

# Popular Searches in Google and Yahoo!: A “Digital Divide” in Information Uses?

**Elad Segev**

*Department of Communication, Hebrew University of Jerusalem, Jerusalem, Israel*

**Niv Ahituv**

*CENTRUM Católica, Pontificia Universidad Católica del, Santiago de Surco, Lima, Peru*

---

**In this article, the authors analyze the popular search queries used in Google and Yahoo! over a 24-month period, January 2004–December 2005. They develop and employ a new methodology and metrics to examine and assess the digital divide in information uses, looking at the extent of political searches and their accuracy and variety. The findings indicate that some countries, particularly Germany, Russia, and Ireland, display greater accuracy of search terms, diversity of information uses, and sociopolitical concern. Also, in many English-speaking and Western countries most popular searches were about entertainment, implying a certain gap within these countries between the few who search for economic and political information and the many who do not.**

---

**Keywords** classification of search queries, digital divide, entertainment searches, Google Zeitgeist, Internet search, political and economic searches, online information uses, Yahoo! Buzz

---

Received 7 August 2008; accepted 15 September 2009.

We would like to convey our gratitude and appreciation to Professor Costas Constantinou for his careful reading and useful comments and to the Research Institute for Law, Politics and Justice at Keele University for supporting this research. A special thanks to Doctor Elizabeth Carter for reading and suggesting ways to improve the article, to Professor Gadi Wolfsfeld for his useful ideas, to Marion Lupu and John Tresman for their proofreading of the article, and to the editor-in-chief and reviewers of *The Information Society* for their very constructive comments. Finally, a very special thanks to Regula Miesch for her encouragement and support all the way.

Address correspondence to Elad Segev, PO Box 4092, Zichron Yaaqov 30900, Israel. E-mail: info@eladsegev.com; Web site: <http://www.eladsegev.com>

The term *digital divide* refers to a variety of phenomena, such as the gap and inequality in accessing online information, the capacity and skills of information and communication technologies (ICT) use, the technical quality and width of the network, the governmental and social investment for online infrastructure and education, the overall ability to translate and evaluate information, and the social diversity of users (Anderson et al. 1995). The digital divide is particularly problematic in the case of the Internet because it is critical infrastructure that enables individuals to find jobs, acquire education, access governmental information, and participate in political panels and support groups. People who are unable or less able to access the Internet therefore have fewer political, economic, and social opportunities and find themselves in a disadvantageous position.

Moreover, the interactive and complex properties of the Internet enable those who possess better skills, education, and multimedia literacy to retrieve more relevant and useful information, which can be translated into social, economic, and political advantages. Therefore, provision of access to technology by itself will not bridge the divide. Warschauer (2004) gives various examples of unsuccessful governmental projects in India, Ireland, and Egypt that attempted to decrease the digital divide primarily through massive investments in information technology infrastructure. Instead of access, he suggests encouraging meaningful access, which also includes information use, literacy and education, and community and institutional structures. In particular, he proposes encouragement of community involvement and production of local content and applications in the fields of politics, economy, health, education, and local news. Similarly, a more recent study entitled E-government for Low Socioeconomic Status (E-LOST 2008) sponsored by the European Union (EU) investigated the digital divide in the use of e-government services by

citizens of low socioeconomic groups. The findings, based on survey and comparative analysis of six countries, support the concern that the transfer of governmental services to the Internet may widen the digital divide. The report suggests a number of actions to alleviate the problem, including attitude change, training programs, raising more awareness, and new access modes.

DiMaggio, Hargittai, Celeste, and Shafer (2004) also emphasize the need to go beyond physical access to information use, identifying five forms of information inequality: inequality of technical means (e.g., hardware, software and connection), the extent of autonomy in using the Web (e.g., monitored or limited use), inequality of information skills (e.g., knowledge of the interface, software, and hardware), inequality of social support, and, finally, the different purposes of information uses. In terms of the latter, they found greater socioeconomic differences between users who searched for health-, politics-, and employment-related information and users who searched for entertainment-related information. Similarly, Robinson, DiMaggio, and Hargittai (2003) found that college-educated online users possess clear advantages over high school-educated online users. They suggested that the former used the Internet much more to search for jobs, health, education, and other economic and political purposes. Their conclusion was that the digital divide is further widened by differences in the ability to use online information, especially for political and economic purposes.

Two important dimensions of the digital divide emerge from these studies: volume of and control over information uses. Whereas *volume* refers to the variety of information available, *control* refers to the ability to extract relevant information and use it skillfully. On the same line, Bonfadelli (2002) suggests that the digital divide should be assessed on the basis of the ability of skilled users to retrieve deeper and more valuable information. Skilled users can retrieve a relatively high volume and variety of information and, at the same time, customize, exploit, and control it successfully. To that end, search engines have an increasingly important role in empowering skilled users, as they cover an extensive amount of information (highest volume), and enable online users to search, organize, customize, and retrieve the most desirable and relevant information (highest control). Together with content analysis (i.e., the extent of political and economic searches), the dimensions of volume and control will be measured in this study.

The digital divide in information uses therefore has important implications for the structure, hierarchy, and future of the information society. Although the Internet was initially thought to provide equal opportunities and freedom of information for all (Rheingold 1993; Negroponte 1995), there is growing empirical evidence indicating the strengthening of the strata of the population possessing

more resources and information skills (Hargittai 2000; 2003; Norris 2001; DiMaggio et al. 2001; Ciolek 2003; Castells 2004; Rogers 2004).

Pippa Norris (2001) specifies three types of digital divide: the *global divide*, which is the difference among different countries; the *social divide*, which is the difference among diverse social groups; and the *democratic divide*, which refers to the different applications and uses of online information to engage, mobilize, and participate in public life. In this article, we focus mainly on the global divide and the democratic divide—in other words, the democratic divide on the global level. It therefore deals less with the inequality of access and more with the inequality of uses and skills across countries.

## THE METHODOLOGY FOR ANALYZING SEARCH QUERIES

The object of our analysis is the most frequent search queries initiated by users in different countries, mainly in Google, but also in Yahoo! (for information uses in the United States, see later in this article). The analysis of search queries can provide an important insight into popular information trends in different countries. Several studies explore how people search the Web (e.g., Silverstein et al. 1999; Wolfram et al. 2001). Bar-Ilan (2004) and Jansen and Spink (2004) suggest that Web-searching studies fall into three categories: (1) those that examine search queries, (2) those that incorporate user surveys and observations, and (3) those that examine issues related to or influencing Web searching (e.g., Web structure, interface design, social and environmental conditions).

Hargittai (2002) attempts to assess the digital divide of information uses by providing a random sample of users with a list of information search tasks. Individuals were asked to find information about local cultural events, political candidates, tax forms, and so on. Information skills were defined as the ability to find the desired information and by the time required. The findings indicated that young and experienced users were more likely to succeed in completing the tasks quickly, whereas old users and newcomers were much slower and sometimes could not complete all tasks.

Other studies (e.g., Silverstein et al. 1999; Jansen et al. 2000; Jansen and Spink 2003) focus on the search queries people use, but not from a digital divide perspective. Jansen and Spink (2004) conducted a longitudinal study from 1997 to 2003, looking at search queries in Excite, Alta Vista, Ask Jeeves, and alltheweb.com to explore how and what people search on the Web in Europe and the United States. Their studies included pornography-, health-, and business-related search queries, but did not look at the political, social, or cultural implications of the searches.

In terms of query length and search session length, Jansen and Spink (2003) indicated very little changed over the years, with most users entering two to three terms per query and viewing about five Web documents per query, without query reformulation or modification. In the distribution of search queries, Jansen et al. (2000) assume the existence of the power law; that is, a few terms were used repeatedly and many terms were used only once. This further increases the importance of studying popular search queries in Google Zeitgeist (see subsequent discussion), which can shed light on information-searching habits of many users worldwide.

In another study, Spink, Jansen, Wolfram, and Saracevic (2002, 108) attempted to classify search queries from the Excite search engine into eleven “non-mutually exclusive, general topic categories,” such as “entertainment or recreation,” “health or sciences,” “commerce, travel, employment, or economy,” and “people, places, or things.” However, the reasoning behind their taxonomic system is unclear, and the categories themselves seem to overlap (e.g., entertainment and people).

Chau, Fang, and Yang (2007) compare popular Chinese search queries in a Hong Kong-based search engine with those in English search engines for content variety, query length, and the use of search operators. Their study indicated similarity to English search engines in search topics and the average query length. Apart from pornography-related queries among the top 100 search queries, there were many queries related to travel, e-commerce, and music downloading.

Finally, Ross and Wolfram (2000) analyze popular queries from the Excite search engine, identifying various topics using cluster analysis. Similarly, Pu, Chuang, and Yang (2001) classify popular search queries in three Taiwanese search engines. Both studies attempt to apply automatic systems to classify popular search queries into topical categories, resulting in several well-defined clusters of subjects. Their logic was to obtain highly ranked Web documents based on each search query, and then to analyze the content of these documents, and to identify their main topics.

Previous studies of search queries suggested methodologies for manual or automatic divide. This study continues the investigation of online search, attempting to shed light on the digital divide of information uses by analyzing the content, diversity, and accuracy of search queries. It also makes a contribution by developing a crossnational comparison of popular searches in a relatively large number of countries, a comparison that has not been done in previous analyses of search queries.<sup>1</sup>

We developed three indices to examine three different aspects of the digital divide in information uses: the Economic and Political Value (EPV) of search queries, the Variety of Uses (VoU),<sup>2</sup> and the Specificity of Search (SoS).

Observations were made over a twenty-four-month period, from January 2004 to December 2005. The relationships between those indices are examined, and subsequently countries are clustered based on the different attributes of the searches. The implications and limitations of this study are discussed, calling for further development and implementation of search query databases and new analytical tools to study the digital divide in information uses.

## DATA SOURCES

Most data have been automatically gathered and published in the Google Zeitgeist Web site. The term *Zeitgeist* is commonly attributed to J. G. Herder’s German translation of the Latin expression “genius seculi,” referring to the spirit of the century (Barnard 2004). The drawback of Google Zeitgeist is that it does not regularly provide data on popular searches in the United States.<sup>3</sup> However, Yahoo! also provides a weekly summary of the most popular search queries in general (also known as Yahoo! Buzz). Yahoo! does not divide information uses by countries, and therefore may provide a more global perspective on information use. However, it is estimated by several sources that the highest share of users who search in Yahoo.com are by far American,<sup>4</sup> and thus the data provided by Yahoo! gave an indication of information trends in the United States.<sup>5</sup>

Jansen and Spink (2004) make a similar attempt to compare the use of information in various countries by looking at popular search queries in several search engines (i.e., Fireball, a predominantly German Web search engine; BWIE, a Spanish Web search service; and Excite, a U.S.-based Web search engine). In another study, Spink, Ozmutlu, Ozmutlu, and Jansen (2002) examine search queries during 2001 to FAST (also known as alltheweb.com), which was largely used by Europeans at that time. Popular queries to FAST were compared with those to Excite (used mainly by American users), suggesting that FAST’s users searched more for people and places, whereas Excite’s users focused on e-commerce. With regard to methodology, these studies suggest that there are some functions, such as the content of search queries and time of search sessions, that are comparable across different search engines in different countries. However, the comparison of some interface-dependent functions, such as the use of search operators, is less straightforward.

