

The Unbreakable Google Monopoly

T. Goldsmith, Azinet LLC
January 2007 REV 4/20/07

[Search Engine Honesty](#)

This paper summarizes the situation regarding the possibility of significant competition for Google search engine and targeted advertising activities and concludes that Google has an unbreakable functional monopoly for the high-quality search business that is similar to and possibly more extensive and disturbing than Microsoft's position regarding operating systems.

In the past, major search engines such as Altavista and Excite seemed to have very dominant positions and were nonetheless subsequently almost completely superseded by newer search services particularly Google. However, the current situation is very different and Google's dominant position regarding high-quality Internet search is very secure because of a number of factors that did not apply to previous search operators including:

- Massive Investment Requirement
- Maturity of Search and Web Technology
- Operations Data
- Economies of Scale
- Tracking Data Volume
- Anti-Competitive Practices
- Declining Captive Audiences

Massive Investment Requirement

Given informed choice, search engine users choose a search service based on the following parameters: (Our poll indicates that the importance is in the stated order.)

- Quality of web sites listed in results¹ (relevance, freedom from spam)
- Freshness of results (speed with which new web information is incorporated into search results)
- Comprehensiveness of results (number of sites and pages that can potentially be found with a search)
- Speed with which results displayed during a search

We can define "search quality" as encompassing all of these parameters. The cost of satisfying each of these parameters is driven by the size of the web, which now encompasses billions of web pages. Where a university project could produce a competitive search product ten years ago, explosive growth in the web has resulted in a

¹ **Search Results:** Specifically non-sponsored listings as opposed to paid, "sponsored results."

situation in which enormous up-front investment would be required to produce a search service that could plausibly compete with Google in terms of the above parameters.

A search engine must acquire and store all the web data that it covers. A search engine must repeat this acquisition often to achieve freshness. A search engine must be able to search through all this data in fractional seconds. Google is said to have 150,000 computers and massive Internet connectivity.

It is true that a search engine that is superficially similar to Google could be (and has been) produced with vastly fewer resources. Google indexes approximately 10 *billion* web pages and revisits them approximately monthly. (Many pages are visited much more often.) Imagine a search engine that only indexed 10 *million* pages and only visited them every six months. This engine would require about 1,000 times less computational resources and 6,000 times less connectivity, an enormous cost reduction. A search for a popular search term (Britney Spears, hot chicks, MP3 files) might well produce results acceptable to a casual user. However, apparently people (even casual and relatively unsophisticated searchers) want the superior results they can get from a better search service.

Maturity of Search and Web Technology

Search and the underlying Internet web technology including servers and browsers are now relatively well developed. There is little or no chance of a revolutionary new invention or technological breakthrough that would allow a competitive advantage sufficient to override the other factors. Because of the size of the web (number of deployed sites, pages, servers, browsers, etc.) any technological change would tend to spread relatively slowly, allowing Google time to catch up. In the event a new company somehow developed new technology that promised an advantage, Google's enormous market cap would allow it to outbid most, likely all, competitors.

Operations Data

Google Search is based on innovations² made by cofounders Brin and Page that produced substantial improvements in search quality. Patents on these and other innovations contribute to Google's monopoly position. These innovations are augmented by years of accumulated data and incremental development resulting from the actual day-to-day operation of the Google system in today's web environment. A competitor with comparable technology and comparable installed computer capacity would still be years behind in the accumulation of this operating data and would likely never catch up.

² *Brin, S. and Page, L., [The Anatomy of a Large-Scale Hypertextual Web Search Engine](#), Stanford University, 1998*

Economies of Scale

Google has more than twice the search traffic³ of its nearest competitor (Yahoo Search) and more than five times the traffic of its next nearest competitor (MSN Search). Because of the large portion of search system cost represented by “front-end” infrastructure, that is, costs that must be expended regardless of search volume, Google has a major advantage over rivals with lower traffic. Income is proportional to traffic; cost is not.

Tracking Data

Google’s large advantage in search traffic and extremely pervasive AdSense advertising placements allow it to acquire *tracking data*⁴ on an extremely large number of web users relative to any existing competitor much less any new competitor. This tracking data advantage allows Google to control spam⁵ to a much larger degree than otherwise possible and therefore allows Google to deliver results of higher quality than competitors. Web spam has been and continues to be a major problem for all search engines. Because spammers evolve and for other reasons, spam control represents recurring cost.

