

QUT Digital Repository:
<http://eprints.qut.edu.au/>



Spink, Amanda H. and Jansen, Bernard J. (2007) Commerce Related Web Search : Current Trends. In *Proceedings 18th Australasian Conference on Information Systems (ACIS2007)*, Toowoomba, Australia.

© Copyright 2007 Amanda H. Spink and Bernard J. Jansen

Commerce Related Web Search: Current Trends

Amanda Spink
Faculty of Information Technology
Queensland University of Technology
Brisbane QLD 4001 Australia
Email: ah.spink@qut.edu.au

Bernard J. Jansen
College of Information Sciences and Technology
The Pennsylvania State University
State College, PA
Email: jjansen@ist.psu.edu

Abstract

The level of commerce related Web search is a major indicator of the growth of e-commerce. Web queries are a key element in how people purchase services and products via the Web. People create Web queries in order to translate their information problems into a demand for products and services. Our paper provides results from studies analyzing Web search logs to determine the level of commerce related searches on commercial Web search engines for 2005 and 2006. Within commerce, we include business, e-commerce, travel, employment or economic related Web queries. Results show that commerce related queries currently represented between 28% and 30% of Web queries. The commerce related terminology on the Web is incredibly various and often difficult for the average user to predict with total accuracy. The implications of the findings and further research are discussed.

Keywords

Web search engines, commerce, queries, Dogpile

Introduction

Web searching is now a daily activity for many people. Web traffic has increased exponentially as people are using Web search engines as a major tool to dig their way through Web-based information. From the broad perspective, large-scale, quantitative or qualitative studies have explored how users' search the Web (Spink & Jansen 2004). To support information behaviors, we are seeing the development of a new generation of Web tools, such as Web meta-search engines, to help users persist in electronic information behavior is needed to help people resolve their information problems. Studies show that in general most Web queries are short, without much refinement or modification, and are simple in structure with few words (Spink & Jansen 2004). Few queries incorporate advance search techniques, and when such techniques are used many mistakes result. Frequently, people retrieve a large number of Web sites, but view few result pages and tend not to browse beyond the first or second results pages (Spink & Jansen 2004). Overall, a small number of terms are used with high frequency and many terms are used once. Web queries are very rich in subject diversity and some unique.

In addition, the Web has become a major source for all kinds of commerce related information, including business, e-commerce, travel, employment and economics. Since the inception of commerce related information on the Web, and easy and relatively less costly access to Web services, customers are using the Web has become a transaction medium for e-commerce purchases and commerce related information (Zwass, 2000). Web searching services are now everyday tools for providing commerce related information. Web search engine use is edging towards email usage as the primary Internet application (Raine, 2005). In the midst of a growing Web, what proportion of Web queries are commerce or related? Has the increase in commerce related Website content led to an increase in commerce related queries on Web search engines? What are the trends in commerce related Web searching? In this paper we report results from major studies of Web searches on commercial Web search engines. In this paper we examine the level of Web queries that are commerce related to help address these questions. This is a significant area of research for Web companies, Web site developers and providers, e-commerce providers and people in general who search the Web.

Related Studies

E-commerce is a growing force in the world economy and the Web is becoming a major source of business products, services and information for many people worldwide (Koufaris et al. 2001/2002). E-commerce is the use of electronic means and technology to conduct commerce, within-business, business-to-business and business-to-consumer (Chen, Chen & Shao 2003). E-commerce involves sharing information, developing and sustaining business ties and monetary transactions across telecommunications networks. Estimates suggest that e-commerce was worth \$1 trillion by 2002 (E-Commerce Times 2002). Major digital retail sites are continuing to develop and build brands, and creating market spaces for information, goods and services. Studies have shown that by 1999 most Web content was business-related (Lawrence & Giles 1999) with 83% of Web servers contain commercial content. Many studies have identified the factors affecting Internet sales (Phau & Poon 2000), the role of technology on Internet sales (Aberg & Shahmehri 2000), and the role of trust in e-commerce relationships (Hampton-Sosa & Koufaris 2005). Other studies have examined various aspects of business Web sites and the growth of online shopping (Li et al. 2006). However, limited studies have investigated the level of and trends in commerce-related Web searching. We have little idea how people are locating commerce Web sites, although we know that people use Web search engines to locate information on the Web. There is certainly great value in understanding how searchers express their commerce information needs.