In this study the comparison between Google’s and Yahoo!’s popular search queries was in terms of content, variety, and accuracy, which are not interface-dependent functions and can therefore be compared. A possible methodological risk in such a comparison is that some people may use both Google and Yahoo!, but for different purposes (e.g., Google for information-seeking, and Yahoo! for entertainment purposes). This, however, is

very unlikely, because both Google and Yahoo! are general rather than niche search engines, consisting of very similar functions.<sup>6</sup> Moreover, the data used in our analysis consist of the most popular search queries, which are almost always general, rather than specific, queries. In any case, Google Zeitgeist displayed popular search queries in Google.com for some months. These data were used to validate the results of Yahoo!, and clearly confirmed and supported the results, which showed that the most popular search queries to the parent sites were about entertainment.<sup>7</sup>

In January 2004, Google Zeitgeist displayed the most popular search queries in nine countries: the United Kingdom, Canada, Germany, Spain, France, Italy, the Netherlands, Australia, and Japan. Another seven countries, Brazil, China, Denmark, Finland, South Korea, Norway, and Sweden, were added to the report in July 2004. Finally, four more countries, Ireland, India, New Zealand, and Russia, were included in January 2005 (see appendix B for the complete list of countries and the dates of inclusion in Google Zeitgeist). In sum, this study exploits data from Google's archive on twenty countries,<sup>8</sup> and from Yahoo!'s archive on general information searches, which mainly refers to the United States.

It is possible for users in one country to connect to search engines in other countries. For example, users in Germany could search with Google.de, but also with the parent U.S. site, Google.com, or the French version, Google.fr. The monthly report of Google Zeitgeist shows the most popular search queries used to search in Google's national interfaces. Thus, for example, search queries that were counted for Google Germany are those that were used to search in Google.de. This does not necessarily imply that all the users who searched Google.de reside in Germany. They could theoretically be in China, in the United States, or anywhere else. However, it does mean that users who searched in Google.de, and made up the monthly statistics of popular search queries, were most likely to be familiar with the German language, since the interface of Google.de is in German. This means that language is an important factor in the analysis.

Nonetheless, customization mechanisms in Google also promote and reinforce local and national factors. Google automatically recognizes the IP number of its users, and thus the location from which they search. Subsequently, it automatically loads the interface that is appropriate to their country by default. This "user-friendly" process considerably increases the probability of local users employing the national interface of the country from which they search. It is therefore reasonable to assume that Google Zeitgeist broadly presents popular information searches of people in different countries speaking different languages.

This study analyzes between 150 and 200 popular search queries from each of the national interfaces of

Google during the study period, from January 2004 to December 2005.<sup>9</sup> All together, 4,474 different search queries were analyzed. In response to our inquiries,<sup>10</sup> Google has indicated that data in Google Zeitgeist were compiled using the internal version of more recent tools (i.e., Google Trends and Insights for Search). This list reflects the most popular searches in Google Search, excluding porn-related queries, duplicate entries (including misspellings), and spam results.<sup>11</sup> Obviously, many search engine companies hesitate to share a large volume of search queries or reveal the processes behind the data collection and reporting in their different services.<sup>12</sup> Nevertheless, the longitudinal investigation enabled gathering of a relatively large volume of data for a crossnational comparison, which was very instrumental in the use and demonstration of new methodologies for studying the digital divide in information use.

## A CROSSNATIONAL COMPARISON

There are two main reasons for conducting a cross-national comparison in this study. First, as previously mentioned, literature on the digital divide often describes the technology and information differences between states, as well as within each state (Norris 2001). On the most basic level, there is a technical divide, which refers to the Internet infrastructure, and the physical differences in access to the network in terms of equipment, Internet service providers (ISP), costs, and so on. Furthermore, different countries have different economic power, which is strongly related to the different percentages of online population (see also table A4). Then there are also political differences between countries. Democratic regimes allow access to most Web sites in the Web, whereas undemocratic regimes impose censorship and restrict access to certain Web sites for political, cultural, and social reasons. These restrictions and limitations are mostly exercised at the national level. Finally, the digital divide in information use is also a result of the different knowledge of languages. Most Web sites provide information in English. The official national language and learned second languages are a direct result of national policies, and thus it is more likely that national division also affects language division, and both have crucial implications on the digital divide of information uses. Subsequently, our analysis uses countries as a unit of comparison in the study of the digital divide.

The second reason for conducting a crossnational comparison is the nature of the data in Google Zeitgeist, which also divides search queries by country. It is therefore a straightforward process to exploit these data, analyze the different popular information searches in different countries, and discuss their implications for the digital divide. The methodology developed in this article could be

also applied in future studies to explore the digital divide within states, looking at popular search queries of users of different age, ethnicity, region, and the like.

### MAIN CLASSIFICATION SYSTEM

To compare information uses in different countries, this study employed a classification system of content integrated in Google Search, called the Open Directory Project (ODP), which is a directory developed and constantly updated by the online community. Each editor qualified to add to and maintain the open directory is chosen on the basis of knowledge of the language, the culture, and the field of the category to be edited. New information appearing on the Web is constantly classified by the network community itself, and categories and subcategories are added and edited through a system of checks and balances and quality assurance. The Open Directory powers the core directory services in Google, AOL/Netscape Search, and many other large and popular search engines and portals (DMOZ 2005).

One of the main principles that ODP editors are required to follow is to organize Web sites by topics (e.g., news, business, and games), rather than simply by region. This principle works well with the concept of functionality and usability of information, and refers directly to the research problem of this study. Furthermore, there are two main advantages in exploiting the ODP classification system in this study. First, content has already been classified, which means consistency and accuracy of the classification process. Different coders who use the ODP classification system will always attain similar results. The second advantage is that the ODP enterprise is international, and its editors are local. It therefore already contains wide knowledge and experience, and provides an expert-specific classification of content by culture and language. Because ODP editors are required to have the cultural, language, and even topical background of the category they manage, it is reasonable to assume that they classify and sort information more accurately than people who do not know the field, the language, or the cultural context of the classified content. The ODP central management, the hierarchical structure of editors, and the developed system of checks and balances ensure consistency and accuracy of classification, even when done by different editors.

Google Web Directory, based on the ODP, provides fourteen different topical categories. For each category there are between one and seventeen subcategories, which are divided again into third-level subcategories, and so on. The main categories of Google Web Directory are: arts (with twelve subcategories, including movies, music, and television), business (with eight subcategories, including employment, financial services, and investing),

computers (with seven subcategories, including hardware, the Internet, and programming), games (with six subcategories, including gambling, role-playing games, and video games), health (with four subcategories, including alternative, beauty, and nutrition), home (with four subcategories, including do-it-yourself, cooking, and family), news (with four subcategories, including breaking news, online archives, and weather), recreation (with thirteen subcategories, including humor, outdoors, and travel), reference (with five subcategories, including education, dictionaries, and maps), science (with three subcategories, astronomy, technology, and earth sciences), shopping (with eight subcategories, including auctions, clothing, and flowers), society (with nine subcategories, including chats and forums, government and religion, and spirituality), and sports (with seventeen subcategories, including basketball, football and soccer). Appendix B displays the full list of categories and subcategories.

A search query submitted to the ODP or to Google Web Directory provides in return not only a list of results with their specific classifications, but also the main and most frequent classification of most results.<sup>13</sup> In this way, it was possible to ascribe the most frequent and common ODP classification for each popular search query automatically. Even though the process of classifying search queries into categories and subcategories was mostly automatic, there was careful human control for each query that involved checking the integrity, and filtering the regional effect of the classification process. Hence, for example, the query “herr der ringe” (in English: “The lord of the ring”) appeared in Google Germany in January 2004, and was automatically classified as World > Deutsch > Arts > Films > Titles > H, since the query was written in German. In this case, the researcher manually filtered the first regional categories and started with Arts > Films > Titles, as the three categories to be checked and compared. When a query was classified automatically as regional, the subcategories were used as the main classification to maintain integrity with results from other national interfaces. Another example is the query “eastenders” in Google UK. This query was classified automatically as Regional > Europe > United Kingdom > Arts > Television > Programs, and was counted only as Arts > Television > Programs, for research purposes. The only case when a query was classified in this research as regional was when the query itself was a region, like the query “france” in Google France in July 2004. Apart from exceptional and very rare regional queries, all search queries were classified first by their topic and usability, using the automatic subcategories suggested.

To reiterate, although the main method of classifying search queries was Google Web Directory, the automatic classification process of each query was also manually monitored to maintain the integrity of the results and to

filter the regional effect, resulting in a better comparative exercise.

### RELIABILITY OF CODING: THE HIDDEN INTENTION

Even knowing the search query, it is impossible to be completely sure what kind of information each individual user intended to acquire. However, it is possible to follow the main theme of each query and assign its relevant topic with a high degree of confidence. The classification process is based on the majority of search results and the subsequent open directory classification. Additionally, as mentioned earlier, the automatic classification was manually controlled by the researchers to keep its coherence with other results.

In most cases the classification was straightforward, because most queries were very simple and popular. For example, the query “Britney Spears” was classified as Arts > Music > Bands and Artists. However, in some specific cases the classification process was not as straightforward. For example, the query “heart” is very general, and could be classified in different ways by the ODP. By searching for the query “hjärtan” (which was actually one of the popular search queries in Google Sweden in February 2005 and translates into “heart” in English), not everyone intends to find the same kind of information. Some may refer to Society and Relationships and others to Health. It is therefore the duty of the coder to analyze the relevant results, to refer to data from other months in order to develop a comprehensive picture, and finally to decide what is the most common information retrieved, or reasonably intended to be retrieved, by using this query.<sup>14</sup> Similar to the query “hjärtan,” less than 1 percent of the search queries were too general or vague, leading to a classification process not being straightforward. Because those cases were very rare, it is unlikely that a mistake in an intelligent guess would have adversely affected the results.