In addition, the tracking data advantage allows Google to better control click-fraud and other issues that would otherwise very adversely affect its AdSense advertising system. This advantage allows Google to achieve a higher level of automation and therefore lower per-partner-site cost in managing its AdSense advertising partners. Consequently, it is unlikely that any competitor will be able to develop an advertising system with the degree of automated management possessed by the Google AdSense system. Competitor’s per-site labor costs will therefore always be higher, limiting their ability to place ads on smaller web sites and eliminating their access to an entire market segment (small web site advertising). Google will therefore retain its substantial monopoly in small web site pay-by-click advertising. Notice the circular situation: Google’s *ability* to place ads on small web sites is, in part, enabled by the fact that they *are* placing ads on small sites and therefore possess more tracking data.

Google’s advertising system is so effective that Google can economically and effectively display targeted text ads on essentially any web site with apparently zero manual review. No competitor has a similar capability.

³ **Search Traffic:** Nielsen//Netratings measured searches performed by U.S. home and work web surfers for July 2006: Google and its partners (AOL Search) using Google data 55.5 percent of searches, Yahoo 23 percent, MSN Search 9.6 percent, Ask 2.6 percent, others 8.5 percent.

⁴ **Tracking data:** Data showing the web usage (pages visited, time spent on each page, time-of-day visits were performed, subject matter of pages visited, etc.) of individual web users. Usage patterns of legitimate users differ from click-fraud participants and can also be used to identify spam sites. Tracking data can be obtained anytime a user visits a web site owned by a search engine or one of its advertising partners.

⁵ **Spam:** A web site listed in search results that turns out to have little or no useful content or that is not relevant to the subject of the search.

Google's purchase of major web advertising network DoubleClick (if approved by regulators) will increase Google's tracking advantage.

Google Anticompetitive Practices

Google (as well as other search engines) can use *manual censoring*⁶ of non-sponsored search results or *site-unique bias*⁷ to suppress access by their users to competitive services such as directories. Our studies⁸ show that search results pointing to small-business directories are very heavily censored by Google. These techniques⁹ can also be used to suppress user access to sites that compete with Google's activities in other areas (selling of things, video, pictures, message boards, etc.) or which compete with Google partners. Manipulation of unpaid (non-sponsored) search results for anti-competitive purposes is currently unregulated.

Declining Captive Audiences

Many MSN users use MSN Search because it has been integrated into MSN as the "default" search service. Similar default settings can be found on ISP supplied browsers and Microsoft Internet Explorer. As the user population becomes more educated and knowledgeable regarding the Internet, fewer will use "beginner" services such as MSN and more will change default settings to use Google. Users will graduate to Internet service providers that provide superior or cheaper Internet access with fewer beginner features such as cable Internet services or telephone company DSL or fiber optic services. Simultaneously, MSN's more sophisticated users will tend to choose Google because of superior performance as opposed to using the "pre-chosen" MSN search service. Similar issues will tend to erode Yahoo's share of the search business. Google's search share will therefore continue to increase.

Conclusion

Google has an effective unbreakable monopoly regarding high-quality Internet search and associated advertising. Google's stock price, relative to Yahoo, suggests that the market has reached a similar conclusion. Because of the factors noted above, it is very unlikely that Yahoo Search, MSN Search, or an emerging competitor would be successful in matching Google in this arena and so potential competitors are likely to expend their resources elsewhere. It is therefore likely that the differences in quality between Google and other search services will increase.

⁶ **Manual Censoring:** Google admits to manual censoring ("banning") in which individual small-business web sites are excluded from the Google search index and therefore cannot be found in a Google search. Google does not disclose its reasons for censoring a particular site. Small-business directory sites are very heavily censored by Google.

⁷ **Site-Unique bias:** The use of a search engine ranking or depth-of-crawl algorithm that contains or refers to a list of domain names, trademarks, site-unique phrases, or other site-unique information in order to negatively or positively bias search result ranking or page indexing of individual, hand-selected web sites.

⁸ **Case Studies:** See [Kinderstart Case](#), [SeekOn Case](#), [Search Engine Censoring of Sites Using ODP Data](#)

⁹ **Search Engine Bias Issues:** [Impact of Search Engine Editorial Policies on Small Business and Freedom of Information](#), Azinet LLC, April 2007

Although unsophisticated users and people using search for entertainment can continue to be served by Yahoo search, MSN Search, and other existing or emerging services, Google is the only source of the highest quality search. People who need high-quality search for business, research, and other more serious activities have an increasingly sole-source situation.

There are also freedom-of-information issues associated with a monopoly or even extremely limited competition in high-quality search. Search engines are seen as publishers and therefore currently have total editorial control over the information they “publish” in their non-sponsored search results. They can and do use site-unique bias and outright censoring of search results to impose editorial policies. Other forms of bias including political bias are technically easy to implement and may exist. Disclosure of editorial policies is not required. A single entity controlling such an important information source as world-wide high-quality search is, at the very least, disturbing. Special anti-trust issues surround businesses that control the distribution of information.

[Search Engine Honesty](#)

Copyright 2007 Azinet LLC