

Search Engines and Ethics

First published Mon Aug 27, 2012; substantive revision Tue Aug 11, 2020

What is an Internet search engine? Why are search engines problematic from an ethical perspective? In this entry, the available philosophical literature on this topic will be critically reviewed. However, relatively few academic works on the topic of search engines have been written from a philosophical perspective. And only a handful of the existing publications that focus specifically on the *ethical* aspects of search engines have been contributed by philosophers (see, for example, Nagenborg 2005).

- [1. Introduction and Overview](#)
- [2. Search Engine Development and Evolution: A Short History](#)
 - [2.1 The Pre-Internet Era of Computing and Information Retrieval \(1940s–1970s\)](#)
 - [2.2 The Early Internet \(pre-Web\) Era \(1980s\)](#)
 - [2.3 The \(Early\) Web Era \(1990s\)](#)
 - [2.4 The “Web 2.0” Era \(2000–Present\)](#)
- [3. Ethical Implications and Core Ethical Issues](#)
 - [3.1 Search Engine Bias and the Problem of Opacity/Nontransparency](#)
 - [3.2 Privacy, Consent, and Non-voluntary Disclosure of Personal Information](#)
 - [3.3 Monitoring and Surveillance](#)
 - [3.4 Democracy, Censorship, and the Threat to Liberty and Freedom](#)
 - [3.5 \(Cyber\)Security and the Internet of Things](#)
- [4. Moral Accountability and Social-Responsibility Issues for Search Engine Companies](#)
 - [4.1 Commerce-Related Conflicts for SECs](#)
 - [4.2 Legal Liability, a “Right to erasure,” and Fake News in the Digital Era](#)
 - [4.3 Some Questions Affecting Trust](#)
- [5. Conclusion](#)
- [Bibliography](#)
- [Academic Tools](#)
- [Other Internet Resources](#)
- [Related Entries](#)

1. Introduction and Overview

It may be difficult to imagine today’s world without search engines. Which high school student has not used a Web search engine to query about some topic or subject? Of course, it is quite possible that many Internet users, both young and old, do not consciously distinguish between the search engines they use and Web browsers that now also typically include search engines as a feature within their user interface. But virtually all Internet users have come to expect and depend on the instantaneous results they receive in response to their various search queries. While there is no shortage of definitions of “search engine,” none has been accepted as *the* standard or universally agreed upon definition. For purposes of this entry, however, the

definition of a (Web) **search engine**, put forth by Halavais (2009, 5–6), is “an information retrieval system that allows for keyword searches of distributed digital text.” We note that this definition includes some important technical terms and concepts that, in turn, need defining and further elucidation. Our examination of key technical concepts underlying search engines is intended to provide a useful context for our analysis of the ethical implications. In this sense, Blanke (2005, 34) is correct that an adequate analysis of the ethical aspects of search engines “requires knowledge about the technology and its functioning.”

We begin with a brief sketch of the history and evolution of search engines, from their conception in the pre-Internet era to the development and implementation of contemporary (“Web 2.0” era) search engines such as Google. Our examination of important historical developments of this technology is intended to address our first question, noted above: “What is a search engine?” It also provides a backdrop for analyzing our second major question, “Why are search engines problematic from an ethical perspective?” where a cluster of ethical concerns involving search engine technology is examined. These include **issues ranging from search engine bias and the problem of opacity/non-transparency, to concerns affecting privacy and surveillance, to a set of issues involving censorship and democracy.** We also describe some emerging ethical concerns affecting **(cyber)security**—at the data, system, and (inter)national levels—generated by the use of “discoverable search engines” in the context of the Internet of Things. Additionally, we question whether search engine companies have any special moral obligations, e.g., in light of **their “privileged” placed in society as “gatekeepers of the Web”** (Hinman, 2005, 21), for which they should be held accountable. In analyzing this question, we also examine a key role that trust now plays for users who have come to depend on search engines, and the companies that develop them, **for access to accurate information.**

Summary of Ethical Issues

How trusting search engines enters the picture

Finally, in the concluding section, we briefly mention some impacts that search engines have for broader philosophical issues (especially in the area of epistemology) that may not be solely or mainly ethical in nature. However, an adequate analysis of these issues is beyond the scope of this entry.

2. Search Engine Development and Evolution: A Short History

Because search engines provide Internet users with access to important information by directing them to links to available online resources on a plethora of topics, many are inclined to see search engine technology in a positive light; some might also assume, as Noble (2018) and others note, that this technology is “value-neutral.” However, search engines can raise a number of ethical controversies. Before examining these controversies, however, we first briefly discuss the history of search engine technology via categories that, for our purposes, reflect four distinct eras: (i) Pre-Internet, (ii) Internet (pre-Web), (iii) early Web, and (iv) Web 2.0. We will see how technical developments in each era have had some implications for the cluster of ethical issues examined in [Section 3](#).