### ECONOMIC AND POLITICAL VALUE INDEX

The construction of the first index was inspired by the argument that certain information skills and uses can empower individuals with political and economic advantages (Norris 2001; Bawden 2001; Ciolek 2003; Castells 2004; Rogers 2004; Webber 2000). A recent study of television audiences by Robert Putnam (2000) revealed that the more time people spend on watching news, the greater is their civic and social engagement. In contrast, the more time people spend on watching soap operas and game shows, the less is their civic and social engagement. This does not necessarily suggest a causal relationship (i.e., retrieving information related to politics leads to greater civic

engagement) but it clearly implies a link between the two. When it comes to the Internet, retrieving information about news, tax, law, government, society, or business provides users with economic and political knowledge. This information includes also, for example, new available positions, price comparison, education opportunities, political Web sites, fund-raising, and so on.

Looking at the digital divide among users, DiMaggio et al. (2004) compare the information skills, effectiveness, and productivity of information uses for economic and political purposes. They analyzed data from the 2000 and 2002 General Social Surveys, and subsequently distinguished between uses that are primarily recreational and uses that increase economic welfare (e.g., job seeking, consumer information, education), as well as political and social capital (e.g., following the news, searching for information on public and civic issues).

A recent study (DiMaggio and Hargittai 2002) has linked higher education and economic status with greater use of “capital-enhancing” information, which is financial, political, or governmental information. Similarly, a study by Bonfadelli (2002) found that higher education is positively associated with information and service retrieval, and negatively associated with using the Web for entertainment purposes. The framework developed by DiMaggio and Hargittai (2002) was implemented a year later in a study by Robinson et al. (2003) that linked the digital divide with the use of online information for various purposes. Their study indicated, for example, that higher education and income of users were associated with greater search for jobs, health, education, news and other economic and politically related information. In contrast, lower education and income of users were associated more with searches for entertainment, music, games, sports, and leisure activities. Subsequently, they concluded that the digital divide is deepened by different uses of online information, and particularly political and economic uses.

Following this distinction between the various information uses (see also Warschauer 2004; Howard et al. 2001; World Summit on the Information Society [WSIS] 2003), and in line with Norris’s (2001) distinction of the political divide, our study articulated the Economic and Political Value Index to examine the extent of search queries of high political and economic value (related to governmental information, news, jobs, business, etc.). It is impossible to infer how users will employ information from search queries. However, as previously indicated, there is a strong correlation between searching for political and economic information and greater information literacy and skills.<sup>15</sup> This does not suggest that entertainment-related information cannot empower users and provide them with certain advantages. This distinction was primarily meant to examine the inequality of economic and political opportunities between users worldwide. Thus, there is room

for further studies to examine the link between acquiring entertainment-related information and gaining social and emotional advantages.

Google Web Directory categorized search queries on movies, music bands, and celebrities as arts. Similarly, it categorized search queries on political, economic, and social affairs as news, business, and society, respectively. This division between entertainment and political, economic, and social affairs is crucial for this study, as it

weights were chosen to enable a simple comparison between countries and rank them according to the extent of political and economic value of their searches. This method is based on the abovementioned previous studies on the digital divide of information uses (e.g., DiMaggio et al. 2004; Howard et al. 2001; Robinson et al. 2003; Warschauer 2004). A high weight (e.g., 3) reflects a higher extent of political and economic value, whereas a low weight (e.g., 1) reflects a lower extent of political and economic value.

$$EPV_i = \frac{(\text{Number of high-level queries} \times 3) + (\text{Number of medium-level queries} \times 2) + \text{Number of low-level queries}}{3 \times \text{Total number of queries}} \quad [1]$$

provides an insight into the different national uses of information, and is directly linked with the distinctions made by the digital divide studies previously mentioned. Subsequently, three levels of political and economic value are defined. *High-level* categories refer to the search queries of high economic and political value (i.e., business, news, shopping,<sup>16</sup> and society—only the subcategories: issues, politics, government, organizations, and law). *Medium-level* categories refer to search queries of middle economic and political value (i.e., recreation, home, regional, reference, science, computers, health, and society—apart from the subcategories: issues, politics, government, organizations, and law). These categories are not directly related with economic and political uses, but are also not entirely related to entertainments. However, they have been mentioned by previous studies as more “capital-enhancing” information with a certain political or economic value.<sup>17</sup> And finally, *low-level* categories refer to entertainment-related search queries (i.e., arts, games, and sports), which have relatively lower economic and political value.

The EPV Index was constructed by assigning a weight to each of the suggested categories and subcategories of the search queries, depending on its level of political and economic value. Table 1 shows the relative weights ascribed to each of the categories in this study. Ordinal

The EPV Index of a certain country in a certain month is the number of queries from each EPV group multiplied by its weight and standardized to 1. Definition 1 provides a simple formula for calculating the EPV Index in a country each month.

$$VoU_i = \frac{\mu_i}{\sigma_i} \quad \begin{array}{l} \mu = \text{Average queries in a category} \\ \sigma = \text{Standard deviation of queries} \\ i = \text{For each country each month} \end{array} \quad [2]$$

The EPV Index can range from 0.33, which is the lowest extent of economics- and politics-related searches in a certain month, to 1, which is the highest extent of economics- and politics-related searches. When it is close to 1, it indicates that there are many queries of economic and political value among the popular queries in a certain month. For example, in November 2004 the EPV Index in Google France reached a record of 0.93, entailing mainly business-, news-, and shopping-related queries.<sup>18</sup>

## VARIETY OF USES

Some have argued earlier that skilled and creative users are empowered by their variety of information uses and their control over information. It is therefore suggested that countries with a greater variety of online information

**TABLE 1**  
Weight of economic and political value

Level	Categories	Weight
High-level categories	Business, news, shopping, and society (only the subcategories: issues, politics, government, organizations, and law)	3
Medium-level categories	Society (apart from the subcategories: issues, politics, government, organizations, and law), reference, science, computers, regional, home, health, and recreation	2
Low-level categories	Arts, games, and sports	1

uses may also experience a wider range of opportunities for benefiting from the Internet. Stated differently, they demonstrate a greater understanding of the Internet's potential, and are empowered by online information in various fields, such as economics, politics, education, society, business, and entertainment.

This view has been supported by recent studies. Bonfadelli (2002) notes that the variety of online information empowers skilled users and thus increases information inequalities. Similarly, the study by Robinson et al. (2003) revealed that a variety of information uses, and particularly the search for jobs, health, education, and other economics- and politics-related information, correlated highly with better education and indicated better information skills.

In our study, in January 2004 in the Netherlands, seven of the ten most-popular search queries were about arts and entertainment, two about sports, and one about society. It is clear that popular searches in the Netherlands in January 2004 were relatively homogeneous, and concentrated mainly on entertainment. In contrast, in the same month in Italy there were only four search queries about arts and entertainment, two about games, one about health, one about society, one about science, one about reference, and one about business. This implies that popular search queries in Italy in January 2004 were more diverse than those in the Netherlands. Again, it is impossible to infer from the search queries how users employ the information. However, as previously indicated, a variety of search topics suggests a better understanding of the various applications of online information and was found to be correlated with greater information literacy and skills.

The VoU Index was constructed to study this variety of search topics. It is based on the coefficient of variation (standard deviation/mean), which is a dimensionless number reflecting the spread of search queries among the categories.<sup>19</sup> The reciprocal of the coefficient of variation (mean/standard deviation) was calculated separately for each country in each month. The reciprocal was used because a smaller variation indicates an even spread of queries in each category, and therefore a greater variety of uses. For example, the reciprocal of the coefficient of variation in the Netherlands in January 2004 was 0.38, whereas in Italy it was 0.78, meaning that in January 2004, there was a greater variety of uses in Italy than in the Netherlands. Definition 2 provides a simple formula for calculating the VoU Index for each country each month.

The VoU Index can range from 0.27 to 1.52,<sup>20</sup> as 0.27 indicates that all search queries are related to one category in a country (e.g., Arts), and 1.52 means that information uses are very heterogeneous in a country in a particular month.

## SPECIFICITY OF SEARCH INDEX

Specific searches also mean specific results, which can further provide users with more relevant and immediate information. A focused and detailed search indicates better searching skills. If the VoU Index indicates how heterogeneous and rich information uses are in different countries, the SoS Index indicates how skilled and controlled information uses are in various countries. The SoS Index is therefore another way of assessing the digital divide of information uses.

Each search query can be classified into up to three categories and subcategories. For example, the search query "lord of the rings," which is a very specific one, was classified based on Google Web Directory into three categories and subcategories: Arts > Movies > Titles. The search query "games," which is a very general one, was classified into only one category: Games. Hence, the number of categories and subcategories can help assess whether a search query is more general or more specific. Definition 3 provides a simple formula for calculating the SoS Index in each country each month.

$$\text{SoS}_i = \frac{\text{Number of subcategories}}{3 \times \text{number of queries}} \quad [3]$$

The value of the SoS Index can range from 0.33 to 1.00; the former indicates relatively general search queries and the latter indicates relatively specific search queries.

## INITIAL PREDICTIONS

It was earlier suggested that online information uses derive from socioeconomic, political, and cultural differences between countries. Thus, it was expected that countries leading in economy and technology, such as the United States (in which online networks are well-established and the majority of the population has used the Internet for a relatively long time), would also display greater versatility and accuracy in their use of search queries. Similarly, since most online content is in English (UNESCO 2006; O'Neill, Lavoie, and Bennett 2003; Pastore 2000), it was expected that users from English-speaking countries would demonstrate a greater variety of uses.

## RESULTS AND ANALYSIS

### Economic and Political Value Index

Figure 1 summarizes the average of the EPV Indices over 2004 and 2005 for each country. Because Google Zeitgeist displays only 10–15 popular search queries for each



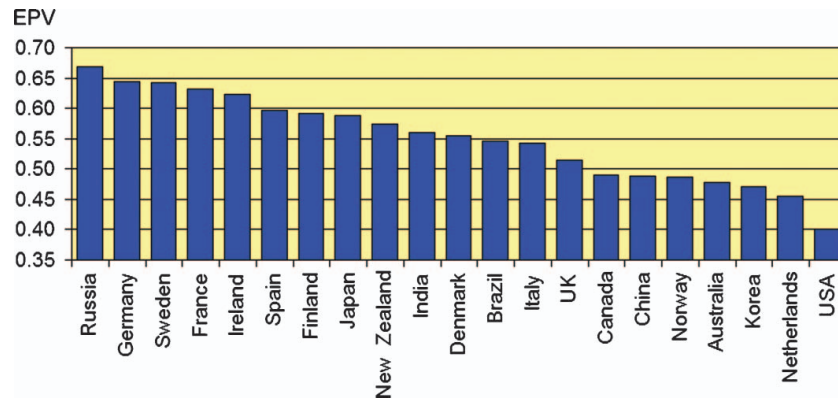


FIG. 1. Average EPV Index, 2004–5.

country each month, the average of EPV Indices over two years provides a more comprehensive estimation of the information trends in different countries.