Large-scale, quantitative or qualitative studies have explored how people are using of Web search engines (Spink & Jansen 2004). Most of these studies have focused on interactions with the non-sponsored Web search engine results. An increasingly important, popular, and uniquely contextual form of information interaction on the Web is sponsored search (Jansen, 2006). These results are not obtained via a crawl of the Web. With sponsored search, content providers pay Web search engines to display these sponsored links in response to user queries alongside the algorithmic links, also known as organic or non-sponsored links. This mechanism plays a critical role in financing the free search upon which so many users now rely. Web studies on sponsored search are just appearing in research publications. Jansen and Resnick (2005) report results of an investigation into the effect of sponsored links on e-commerce information seeking. In their research, 56 participants each engaged in 6 e-commerce Web searching tasks using 60 organic and 30 sponsored Web links for each task. Jansen and Resnick (2005) then studied the perceptions of sponsored links and the factors that influence this bias. Findings showed that there is a strong preference for organic links, a bias against sponsored results, and that more than 56% of the time, the title of the sponsored link was the determining factor in searcher perceived relevance. In related research, Jansen and Molina (2006) examined e-commerce Web searching effectiveness at the individual customer level by focusing on the success of everyday e-commerce searches. Searchers' found that the non-sponsored links were slightly more relevant than the sponsored links. Despite the growing research literature on business/e-commerce Websites, limited studies have analyzed levels and trends in commerce related Web searching. The next section of the paper outlines the research questions explored in our study followed by the data analysis and discussion of the implications of the key findings.

Research Questions

The goal of our analysis is to examine the current state of commerce related Web searching on Web search engines. Specifically we examine the following research questions:

1. What proportion of current Web searches are commerce related?

Research Design

We examined two Dogpile datasets of over 1.5 million queries from 2005 and 2006 to analyse the level of commerce related Web queries. Few published studies are available that have classified commerce related Web queries directly from Google, Microsoft (MSN), America Online or Yahoo! These nine studies are those large-scale studies that have been able to be conducted using large-scale Web logs directly from Web search engine companies. As Spink and Jansen (2004) state, access to large-scale Web logs for academic research analysis has been quite restricted for a number of reasons. Many Web search engine companies have been reluctant to provide large-scale Web logs for academic research. Spink and Jansen (2004) have been fortunate to obtain various large-scale Web logs and provide Web logs (where legally possible) to the Web research academic community.

For each Web query data set analyzed over the studies for 2005 and 2006, a random sample of approximately 2,500 queries was qualitatively classified by the researchers into one of eleven categories shown in Table 1 and Table 2. Examples of queries in the commerce category include, product and company names such as "coca cola", "Walmart" and "Rolex". Travel related queries included, "railway trains in Canada", "beach vacation Bahamas" and "cheap flights". Employment related queries included, "jobs in Miami", "how to join the army"

and “child care positions”. Economic related queries may include finance queries, such as “financial planning”, “government bonds” and “stock market falls”.

Results

Table 1 and Table 2 show the distribution of Dogpile queries topics for 2005 and 2006.

