least for the rest of us" (Googling Google, 2004, p. 10). Such is the fate for all of us when Google the search engine became Google the brand.

This issue of *Library Trends* addresses Web content within the context of Internet commercialization and democracy. These are big ideas and problems, with potentially big solutions, so this issue has cast a wide net, pulling together voices from multiple disciplines, including communication studies, informatics, information management, research programming, computer science, engineering, and library science. I hope this issue highlights the

need for and value of continuing interdisciplinary cooperation and cross-fertilization. We have so much to learn from each other.

The issue is organized into two sections, with one section posing “the problem” of the commercialized Web and the other section posing multiple “solutions.” Part 1, “From Information to Commercial Highway,” examines the political and economic forces that shape and control online content. In “Links and Power: The Political Economy of Linking on the Web,” Jill Walker presents search engines as commercial entities that reinforce the most powerfully funded information—either commercial content or information that is already dominant in the mainstream media. This article provides a base for the next three, which look at the specific consequences of a commercial search environment on student research. In “On Their Own: Students’ Academic Use of the Commercialized Web,” Samuel Ebersole writes about high school students’ use of search engines, concluding that students’ research is inundated with commercial sources and that students do not have sufficient help in negotiating this environment.

In “Student Searching Behavior and the Web: Use of Academic Resources and Google,” Jillian Griffiths and Peter Brophy discuss student searching tendencies at the college level. Among their findings are a heavy student reliance on search engines rather than other academic resources, a general sense that search engines are all inclusive, and a significant difficulty in finding information and resources via search engines, with students trading performance for the path of least cognitive resistance. Finally, Julie Frechette goes one step beyond the world of search engines with an investigation of Web filtering software, which public officials are currently pushing in public libraries and schools. The article “Cyber-Democracy or Cyber-Hegemony? Exploring the Political and Economic Structures of the Internet as an Alternative Source of Information” illustrates how filtering software companies have become the second tier, after search engines, of lucrative Web properties that feed the commercial system. Like search engines, which now act to streamline and control much of the content on the Web, filtering software is suppressing certain kinds of “illicit” content while surreptitiously promoting commercial interests and commercial content.

What these four articles suggest is that, despite the huge amount of trust search engines and filtering tools continue to garner in the public sector, they are far from neutral and relentlessly steer users toward mainstream, mostly commercial content. That would be fine if one was interested in buying shoes. Indeed, the business sector tells us how happy we should be to find shoes that so precisely fit our marketing profile. But students and researchers looking for noncommercial, or at least nonmainstream, content, trying to gather a wide range of information containing as many disparate viewpoints as possible, or trying to access research that is controversial will not be successful, ultimately, in a research environment controlled by commercial interests.

Part 2, "Harnessing the Web for Noncommercial Purposes," offers a glimpse of the many exciting international projects underway that are providing alternate routes to online content. If part 1 of this issue poses the problem of search engine commercialization, part 2 offers multiple solutions. Numerous computer scientists and digital librarians have been developing open source technology, such as the Open Access Initiative for Metadata Harvesting Protocol (OAI-PMH), iVia, and Data Fountains, that offer (and enhance) a user's ability to search across multiple (that is, thousands of) subject gateways. These digital repository harvesting services imitate the functions and interface of a search engine, but they can be moulded to search in specific academic areas. In other words, one can create completely noncommercial searching environments that offer the scope and feel of a search engine. Do not be scared off by the unfamiliar acronyms because these developments have profound implications for academic research.

In "Current Developments and Future Trends for the OAI Protocol for Metadata Harvesting," Sarah Shreeves, Thomas Habing, Kat Hagedorn, and Jeffrey Young report on the latest developments and future directions within the OAI community. Their article provides a succinct history of the OAI Protocol movement, which got its start in 2001 and since then has become widely adopted as an international standard in digital archiving and subject gateway development. Shreeves, Habing, Hagedorn, and Young are at the forefront of the OAI initiative. So are Xiaoming Liu, Kurt Maly, Michael L. Nelson, and Mohammad Zubair, who co-authored "Lessons Learned with Arc, an OAI-PMH Service Provider." This article discusses the success of Arc, the first end-user OAI-PMH service provider. The searchable repository, which currently indexes about seven million records from several hundred subject gateways, has an immense scale and is a model of future academic Web searching. Besides detailing the Arc system, this article outlines the ongoing research devoted to improving Arc.

Beyond the OAI protocol, Steve Mitchell outlines the essence of iVia technology, a virtual library collection-building software platform. In "Collaboration Enabling Internet Resource Collection-Building Software and Technologies," Mitchell details a decade of ongoing research at the library of the University of California that aims to help librarians and subject gateway developers more efficiently build metadata collections. Mitchell discusses iVia's implementation in INFOMINE, a vast subject gateway that holds significant scholarly and educational resources on the Internet. He also mentions the ongoing work on the open source software system called Data Fountains, which expands upon the iVia system in areas of automated data harvesting that are intrinsically tied to the active skills and wisdom of participating librarians. Edward Almas, co-director of the Internet Scout Project, is also dedicated to facilitating the development of easily searchable subject gateway systems for the academic community and beyond. In

“Tools for Creating Your Own Resource Portal: CWIS and the Scout Portal Toolkit,” Almsy describes a user-friendly means for building high-quality subject gateways. He and his colleagues have developed the Scout Portal Toolkit (SPT) and the Collection Workflow Integration System (CWIS), both complementary technical resources for subject gateway development. Almsy offers detailed descriptions of CWIS in particular, as well as information about hardware and software requirements and support in locating all appropriate software.

The next article describes another kind of toolkit. In their contribution, “Gateway Standardization: A Quality Assurance Framework for Metadata,” Brian Kelly, Amanda Closier, and Debra Hiom discuss the various strategies for streamlining metadata when adding new resources and maintaining subject gateways once they are built. Through trial, error, and careful testing since 2001, they have developed a toolkit that works as a straightforward self-assessment tool for subject gateway developers. Finally, Paul Jones, in “Strategies and Technologies of Sharing in Contributor-Run Archives,” illustrates the important developments in the collaborative subject gateway movement. He discusses contributor-run archives such as the Linux Documentation Project, the Degree Confluence Project, and Etree.org—all technologically inventive portals supported by passionate volunteers who cooperate to build these open source services. No librarians are involved—just experts and public citizens dedicated to sharing their knowledge and/or creative efforts with others.

I hope these writers’ contributions in this issue give us ideas when considering our fate as users of a Web that has become dominated by the powerful commercial economy. The ongoing work toward subject gateway development—all of it developed as free, open source software—provides a small but growing countervailing force to the commercialization of “the universe of knowledge.” Underlying all these efforts is the understanding that, for a democracy to function properly, one needs access to all kinds of information, not just information with a commercial purpose. Also underlying these efforts is the understanding that, in our commercial system, librarians and citizens interested in nurturing a public sphere must work together to control our destiny as information specialists—or somebody else will.

## REFERENCES

- Fabos, B. (2004). *Wrong turn on the information superhighway: Education and the commercialization of the Internet*. New York: Teachers College Press.
- Google voted best brand of 2003. (2004). *BBC News* [Electronic version]. Retrieved November 11, 2004, from <http://news.bbc.co.uk/2/hi/business/3456363.stm>.
- Googling Google [Editorial desk]. (2004, May 2). *New York Times*, p. 10.
- Markoff, John, and Ives, Nat. (2005, February 4). Web search sites see clicks add up to big ad dollars [Business/Financial Desk]. *New York Times*, p. 1.