The EPV Index indicates that Russia, Germany, Sweden, France, and Ireland have relatively more search queries of high economic and political value (score 0.62 and above). The data used to compute the EPV Index for the “General” search were taken from Yahoo!, and refer to the top search queries in Yahoo.com over 2004 and 2005. These data mostly reflect information trends in the United States, and display very few search queries of economic and political value. This is also true for the popular information uses in the Netherlands, South Korea, and Australia, all of which have a low EPV Index value (0.48 or lower). The findings suggest that Russia, Germany, Sweden, France, and Ireland demonstrate a relatively high extent of political and economic information uses, whereas the United States, Australia, the Netherlands, and South

Korea demonstrate a low extent of political and economic information uses. In the latter group, most popular search queries are about entertainment.

### Variety of Uses

Figure 2 summarizes the average of the VoU Indices for 2004 and 2005 for each country. Here again, the average of VoU Indices over two years provides a more comprehensive estimation of the information trends in different countries.

Figure 2 shows that, in terms of variety of online information uses, Spain was the leading country during 2004 and 2005, with a relatively high variety of uses, followed by Denmark, Sweden, and Ireland. Not surprisingly, countries like South Korea, the United States, Canada, and Australia, which scored low on the EPV Index because of the dominance of entertainment-related

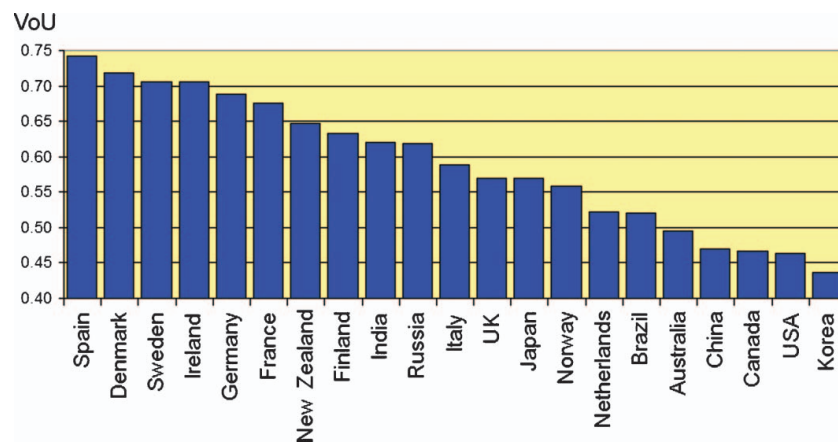


FIG. 2. Average VoU Index, 2004–5.

uses, also have the lowest VoU value, indicating that their uses of online information are more homogeneous. European countries such as Spain, Denmark, Sweden, Ireland, and Germany that reported medium and high EPV Index values also have a high VoU value, because they demonstrate a relatively even distribution of information uses.

### Specificity of Search Index

Figure 3 summarizes the SoS Indices over 2004 and 2005 for each country. Similarly, the average of SoS Indices over two years may provide a more comprehensive estimation of the information trends in different countries.

Figure 3 indicates that in South Korea, the United States, China, India, and Australia, search queries are relatively more specific. The average SoS Index value of 0.97–0.99 reveals that more than 90 percent of search queries in these countries are very specific and detailed. In contrast, in Sweden, Denmark, France, and Finland, search queries are relatively less specific. The average SoS Index value of 0.82–0.88 reveals that more than 40 percent of search queries in these countries are relatively general. These findings are especially interesting when compared with the VoU Index, for which the results were almost the reverse. Sweden, Denmark, and France were among the leading countries in terms of variety of information uses, while in terms of information specificity they lag behind. In contrast, countries such as Korea, the United States, and China had relatively homogeneous information uses but lead in terms of specificity. This is probably because in all of these countries the use of information for entertainment purposes is dominant, and people tend to search for more specific information, such as particular performers, music bands, television programs, and so on.

The fact that users employ the Internet mainly for entertainment purposes does not necessarily mean that they possess less information skill. The ability to use specific search terms to retrieve information more accurately and promptly is another important factor. It indicates that most users know exactly what to look for, and may further imply that online information is highly customized in these countries. Hence, the SoS Index reveals another aspect of the politics of online information: the ability to control and retrieve relevant information. The findings suggest that countries such as South Korea, Australia, and the United States, which display a low extent of economic and political searches and variety of information uses, have relatively more specific information uses. Although they do not exhibit a high variety of information uses, their searches are more focused, and therefore can yield more relevant and immediate results.

**TABLE 2**  
Summary of index rankings

Rank EPV	VoU	SoS
1. Russia (0.67)	Spain (0.74)	South Korea (0.99)
2. Germany (0.64)	Denmark (0.72)	India (0.98)
3. Sweden (0.64)	Sweden (0.71)	Australia (0.98)
4. France (0.63)	Ireland (0.71)	China (0.98)
5. Ireland (0.62)	Germany (0.69)	United States (0.98)
6. Spain (0.60)	France (0.68)	Ireland (0.97)
7. Finland (0.59)	New Zealand (0.65)	Canada (0.97)
8. Japan (0.59)	Finland (0.63)	Russia (0.96)
9. New Zealand (0.57)	Russia (0.62)	Italy (0.96)
10. India (0.56)	India (0.62)	Norway (0.95)
11. Denmark (0.55)	Italy (0.59)	Brazil (0.95)
12. Brazil (0.55)	Japan (0.57)	Germany (0.94)
13. Italy (0.54)	UK (0.57)	New Zealand (0.94)
14. UK (0.51)	Norway (0.56)	Netherlands (0.94)
15. Canada (0.49)	Brazil (0.52)	Japan (0.91)
16. China (0.49)	Netherlands (0.52)	United Kingdom (0.9)
17. Norway (0.49)	Australia (0.5)	Spain (0.89)
18. Australia (0.48)	China (0.47)	Finland (0.88)
19. South Korea (0.47)	Canada (0.47)	France (0.87)
20. Netherlands (0.45)	United States (0.46)	Denmark (0.85)
21. United States (0.40)	South Korea (0.44)	Sweden (0.83)

### Relationships Between Indices

Table 2 summarizes the rankings of the EPV, the VoU, and the SoS, indicating possible relationships between the indices.

In theory, very high scores of the EPV Index mean that most search queries are concentrated in economic- and political-related categories. Similarly, very low scores of the EPV Index mean that most search queries are concentrated in entertainment-related categories. In both extreme cases (of very high and low EPV scores) the VoU Index is supposed to be low, as the spread of search queries is not even among the different categories. Table 2 implies a possible positive correlation between the EPV and the VoU Indices. It indicates that countries with low EPV scores (e.g., the United States, Canada, Australia, Korea, and China) also have low VoU scores, whereas countries with high EPV scores (e.g., Sweden, Ireland, and Germany) usually have also high VoU scores. No countries in

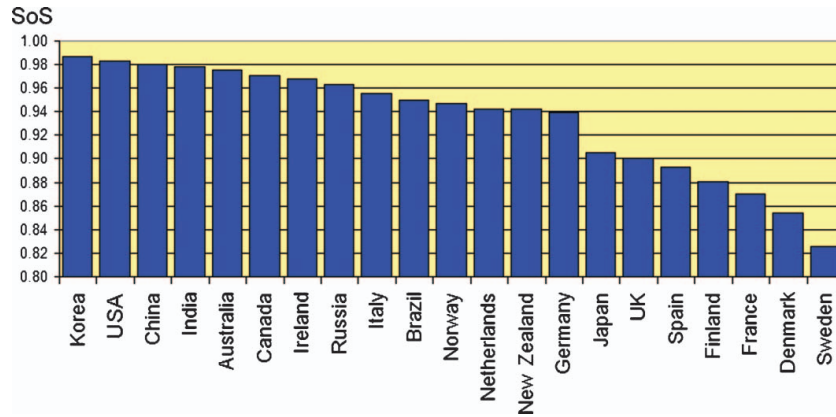


FIG. 3. Average SoS Index 2004-5.

table 2 have high EPV scores and low VoU scores. This is primarily because there are no countries with a very high concentration of economic- and political-related searches. The countries with the highest EPV scores (e.g., Russia, Germany, Sweden, France and Ireland) have still between 20 to 40 percent of entertainment-related searches, and thus display a greater variety of searches than other countries (i.e., greater VoU scores).

Because a positive correlation between the two indices is expected, and there are no assumptions regarding their distribution, a Spearman<sup>21</sup> one-tailed correlation test confirms that the EPV Index and the VoU Index have a strong

positive correlation (0.81,  $p < .01$ ). A combination of the two correlated indices in one graph presents the differences between countries in terms of the content and the variety of searches (see figure 4).

While the EPV Index reflects the content aspect, the VoU and SoS Indices reflect another two other aspects of the digital divide in information uses: volume and control. Table 2 implies that many countries that scored highly on the VoU Index (e.g., Sweden, Denmark, Spain, and France) tend to have low SoS Index scores. Similarly, countries with low VoU scores (e.g., United States, Canada, Korea, and China) tend to have high SoS scores.

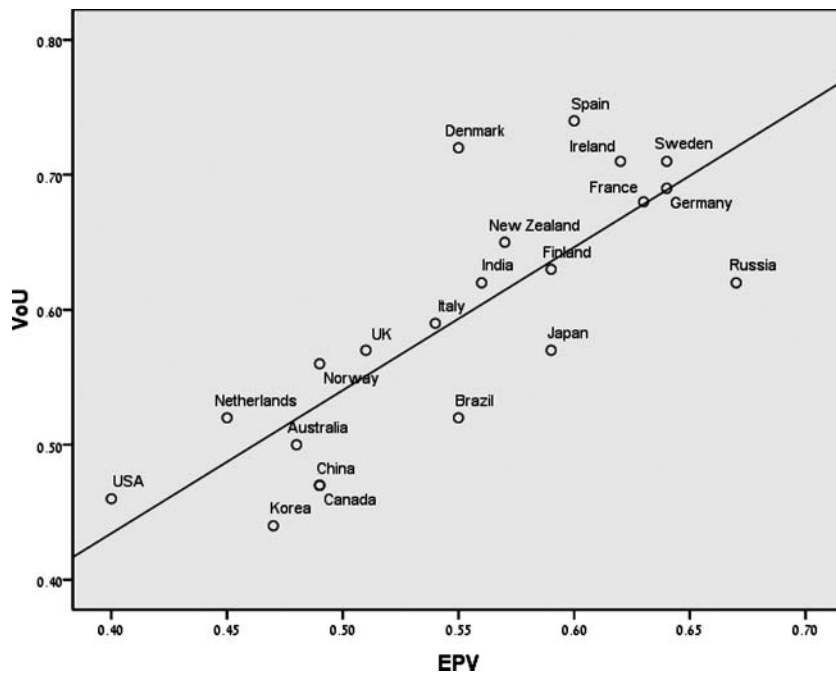


FIG. 4. Content vs. variety of searches.