Table 1: Distribution of Dogpile Web query topic categories for 2005

2005 Dogpile (2500 Queries)	%
People, Places or Things	16.0%
Commerce, Travel, Employment, or Economy	30.4%
Computers or Internet	5.7%
Health or Sciences	8.9%
Education or Humanities	5.6%
Entertainment or Recreation	7.0%
Sex and Pornography	3.8%
Society, Culture, Ethnicity, or Religion	4.7%
Government or Legal	3.65
Performing or Fine Arts	0.5%
Non-English or Unknown	13.2%
	100.0%

Table 2: Distribution of Dogpile Web query topic categories for 2005

2006 (2,500 Dogpile.com Queries)	%
Commerce, travel, employment or economy	28.76%
Entertainment or recreation	16.40%
Sex or pornography	13.12%
People, places or things	10.52%
Health or sciences	8.68%
Education or humanities	5.12%
Society, culture, ethnicity or religion	5.08%
Computers or Internet	4.96%
Unknown or Other	4.44%
Government or legal	2.56%
Arts	.36%
	100%

The classification scheme was developed over various studies of Web searching by Spink and Jansen (2004). The category related to commerce is listed in Table 1 and Table 2 as “Commerce, Travel, Employment or Economy”. Queries classified as commerce were identified and classified as related to business, economics, travel, employment, e-commerce or commerce. Table 1 shows that people use Web search engines to find information on a broad range of information tasks. There has been a shift from 1997 to 2005 towards the use of the Web as a commerce information resource and tool. Commerce related queries increased as a proportion of Web queries from 13.3% in 1997 to now the most frequently occurring category (30.4%), followed by people, places and things, and unknown queries (indiscernible or non-English). The temporal findings across all data sets support the continued increase in commerce as a major topic for Web search engine users. From 1997 to 2005, Web queries related to commerce, computers and people increased as a proportion of all Web queries. Recently, Koshman et al, (2006) found that one in five queries submitted to Vivisimo related to commerce, travel, employment or the economy. This represents a sizable proportion of all queries. In addition, one in seven queries were related to people, places or things. These queries include personal names or the names of locations. The increase in the proportion of the commerce related Web searches could be related to many issues, including the increase in the availability of commerce related Websites.

Discussion

Many Web search engines are currently trying to cater to the broad market and some are catering to more specialized market niches. Niche and specialized Web search engines, such as Google Scholar [<http://www.google.com>] and Kosmix [<http://www.kosmix.com>] are appearing. Within the growth of the Web from 1997, the percent of commerce related queries has been cyclical. People are increasingly searching for commerce related information using Web search engines. The nature of their commerce related information needs are broad and relate to all aspects of commerce. Company (and industry) related queries account for some commerce related Web searches. Some people are searching for company Web sites. A need exists for new Web search tools that more easily facilitate company and product searches. Some people appeared to be searching for shopping sites via Web search engines as some commerce related queries are shopping-related. They are searching under brands names as well as generic types of products, such as "cars" or "golf clubs" (Spink and Jansen, 2004).

The great majority of Web commerce queries are short, infrequently modified and simple in structure (Jansen, et al., 2007; Spink and Jansen, 2004). Few queries include advance search features or Boolean logic. Many queries contain spelling mistakes that lead to failed searches. Despite getting many Web sites in response to their queries, Web users for commerce information view few result pages and do not generally view beyond the first or second page of results (Spink and Jansen, 2004). Overall, commerce related queries are slightly longer than general queries (Spink & Jansen 2004). People are also increasingly querying for commerce related information in question format. The development of Web services that focus on helping people with commerce related information needs in question and answer format is an emerging next step in improving the Web as an e-commerce tool. There is a need for niche Web services and search engines that facilitate more effective and targeted access to commerce sites. As the Web continues to become a key centerpiece for commerce, new tools and new ways of searching for commerce related information on the Web are developing. Searching the Web today is still frustrating for many users. The Web also lacks a standardized approach to search engine functionality and terminology. The commerce related terminology on the Web is also incredibly various and difficult for the average user to predict with total accuracy.

The results of our study have implications for practice. Commercial Web site developers need to be aware of how people search for commerce related information and trends in commercial Web searching. The growth of commerce related Web queries has implications for the practice of commerce via the Web. People are increasingly seeking commerce related information and transactions via Web search engines. The Web search engine companies understand this is occurring and are developing their commerce related Web search capabilities. Sponsored links have emerged as major financial drivers for Web search engine companies (Jansen & Spink 2007). The result of our studies has implications for the development of human computer interaction and information behavior models that seek to represent users' interactions with Web search engines (Spink & Jansen 2004). People's interaction with Web search engines is driven by many different human needs, including information behaviors and purchasing behaviors.