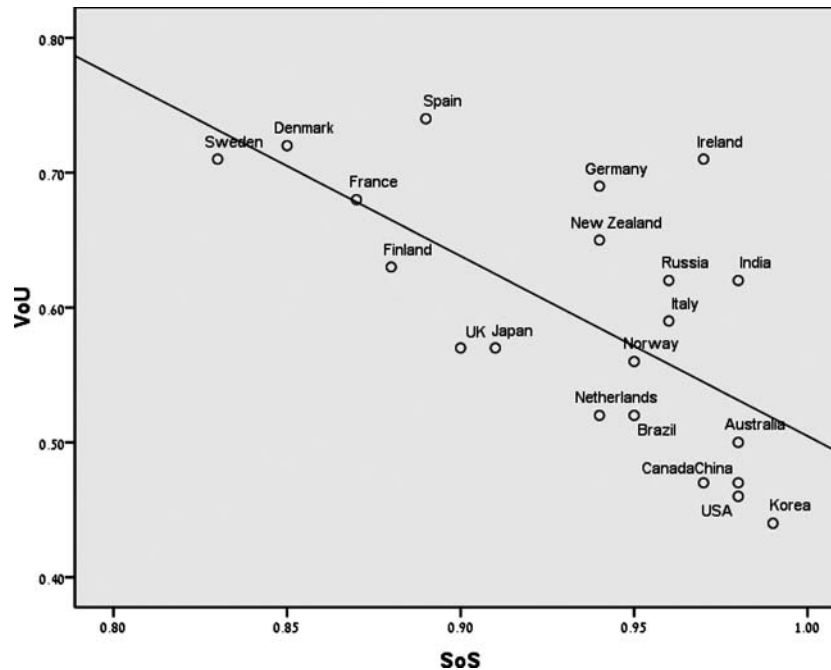


FIG. 5. The trade-off between variety and specificity.

Thus, a negative correlation between the two indices is expected. As there are no assumptions about their distribution, a Spearman<sup>22</sup> one-tailed correlation test confirms that the VoU Index and the SoS Index have a strong negative correlation ( $-0.7$ ,  $p < .01$ ).

This suggests that countries with more specific search queries (i.e., high SoS Index) will usually also display a lower variety of search topics (low VoU Index) and vice versa. In other words, there is a certain trade-off between the variety and the specificity of searches. One possible reason for that is that entertainment-related search queries (e.g., “hilarly duff” or “green day,” which were popular in Canada in February 2005) tend to be more specific and focus on certain people or television programs, whereas politics- and economics-related search queries (e.g., “aftonbladet” or “expressen,” which were popular in Sweden during 2004 and 2005) tend to refer to general news or shopping portals (in which users are often required to continue and search for more specific information). This assumption gets further support in a Spearman<sup>23</sup> one-tailed correlation test that indicates a strong positive correlation between the SoS values and the percentage of entertainment-related searches in each country. Similarly, a strong negative correlation was indicated between the SoS values and the percentage of shopping-related searches, indicating that many shopping-related searches are more general (e.g., referring to general shopping portals rather than specific products and services).

While most countries with high SoS values tend to have a greater concentration of entertainment-related searches and thus less variety, findings also indicate that it is possible to maximize the two. A combination of the VoU and the SoS indices in one graph reveals the differences between countries in terms of the specificity and the variety of searches.

Figure 5 shows the negative relation between the indices. It suggests that countries with more specific search queries exercise greater control and manipulation of online information, whereas countries with a greater variety of searches are exposed to a wider range of information, indicating that they display a better understanding of the various applications of online information. Those who can maximize the opportunities of the search engine as an instrument for providing and retrieving information in a wider range of fields and with greater accuracy and depth display better information skills (see also Bonfadelli 2002). Looking at the international level, the model indicates that countries above the best-fit line exercise a better politics of online information in terms of search accuracy and variety of information uses. In particular, search queries from Ireland and Germany exhibit a higher balance of variety and accuracy than searches from other countries. Although they are as varied as searches from Sweden, Denmark, or France, they are more accurate and specific. Thus, while news-related searches in Sweden and Denmark were for general portal sites, in Germany and Ireland popular searches were more specific,

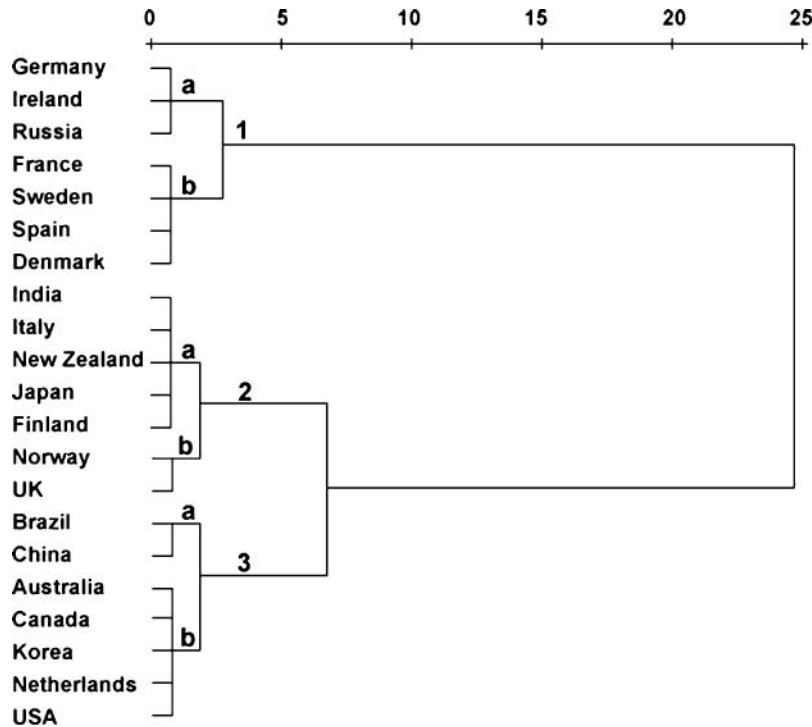


FIG. 6. Hierarchical cluster analysis.

for example, “george bush,” “pope” (or “papst” in German), or “vatican” (or “vatikan” in German).<sup>24</sup>

### Cluster Analysis

A useful integration of the three indices in a comprehensive cross-national comparison can be achieved by employing hierarchical cluster analysis (Aldenderfer and Blashfield 1984; Johnson 1967; Lance and Williams 1967). The purpose of cluster analysis is to measure the distance between each pair of objects (e.g., countries) in terms of the variables suggested in the study (e.g., indices), and then to group objects that are close together. In our case, the cluster analysis is used as a complementary method for validating and supporting previous results, as well as for providing a better insight into the differences in information uses in different countries. Although the various indices indicate the *ranking* of countries in terms of different information uses, cluster analysis allows a more specific look at the *similarities* and *differences* between countries, thus identifying groups of countries with similar information searches.<sup>25</sup>

The clustering was performed based on the Ward (1963) method, which was found to be the most suitable, as it creates a small number of clusters with relatively more countries. Additionally, the Ward method was proved to outperform the other hierarchical methods (Punj and Stewart 1983; Harrigan 1985) in producing homogeneous and interpretable clusters.

Figure 6 shows the results of a hierarchical cluster analysis of countries based on the three indices.<sup>26</sup> The horizontal axis shows the distance between each cluster using the Ward method, in which we identify six clusters with an optimal number of two to five countries in each.

Figure 6 shows that Germany, Ireland, and Russia are included in cluster 1a. The previous analyses (see figures 4 and 5) help to trace the factors behind this classification, indicating that cluster 1a contains the leading countries in terms of all three aspects of the digital divide in information use. They all have a relatively heterogeneous use of online information of high political and economic value. They exercise a strong politics of online information by using accurate and specific search queries. Cluster 1b comprises four countries: France, Sweden, Spain, and Denmark. The common factors for these countries are a variety of political and economic information uses, combined with a low specificity of information use.

Cluster 2a consists of India, Italy, New Zealand, Japan, and Finland, which have a medium variety of searches and medium economic and political value. Cluster 2b comprises Norway and the United Kingdom, also having a medium variety and specificity of searches. However, those countries demonstrate fewer economic and political information uses, and a greater use of online information for entertainment purposes.

Cluster 3a comprises China and Brazil, which demonstrate a low variety and a high accuracy of information

**TABLE 3**  
Summary of cluster analysis

Cluster 1—High scores		Cluster 2—Medium scores		Cluster 3—Low scores	
Cluster 1a:	Cluster 1b:	Cluster 2a:	Cluster 2b:	Cluster 3a:	Cluster 3b:
High variety	High variety	Medium variety	Medium variety	Low variety	Low variety
High accuracy	Low accuracy	High–medium accuracy	Medium accuracy	High–medium accuracy	High accuracy
High EPV	High EPV	High–medium EPV	Medium–low EPV	Medium EPV	Low EPV
Germany	France	India	Norway	China	Canada
Ireland	Spain	Italy	UK	Brazil	Australia
Russia	Sweden	New Zealand			Netherlands
	Denmark	Japan			South Korea
		Finland			United States

uses. They both exercise an extensive use of socially related information, and therefore their EPV Index is medium. Cluster 3b comprises South Korea, the Netherlands, Australia, Canada, and the United States. The common factors of these countries are their low variety of information uses, their extensive use of entertainment-related information, and their high specificity of search queries.

Table 3 summarizes the cluster analysis of countries and the different compositions of information uses and skills in each group.

This digital divide in information use has important political and social implications. Countries that can maximize the variety and accuracy of information search, especially Germany, Ireland, and Russia, also display greater information skills. Other countries, notably the United States, fail to exercise a competitive politics of online information, at least in the context of the suggested framework, based on certain parameters of search queries.

## SUMMARY AND DISCUSSION

One of the early attempts (Sciadas 2003) to monitor the digital divide and construct a Digital Divide Index took into account not only ICT resources, but also information skills (which were measured by education indicators). Subsequently, a report for the WSIS ranked countries by their “info-density,” which is the extent of ICT resources in each country, and “info-use,” which is the uptake and intensity of their uses. Similar to various other recent attempts by United Nations Educational, Scientific and Cultural Organization (2005) to measure the digital divide, the report indicated the very high scores of Western Europe, the United States, Canada, Hong Kong, Singapore, South Korea, Japan, Australia, and New Zealand, compared with the very low scores of developing countries.

Moreover, ranking was highly correlated with GDP per capita.

The methodology for this study was designed to provide a view from a different angle on the digital divide, by looking at the most popular search queries in Google and Yahoo! in various countries. In line with WSIS reports, we expected that the leading countries in terms of economics and technology would display a greater versatility and accuracy in their information search. Additionally, since most content is in English, it was expected that users from English-speaking countries would demonstrate a greater variety of searches, and therefore a better politics of online information.

The findings indicate, however, that many leading countries in terms of economics and technology display a relatively narrow variety and extent of political and economic searches. Countries with higher EPV and VoU scores such as Russia do not lead in terms of gross domestic product (GDP) per capita or percentage of users. Together with Germany, Ireland, Spain, France, and Sweden, they display the greatest variety of searches, as well as the highest extent of political and economic searches. In contrast, countries such as the United States, Canada, Australia, and the United Kingdom, which are all native English-speaking countries, exhibit the lowest EPV and VoU scores, in spite of the fact that, together with South Korea and the Netherlands, they are also the leading countries in terms of percentage of users. Popular search queries in these countries were relatively homogeneous, although more accurate, and concentrated mainly on entertainment.

The narrow range of information uses in some developed countries, such as the United States, Canada, and Australia, matches the increasing Internet commercialization and the dominance of popular channels, which have reinforced highly concentrated Internet traffic. Empirical

studies indicate that the vast majority of visits are aimed at only a small percentage of the Web sites (Hitwise 2008; Webster and Shu-Fang 2002; Waxman 2000). Dominant and popular Web sites continuously customize information and advertisements for the specific interests of their users, reinforcing a narrow range of information uses in favor of commercial and popular content (Turow 2005; Rogers 2004; Barzilai-Nahon 2006; Holtz-Bacha and Norris 2001). The high degree of entertainment-related search queries and the narrow range of popular searches and the very specific queries in the United States, Canada, and Australia reflect this trend, suggesting that information in these countries is highly customized, popularized, and commercialized.

One of the implications of relatively low economic and political searches is the increasing digital divide among users within each of these countries. Although many users focus on entertainment, there are comparatively few information-skilled users who have a greater variety of searches. This empowers them, politically and economically, and therefore can result in social and information inequalities. Norris (2000) argues that the ability to customize information propagates a "virtuous circle" between media and political users, where those who are interested in politics acquire their political content, which in return further empowers them to act politically. Those who are interested in entertainment acquire their preferred content, which in return further reduces their ability and interest to act politically. Hence, the growing ability of users to customize their information through search engines encourages social polarization (Sunstein 2001), and deepens the digital divide between users in these countries.

In contrast, countries with higher scores in all indices, such as Germany, Ireland, and Russia, display greater search skills based on the suggested indicators, which may have several possible reasons. First, it is important to note that there is a significant digital divide of access among the countries observed in this study (see also table A4 for the percentage of users in each country). The high EPV score in Russia, for example, can be attributed to its comparatively early exposure to commercialization and privatization processes. Likewise, because fewer than 17 percent of the Russian population subscribe to ISPs, it could be also argued that there is a higher percentage of information-skilled users among the Russian online community, and among Russian Google users in particular.<sup>27</sup>

Second, when opting out countries with a very low percentage of online users and comparing only countries with more than 40 percent of online users, a negative correlation was found between the percentage of online users and the percentage of business-related searches ( $p < .05$ ). In particular, countries with relatively higher percentages of online users, such as the United States, Canada, the

Netherlands, Sweden, Denmark, and Norway, had a lower percentage of business-related searches, whereas countries with relatively lower percentages of online users, such as Germany, Ireland, France, and Spain, had more business-related searches. A possible explanation for this difference may be the higher percentages among the online users in the former group of countries of youngsters, who usually search more for entertainment-related information (rather than business-related information), and thus further contribute to the low EPV values.

However, as no significant correlation was found between the percentage of online users or the per capita GDP and the EPV scores, we believe that apart from commercialization and Internet usage, there might be some other demographic, social, political, and cultural reasons why certain countries, such as Germany, France, Ireland, and Russia, displayed higher scores in all measurements, and other countries, notably the United States, lag behind. It could be, for example, the result of the intense national political or economic changes that some countries have undergone, engendering greater political and economic concerns among users in these countries. Looking at their popular search queries revealed an ongoing trend of relatively more accurate searches from a wider range of topics (e.g., business, news, and society).

As was previously suggested, this article only opens a path to investigate search queries in the context of the digital divide, suggesting new methods of studying, measuring, and conceptualizing the digital divide of information uses. Obviously, a study that focuses on search query analysis is limited to the users and uses of a specific search engine. It does not and cannot predict, for example, the ability to reach Web sites directly without the help of search engines while acquiring political and economic information. Similarly, it cannot indicate what happens after people search and how they actually use the information available to them. Complementary studies should be designed to observe the demographic profile of search engine users, and examine in depth the processes of information retrieval in various countries, and their economic, political, social, and cultural implications. These kinds of observations may be more limited in scope, but may also help to better understand the reasons behind the current findings.

Finally, it is important to mention that the Internet develops very quickly. While this article covers only two years of study in the phase of the Internet institutionalization and penetration, findings may vary and change in predictable, but also unpredictable, ways in the near future. The effects of the changes in the pattern of use (i.e., Web 2.0 and particularly user-generated content) have to be incorporated into future studies. Together with the increasing ability to customize online information through search engines, and the growing understanding

of its various applications, which also provide advantages to more sophisticated users and corporations, it is expected that the digital divide of information uses will widen, unless governments, international organizations, and non-governmental organizations raise this issue to the top of their priority list.

## NOTES

1. See also the section on crossnational comparison and its limitations.

2. The variety of search analysis is based on a method developed in a previous longitudinal study (Segev, Ahituv, and Barzilai-Nahon 2007) that examined the diversity of content and form of the national homepages of MSN and Yahoo!.

3. Between 2004 and 2005, Google Zeitgeist regularly displayed the most popular search queries in general only for its local interfaces (Google.co.uk, Google.co.jp, etc.). For its main interface, Google.com, which has a majority of American users, it displayed the most popular search queries in specific topics (news, sports, television, etc.). For this study, we could not get additional data on popular searches beyond the publicly available data from Google.

4. See also table A3, as well as more recent data available from Alexa (2009).

5. Following the information provided in Yahoo! Buzz (Yahoo! 2005), it is estimated that around 0.5 percent of the online users in Yahoo! searched for one of the most popular search queries. Thus, each popular query that appeared in Yahoo! Buzz was searched by more than 1 million unique users. Consequently, it could be estimated that for the national interfaces in Google, each query was searched by between 50,000 and 500,000 unique users.

6. See also appendix A for data on the most popular search engines.

7. In Google.com 70 percent of search queries in 2004 and 2005 were about entertainment, which is identical to the results in Yahoo.com.

8. Another seven countries—Chile, Greece, Israel, Poland, South Africa, Turkey, and Vietnam—were added to Google Zeitgeist in September 2005, but were not included in our analyses because of the relatively small number of search queries available and the short duration of sampling.

9. During the sampling period Yahoo!'s archives displayed twenty and Google's archives ten-to-fifteen popular search queries for each country each month. For the complete list also see table B1.

10. These data were obtained following e-mail correspondence between the researchers and Google representatives in December 2008.

11. Similar filters are used in Google SafeSearch, Google Hot Trends, and Google Suggest. Also, see the explanation on the construction of Google Hot Trends data (<http://google.com/intl/en/trends/about.html>) and the filtering of porn-related queries in Google Suggest (<http://labs.google.com/suggestfaq.html#q12>).

12. Google maintains a very tight policy regarding the exposure or sharing of corporate data and was even summoned to court for refusing to hand in a list of search queries to the U.S. government (Hafner and Ritchel 2006).

13. The features and data of the ODP and Google Web Directory were observed and analyzed in 2005 and 2006, and may change in the future.

14. In the same month and the same country the other query was "alla hjärtans dag" (which in English means "the day of the hearts" and refers to Valentine's Day). In this way, the query "alla hjärtans dag" could provide new information about the query "hjärtan," which was subsequently classified as Society > Holidays > Valentine's Day. It is still possible that there were some users who used the word "hjärtan" in Valentine's Day to find information about anatomy, health, or recreation, but the initial goal was to intelligently identify the major purpose and the main interest that most users have in a particular month and a particular country.

15. Hereafter the term *information uses* refers to the use of information in Google and Yahoo! and particularly to popular search queries, not to what people actually do with the information they acquire.

16. Shopping-related search queries, which mostly referred to e-commerce, consumer information and price comparison portals, were considered as economic-related searches. See also DiMaggio et al., 2004.

17. DiMaggio et al., 2004 stress the importance of socially-related information, which is associated with higher education, higher income, and the digital divide, particularly by enhancing the social networks and opportunities of online users. Similarly, Robinson et al., 2003 stressed the importance of education- and health-related information, which is also associated with greater information literacy and skills. Subsequently, these categories were considered as medium-level categories for the purpose of this investigation.

18. Interestingly, despite the local elections in France, many popular search queries in this month were related to shopping and did not refer to this specific event. In most cases, popular search queries tend to be more constant and similar from one month to another, reflecting more general trends. In contrast, the "top-gaining" search queries (which were not included in this study) tend to be more "sensitive" to local social and political changes and therefore reflect more specific trends. In some rare occasions, such as the U.S. presidential election in 2004, the death of Pope John Paul II, or the tsunami in Southeast Asia, popular search queries reflected these regional and global events in many countries. However, the longitudinal investigation over twenty-four months helped to minimize the possible effect of specific events on the findings.

19. It is important to mention that not all categories in Google Web Directory are associated with a similar variety of information. For example, health or games may be associated with less information than arts and entertainments. Nonetheless, Google Web Directory covers a wide range of topics and the VoU Index is designed to provide a more general distinction between countries that display mostly entertainment-related search queries and countries that display also politics-, economics-, and society-related searches. See also Segev et al. 2007 for the methods of diversity analysis of information.

20. The lower limit of the VoU (0.27) is a theoretical case in which all ten queries are from one of the fourteen categories (e.g., all ten search queries are entertainment-related). The upper value of the VoU (1.52) is a theoretical case where all ten queries are divided equally along the fourteen categories (ten categories have one query and four categories have no queries). In August 2005, Google Zeitgeist started reporting fifteen search queries per a country, and therefore the upper value of the VoU could theoretically reach 4.01 (i.e., thirteen categories have one query and one category has two queries), whereas the lower limit stays at 0.27. For the United States, Yahoo! Buss reported twenty search queries, and therefore the upper value of the VoU could theoretically



reach 2.78 (i.e., six categories have two queries and eight categories have one query), whereas the lower limit stays at 0.27. In practice, the highest value of the VoU was found to be 1.17 both in France in August 2004 and in Spain in October 2004. The lowest value of the VoU was 0.3 in South Korea in September 2004 and in January 2005, when nine out of ten search queries were entertainment-related.