Conclusion

Further research is needed to track the trends in commerce related Web searching. We are conducting further studies on additional Dogpile Web query datasets to examine trends in commerce related searching. The authors are conducting ongoing studies into many aspects of commerce related Web searching using data from Dogpile.com.

References

- Aberg, J & Shahmehri, N 2001, 'The Role of Human Web Assistants in E-Commerce: An Analysis and Usability Study', *Internet Research: Electronic Networking Applications and Policies*, vol. 10, no. 2, pp. 114-25.
- Chen, M, Chen ANK & Shao, BBM 2003, 'The Implications and Impacts of Web Services to Electronic Commerce Research and Practices', *Journal of Electronic Commerce Research*, vol. 4, no. 4, pp. 128-39.
- E-Commerce Times.com 2002, 'Study: E-Commerce To Top \$1 Trillion in 2002', [<http://www.ecommercetimes.com/story/16314.html>]
- Hampton-Sosa, W & Koufaris, M 2005, 'The Effect of Web Site Perceptions on Initial Trust in the Owner Company', *International Journal of Electronic Commerce*, vol. 10, no. 1.
- Jansen, B J 2006, 'Paid Search', *IEEE Computer*, vol. 39, no. 7, pp. 88-90.

- Jansen, BJ & Molina, P 2006, 'The Effectiveness of Web Search Engines for Retrieving Relevant Ecommerce Links', *Information Processing & Management*, vol. 42, no. 4, pp. 1075-98.
- Jansen, BJ & Resnick, M 2005, 'Examining Searcher Perceptions of and Interactions with Sponsored Results', paper presented at the Workshop on Sponsored Search Auctions. Sixth ACM Conference on Electronic Commerce (EC'05), Vancouver, Canada, June 5-8.
- Jansen, BJ, Spink, A, Blakely, C & Koshman, S 2007, 'Web Searcher Interaction with the Dogpile Web Search Engine', *Journal of the American Society for Information Science and Technology*, vol. 58, no. 5, pp. 744-55.
- Koshman, S, Spink, A & Jansen, BJ 2006, 'Web Searching on the Vivisimo Search Engine,' *Journal of the American Society for Information Science and Technology*, vol. 57. no.14, pp. 1875-87.
- Koufaris, M, Kambil, A & Labarbera, PA (Winter 2001/2002), 'Consumer Behavior in Web-Based Commerce: An Empirical Study', *International Journal of Electronic Commerce*, vol. 6, no. 2.
- Lawrence, S & Giles, CL 1999, 'Accessibility and Distribution of Information on the Web', *Natur.*, vol. 400, pp. 107-9.
- Li, D, Browne, GJ & Wetherbe, JC 2006, 'Why Do Internet Users Stick with a Specific Web Site? A Relationship Perspective', *International Journal of Electronic Commerce*, vol. 10, no. 4, pp. 105.
- Phau, I & Poon, SM 2000, 'Factors Affecting the Types of Products and Services Purchased Over the Internet,' *Internet Research Electronic Networking Applications and Policies*, vol. 10, no. 2, pp. 102-13.
- Rainie, L 2005, 'Search Engine Use Shoots Up in the Past Year and Edges Towards Email as the Primary Internet Application', *Pew Internet and American Life Project* [http://www.pewInternet.org/PPF/r/167/report_display.asp].
- Spink, A & Jansen, BJ 2004, *Web Search: Public Searching of the Web*, Dordrecht, Springer.
- Zwass, V 2000, 'Structure and Macro-Level Impacts of Electronic Commerce: From Technological Infrastructure to Electronic Marketplaces', *Foundations of Information Systems - E-Commerce Paper* [<http://www.mhhe.com/business/mis/zwass/ecpaper.html>].

Copyright

Amanda Spink and Bernard J. Jansen © 2007. The authors assign to ACIS and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to ACIS to publish this document in full in the Conference Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.