21. A Pearson correlation test yielded similar results, supporting the positive correlation between the two indices.

22. A Pearson correlation test yielded similar results, supporting the negative correlation between the two indices.

23. A Pearson correlation test yielded similar results with a *p*-value of less than 0.01.

24. It is very possible that Google is used in Sweden, Denmark or France relatively more as a general gateway to local news and shopping portals, where more specific second-level searches are made. To this end, our study is limited to the examination of the search differences in Google and Yahoo!.

25. Cluster analysis was also applied as a complementary method in Segev et al., 2007, which examined the similarities and differences between homepages of MSN and Yahoo! in terms of content and form.

26. Standardized values of the indices were used for this analysis. As Punj and Stewart 1983 suggest, using standardized variables in a cluster analysis reduces the effect of the outliers, enabling the examination of all the countries in the dataset.

27. This assumption requires further investigation. The data indicate that in Russia, there are more than 20 million users; thus, even if the majority of users are relatively highly information-skilled, they still represent a significant number of users. Moreover, there was no significant correlation between the percentage of online users in different countries and their index values. In India, Brazil and China, for example, fewer than 15 percent subscribe to an ISP, but their EPV and VoU scores are medium or low.

## REFERENCES

- Aldenderfer, M. S., and R. K. Blashfield. 1984. *Cluster analysis*. Beverly Hills, CA: Sage.
- Alexa. 2009. Site information for Yahoo!. <http://www.alexa.com/siteinfo/yahoo.com> (accessed September 17, 2009).
- Anderson, R. H., T. K. Bikson, S. A. Law, and B. M. Mitchell. 1995. *Universal access to email: Feasibility and societal implications*. Santa Monica, CA: RAND.
- Bar-Ilan, J. 2004. The use of Web search engines in information science research. In *Annual review of information science and technology*, ed. B. Cronin, vol. 33, 231–288. Medford, NY: Information Today.
- Barnard, F. M. 2004. *Herder on nationality, humanity, and history*. Montreal: McGill–Queen’s University Press.
- Barzilai-Nahon, K. 2006. Gatekeepers, virtual communities and their gated: Multidimensional tensions in cyberspace. *International Journal of Communications, Law and Policy* 11:1–28.
- Bawden, D. 2001. Information and digital literacies: A review of concepts. *Journal of Documentation* 57(2):218–59. <http://www.emeraldinsight.com/Insight/viewContentItem.do?contentType=Article&hdAction=Inkpdf&contentId=864156> (accessed January 5, 2007).
- Bonfadelli, H. 2002. The Internet and knowledge gaps: A theoretical and empirical investigation. *European Journal of Communication* 17:65–84.
- Castells, M. 2004. Informationalism, networks, and the network society: A theoretical blueprinting. In *The network society: A cross-cultural perspective*, ed. M. Castells, 3–48. Northampton, MA: Edward Elgar.
- Chau, M., X. Fang, and C. C. Yang. 2007. Web searching in Chinese: A study of a search engine in Hong Kong. *Journal of the American Society for Information Science and Technology* 58(7):1044–1054.
- Cho, J., and S. Roy. 2004. Impact of search engines on page popularity. In *Proceedings of the 13th International Conference on World Wide Web, May 17–20, New York, 20–29*. New York: Association for Computing Machinery.
- Choo C. W., B. Detlor, and D. Turnbull. 2000. Information seeking on the Web—An integrated model of browsing and searching. *First Monday* 5(2). [http://www.firstmonday.dk/issues/issue5\\_2/choo](http://www.firstmonday.dk/issues/issue5_2/choo) (accessed August 12, 2005).
- Ciolek, T. M. 2003. *The Internet and its users: The physical dimensions of cyberpolitics in Eastern Asia*. Paper presented at the From the Book to the Internet: Communications Technologies, Human Motions, and Cultural Formations in Eastern Asia Conference, University of Oregon, Eugene, October 16–18. <http://www.ciolek.com/PAPERS/oregon-2003-text.html> (accessed December 6).
- comScore. 2009. *comScore releases June 2009 U.S. search engine rankings*. July 16. [http://www.comscore.com/Press\\_Events/Press\\_Releases/2009/7/comScore\\_Releases\\_June\\_2009\\_U.S.\\_Search\\_Engine\\_Rankings](http://www.comscore.com/Press_Events/Press_Releases/2009/7/comScore_Releases_June_2009_U.S._Search_Engine_Rankings) (accessed September 16, 2009).
- comScore. 2009b. *Global search market draws more than 100 billion searches per month*. August 31. [http://www.comscore.com/Press\\_Events/Press\\_Releases/2009/8/Global\\_Search\\_Market\\_Draws\\_More\\_than\\_100\\_Billion\\_Searches\\_per\\_Month](http://www.comscore.com/Press_Events/Press_Releases/2009/8/Global_Search_Market_Draws_More_than_100_Billion_Searches_per_Month) (accessed September 16, 2009).
- DiMaggio, P., and E. Hargittai. 2001. *From the digital divide to digital inequality*. Paper presented at the Annual Meeting of the American Sociological Association, August 16–19, Chicago.
- DiMaggio, P., E. Hargittai, C. Celeste, and S. Shafer. 2004. From unequal access to differentiated use: A literature review and agenda for research on digital inequality. In *Social inequality*, ed. K. Neckerman, 355–400. New York: Russell Sage Foundation.
- DiMaggio P., E. Hargittai, W. R. Neuman, and J. P. Robinson. 2001. Social Implications of Internet. *Annual Review of Sociology* 27:307–36.
- Dmoz.org. 2005. Open directory project. <http://dmoz.org/about.html> (accessed January 23, 2005).
- ELOST. 2008. e-Government for low socioeconomic status-groups. <http://www.elost.org> (accessed July 21, 2008).
- Google. 2007. Google 2004 and 2005 Zeitgeist Archive. <http://www.google.com/intl/en/press/zeitgeist/archive2004.html> (accessed March 12, 2006).
- Hafner, K., and M. Ritel. 2006. Google resists U.S. subpoena of search data. *New York Times*. January 20. <http://www.nytimes.com/2006/01/20/technology/20google.html> (accessed December 13, 2008).
- Hargittai, E. 2000. Open portals or closed gates? Channeling content on the World Wide Web. *Poetics* 27:233–253.
- Hargittai, E. 2002. Second-level digital divide: Differences in people’s online skills. *First Monday* 7(4), <http://chnm.gmu.edu/digitalhistory/links/pdf/introduction/0.26c.pdf> (accessed November 27, 2009).

- Hargittai, E. 2003. The digital divide and what to do about it. In *New economy handbook*, ed. D. C. Jones, 822–39. San Diego, CA: Academic Press.
- Harrigan, K. R. 1985. An application of clustering for strategic group analysis. *Strategic Management Journal* 6:55–73.
- Hitwise. 2006. *Hitwise search engine ratings*. <http://searchenginewatch.com/showPage.html?page=3099931> (accessed January 4, 2007).
- Hitwise. 2008. *Hitwise US—Top 20 Web sites—June 2008*. <http://www.hitwise.com/datacenter/rankings.php> (accessed July 22, 2008).
- Holtz-Bacha, C., and P. Norris. 2001. To entertain, inform and educate: Still the role of public television in the 1990s? *Political Communications* 18(2):123–40.
- Howard, P., L. Rainie, and S. Jones. 2001. Days and nights on the Internet: The impact of a diffusing technology. *American Behavioral Scientist* 45(3):383–404.
- Jansen, B. J., and A. Spink. 2003. An analysis of Web information seeking and use: Documents retrieved versus documents viewed. In *Proceedings of the 4th International Conference on Internet Computing, Las Vegas, NV, June 23–26*, 65–69. Las Vegas, NV: CSREA.
- Jansen, B. J., and A. Spink. 2004. How are we searching the World Wide Web? A comparison of nine-search engine transaction logs. *Information Processing and Management* 42:248–63.
- Jansen, B. J., A. Spink, and T. Saracevic. 2000. Real life, real users and real needs: A study and analysis of users' queries on the Web. *Information Processing and Management* 36(2):207–27.
- Johnson, S. C. 1967. Hierarchical clustering schemes. *Psychometrika*, 38:241–54.
- Lance, G. N., and W. T. Williams. 1967. A general theory of classificatory sorting strategies. *Computer Journal* 9:60–64.
- Negroponte, N. 1995. *Being digital*. New York: Knopf.
- Nielsen/NetRatings. 2009. Nielsen (2009), Nielsen announces July U.S. search share rankings, August 12. [http://en.us.nielsen.com/main/news/news\\_releases/2009/august/Nielsen\\_Announces\\_July\\_U.S.\\_Seach\\_Rankings](http://en.us.nielsen.com/main/news/news_releases/2009/august/Nielsen_Announces_July_U.S._Seach_Rankings) (accessed September 14, 2009).
- Norris, P. 2000. *A virtuous circle: Political communications in post-industrial societies*. New York: Cambridge University Press.
- Norris, P. 2001. *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. New York: Cambridge University Press.
- O'Neill, E. T., B. F. Lavoie, and R. Bennett. 2003. Trends in the evolution of the public Web: 1998–2000. *D-Lib Magazine* 9(4), <http://www.dlib.org/dlib/april03/lavoie/04lavoie.html>.
- Pastore, M. 2000. Web pages by language. <http://www.clickz.com/showPage.html?page=408521/>.
- Pu, H. T., S. L. Chuang, and C. Yang. 2001. Subject categorization of query terms for exploring Web users' search interests. *Journal of the American Society for Information Science and Technology* 53(8):617–30.
- Punj, G., and D. W. Stewart. 1983. Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research* 20(2):134–48.
- Putnam, R. D. 2000. *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Rheingold, H. 1993. *The virtual community: Homesteading on the electronic frontier*. Reading, MA: Addison-Wesley.
- Robinson, J. P., P. DiMaggio, and E. Hargittai. 2003. New social survey perspectives on the digital divide. *IT & Society* 1(5):1–2.
- Rogers, R. 2004. *Information politics on the Web*. Cambridge, MA: MIT Press.
- Ross, N., and D. Wolfram. 2000. End user searching on the Internet: An analysis of term pair topics submitted to the excite search engine. *Journal of the American Society for Information Science* 51(10):949–58.
- Sciadas, G., ed. 2003. *Monitoring the digital divide ... and beyond. Technical Report, Orbicom*. Montreal: Orbicom International Secretariat. [http://www.orbicom.uqam.ca/projects/ddi2002/2003\\_dd.pdf\\_en.pdf](http://www.orbicom.uqam.ca/projects/ddi2002/2003_dd.pdf_en.pdf) (accessed December 9, 2007).
- Webster, J. G., and L. Shu-Fang. 2002. The Internet audience: Web use as mass behavior. *Journal of Broadcasting & Electronic Media* 46(4): 656–63.
- Segev, E., N. Ahituv, and K. Barzilai-Nahon. 2007. Mapping diversities and tracing trends of cultural homogeneity/heterogeneity in cyberspace. *Journal of Computer-Mediated Communication* 12(4). <http://jcmc.indiana.edu/vol12/issue4/segev.html> (accessed July 23, 2007).
- Silverstein, C., M. Henzinger, H. Marais, and M. Moricz. 1999. Analysis of a very large Web search engine query log. *ACM SIGIR Forum* 33(3): 6–12.
- Spink, A., B. J. Jansen, D. Wolfram, and T. Saracevic. 2002. From e-sex to e-commerce: Web search changes. *IEEE Computer* 35(3):107–11.
- Spink, A., S. Ozmutlu, H. C. Ozmutlu, and B. J. Jansen. 2002. U.S. versus European Web searching trends. *SIGIR Forum* 36(2): 32–38.
- Sunstein, C. 2001. *Republic.com*. Princeton, NJ: Princeton University Press.
- Turow, J. 2005. Audience construction and culture production: Marketing surveillance in the digital age. *Annals of the American Academy of Political and Social Science* 597:103–21.
- UNESCO. 2005. *Towards knowledge societies*. Paris: United Nations Educational, Scientific and Cultural Organization. <http://unesdoc.unesco.org/images/0014/001418/141843e.pdf> (accessed March 4, 2006).
- UNESCO. 2006. In focus: Measures and indicators. [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=20937&URL\\_DO=TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=20937&URL_DO=TOPIC&URL_SECTION=201.html)
- Warschauer, M. 2004. *Technology and social inclusion: Rethinking the digital divide*. Cambridge, MA: MIT Press.
- Ward, J. H. Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58(301):236–44.
- Waxman, J. 2000. The old 80/20 rule take one on the jaw. *Internet trend report 1999, Review*. San Francisco: Alexa Research.
- Webber, S. 2000. Conceptions of information literacy: New perspectives and implications. *Journal of Information Science* 26(6): 381–97.
- Wolfram, D., A. Spink, B. J. Jansen, and T. Saracevic. 2001. Vox populi: The public searching of the Web. *Journal of the American Society for Information Science and Technology* 52(12): 1073–1074.
- WSIS. 2003. Declaration of principles, building the information society: A global challenge in the new millennium. World Summit on the Information Society. Document WSIS-03/GENEVA/DOC/4-E, December 9. <http://www.itu.int/wsis/docs/geneva/official/dop.html> (accessed October 8, 2007).
- Yahoo! 2005. *Buzz Index—Top Yahoo! Web Searches*. <http://buzz.yahoo.com/> (accessed March 30, 2005).

## APPENDIX A—GENERAL STATISTICS

Table A1 displays the share of searches conducted by U.S. Web surfers in June and July 2009 in each search engine, as published by two different search marketing companies—comScore (2009) and Nielsen/NetRatings (2009).

Table A2 displays the share of searches done by World Wide Web surfers in July 2009 in each of the search engines. The data were extracted from statistics provided by comScore (2009b).

Because the experiment presented in this article is based on popular search queries in Google Zeitgeist from 2004 and 2005, Table A3 displays the share of searches in different countries during this period. It shows that Yahoo! used to be much more popular search engine in the United States, while Google kept its lead in Canada, France, and the United Kingdom.

### Percentage of Online Users

Table A4 summarizes the percentage of online users in each of the observed countries in 2006. The data were extracted from the CIA *World Factbook*.

## APPENDIX B—DATA SOURCES AND METHODS

### Countries in Data Set

Table B1 summarizes the countries that appeared in Google Zeitgeist each month during 2004 and 2005, and were used in the analysis of popular search queries.

### Categories and Subcategories

Table B2 summarizes the first two level categories of search queries as were classified in this study by the ODP. The numbers in brackets indicate how many queries were classified under each category or subcategory.

**TABLE A1**  
Share of searches in the United States

Search engine	Share of searches, ComScore (June 2009)	Share of searches, Nielsen/NetRatings (July 2009)
Google	65.0%	64.8%
Yahoo!	19.6%	17.0%
Microsoft	8.4%	9.0%
Ask	3.9%	3.1%

Source: comScore qSearch and Nielsen/NetRatings.

**TABLE A2**  
Share of searches worldwide

Search engine	Share of searches
Google	67.5%
Yahoo!	7.8%
Baidu.com	8.7%
Microsoft	2.9%

Source: comScore qSearch, accessed in September 2009 (<http://www.comscore.com>).

**TABLE A3**  
Share of searches worldwide, April 2004

Site	Canada	France	United Kingdom	United States
Google Sites	70%	80%	77%	44%
Yahoo! Sites	17%	10%	14%	37%
MSN-Microsoft Sites	13%	10%	9%	19%

Note. Source: comScore qSearch, accessed in October 2004 (<http://www.comscore.com>).

**TABLE A4**  
Percentage of online users in the observed countries

Country	Online populations (millions)		
	Population	Internet Users	% online users
New Zealand	4.08	3.2	78.4
Sweden	9.02	6.8	75.4
Australia	20.26	14.18	70.0
South Korea	48.85	33.9	69.4
Denmark	5.45	3.76	69.0
United States	298.44	203.82	68.3
Norway	4.61	3.14	68.1
Japan	127.46	86.3	67.7
The Netherlands	16.49	10.81	65.6
Canada	33.1	20.9	63.1
Finland	5.23	3.29	62.9
United Kingdom	60.61	37.8	62.4
Germany	82.42	48.72	59.1
Ireland	4.06	2.06	50.7
Italy	58.13	28.87	49.7
France	60.88	26.21	43.1
Spain	40.4	17.14	42.4
Russia	142.89	23.7	16.6
Brazil	188.1	25.9	13.8
China	1310	111	8.5
India	1100	50.6	4.6

**TABLE B1**  
Countries that appeared in Google Zeitgeist report

Country	Period 2004	Period 2005	Total queries
Australia	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	220
Brazil	January 4–November 4, 2004*	January 5–September 5, November 5, 2005	175
Canada	January 4–November 4, 2004*	January 5–July 5, 2005	190
China	January 4–November 4, 2004*	January 5–June 5, September 5, 2005	135
Denmark	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	160
Finland	January 4–November 4, 2004*	January 5–September 5, 2005	155
France	January 4–November 4, 2004*	January 5–September 5, November 5, 2005	235
General (United States)	January 4–November 4, 2004* (in Google)	July 4–October 5, 2005 (in Yahoo!)	820
Germany	January 4–November 4, 2004*	January 5–July 5, November 5, 2005	205
India	—	January 5–September 5, November 5, 2005	110
Ireland	—	Jan-05–Jul-05, Sep-05	85
Italy	January 4–August 4, October 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	220
Japan	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	220
Korea	January 4–November 4, 2004*	January 5–July 5, August 5, 2005	145
The Netherlands	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	220
New Zealand	—	January 5–July 5, September 5, November 5, 2005	100
Norway	January 4–November 4, 2004*	January 5–July 5, 2005	130
Russia	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	150
Spain	January 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	220
Sweden	January 4–November 4,	January 5–July 5, September 5, November 5, 2005	150
United Kingdom	January 4, May 4, July 4–November 4, 2004*	January 5–July 5, September 5, November 5, 2005	175

\*Google Zeitgeist provided annual data about the popular search queries in general during 2004. These data were used to validate and support the results of the monthly data.

**TABLE B2**  
Categories of search queries (first two levels)

Category	Subcategories
Art (1950)	Animation (165), architecture (1), body art (8), celebrities (187), design (6), entertainment (16), events (4), literature (17), magazines and e-zines (2), movies (174), museums (2), music (839), performing arts (265), photography (4), radio (5), television (25), visual arts (1)
Games (167)	Board games (1), gambling (14), eye-hand coordination (1), online (34), paper and pencil (9), trading card games (3), video games (83), (22 without subcategories)
Sports (473)	Baseball (8), basketball (14), boxing (1), cricket (13), cycling (28), darts (1), equestrian (18), events (33), fencing (1), football (32), golf (2), handball (1), hockey (12), martial arts (1), motor sports (44), on the Web (12), paintball (8), skating (1), soccer (185), strength sports (1), tennis (28), water sports (4), winter sports (1), wrestling (20),(4 without subcategories)
Business (173)	Advertising (1), agriculture and forestry (1), business services (6), conglomerates (1), construction and maintenance (1), employment (46), financial services (53), food and related products (1), hospitality (2), industrial goods and services (1), international business (1), investing (1), marketing and advertising (1), real estate (14), shopping (6), telecommunications (36)
Computers (86)	Data communications (2), hardware (11), Internet (20), multimedia (1), programming (3), security (8), software (41)
News (346)	Breaking news (110), directories (5), online archives (107), weather (124)
Shopping (180)	Auctions (35), autos (3), beauty (1), classifieds (7), clothing (15), computers (2), consumer electronics (9), entertainment (4), flowers (15), food (1), general merchandise (52), gifts (1), home and garden (28), office products (1), price comparisons (2), sports (1), vehicles (1), (1 without subcategories)
Science (12)	Agriculture (1), astronomy (8), biology (1), earth sciences (2), technology (10)
Recreation (418)	Autos (49), boating (2), collecting (12), crafts (1), drawing and coloring (2), food (5), gardening (1), humor (12), motorcycles (5), online (41), outdoors (2), parties (1), pets (25), sauna (4), theme parks (7), travel (247), (1 without subcategories)
Society (418)	Chats and forums (69), ethnicity (2), folklore (8), government (35), history (2), holidays (181), issues (9), law (2), organizations (2), people (13), politics (15), relationships (42), religion and spirituality (38)
Reference (197)	Dictionaries (48), directories (36), education (47), encyclopaedias (12), flag (1), libraries (2), maps (47), units of measurement (3)
Health (18)	Alternative (1), beauty (1), conditions and diseases (1), dentistry (3), nutrition (9), organizations (3)
Home (12)	Apartment living (4), consumer information (1), cooking (1), do-it-yourself (1), family (1), food (4), home improvement (1)
Regional (17)	Africa (2), America (1), Asia (2), Europe (12)

Copyright of Information Society is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.