2.1 The Pre-Internet Era of Computing and Information Retrieval (1940s–1970s)

Today, we tend to associate search engines with computer technology, and perhaps more specifically with Internet-based computing and electronic devices. Yet, the early work in search/information retrieval systems was carried out independently of developments in electronic computing. Whereas the first (general purpose) electronic computer—the ENIAC (Electronic Numerical Integrator And Computer)—was completed in November 1945 and announced in February 1946 (Palfreman and Swade, 1991), several decades would pass before Internet search engines became available. Because ENIAC and other early computers were designed primarily to “crunch numbers,” relatively little thought had been given to the kinds of information-retrieval

systems that could be used to *search* through the large amount of data that those non-networked (or “stand-alone”) computers were capable of storing. However, some information-retrieval theorists, as we shall see, had already begun to worry about the amount of information that was becoming available during this period and that, in all likelihood, would proliferate with the advent of computers. In particular, they were concerned about how an ever-expanding repository of information could be organized and retrieved in a practical way. Halavais (2009, 13) notes that early computers “drew on the ideas of librarians and filing clerks” for arranging the stored information that would be retrieved. But some of the leading thinkers in the emerging field of information retrieval (or IR), which Van Couvering (2008) describes as a “hybrid” academic discipline combining elements of information science and computer science, saw that traditional methods for retrieving information would not be effective in the era of electronic computer systems.

One visionary who saw the need for a new kind of organizing and retrieval scheme to manage the expanding volume of information was Vannevar Bush, perhaps the most important figure in the history of information-retrieval/search-engine theory in the pre-Internet era. In his classic article, “As We May Think” (*Atlantic Monthly*, July 1945), published approximately six months before ENIAC’s official announcement, Bush remarked,

The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships.

However, Bush believed that a technological solution to this problem was possible through a system he called *memex*, which he described as a

device in which an individual stores all his books, records, and communications, and which is mechanized so that it can be consulted with exceeding speed and flexibility.

Bush envisioned the memex behaving like an “intricate web of trails” similar to the function of the human mind, which he believed works by a method of “association” and not via an alphabetical index (of the kind typically used in libraries and other cataloging schemes). According to Levy (2008, 508), the most “innovative feature” of Bush’s memex system was the establishing of

associative indices between portions of microfilmed text—what we now call hypertext links—so that researchers could follow trails of useful information through masses of literature.

Via Bush’s “associative indexing” scheme, different pieces of information could be linked or tied together, “as any single item may be caused at will to select immediately and automatically another.” Thus, Bush is often credited with having anticipated the kinds of search engine functions that would eventually be used on the Internet and the World Wide Web.

Two other important figures in the history of search engine theory who made significant contributions during the pre-Internet era were Gerald Salton and Ted Nelson. Salton, who some consider the “father of modern search technology,” developed the SMART (Salton’s Magic Automatic Retriever of Text) information retrieval system. And Nelson, who developed hypertext in 1963, significantly influenced search engine theory through his Project Xanadu (Wall 2011). Although several years passed before Salton’s and Nelson’s contributions could be incorporated into modern search engines, it is worth noting that some very “primitive” search functions had been built into the operating systems for some pre-Internet-era computers. For example, Halavais points out that the UNIX operating system supported a search utility called “Finger.”

Via the Finger command, a UNIX user could search for one or more users who also had active accounts on a particular UNIX system. To inquire about a UNIX user named “Jones,” for example, one could simply enter the command “Finger Jones” at the user prompt on the command line. However, this search function was very limited, since the only kind of information that could be retrieved was information about whether one or more users were currently logged into the system and about which time those users logged in/out. But, as Halavais points out, this rudimentary search facility also enabled UNIX users to arrange limited social gatherings—e.g., users could “Finger” one another to set up a time to play tennis after work (provided, of course, that the users were logged into their UNIX accounts at that time).

Some of the conceptual/technological breakthroughs that occurred during the pre-Internet era of search engine development made possible two kinds of ethical issues examined in [Section 3](#). First, Bush’s “associative indexing” scheme for retrieving information, as opposed to more traditional cataloging schemes based on straight-forward inferential rules and techniques, enabled (even if unintentionally) some of the kinds of “bias” and objectivity-related concerns affecting users’ search results that we examine in [Sections 3.1 and 4.1](#). Second, the kind of search function made possible by the UNIX “Finger” utility, enabling UNIX users to retrieve information about the availability of fellow UNIX users and to acquire information about which times those users logged into and logged out from the system, generated some privacy-and-monitoring-related concerns that are included among the ethical issues we examine in [Sections 3.2 and 3.3](#).

2.2 The Early Internet (pre-Web) Era (1980s)

By the 1960s, plans for developing a vast network of computer networks (i.e., what was eventually to become the Internet) were well underway. And by 1970, work had begun on the ARPANET (Advanced Research Projects Agency Network), which is commonly viewed as the predecessor of the Internet. This US-based project was funded by DARPA (Defense Advanced Research Projects Agency) into the late 1980s, when the National Science Foundation Network (NSFnet) took over the project (Spinello 2011). Although multiple computer networks existed during this period, they were not easily able to communicate and exchange data with one another; a common protocol was needed for the various networks to exchange data between systems. The Transmission Control Protocol/Internet Protocol (TCP/IP) architecture was eventually selected as the standard protocol for the newly emerging Internet. With the implementation of this new standard, there was considerable optimism (especially among many in the academic research community) about the potential for sharing the data that resided in the various computers systems comprising the fledgling Internet. However, **one very important challenge still remained: How could Internet users locate the rich resources potentially available to them? To do this, a sophisticated search program/utility, with a robust indexing system, was needed** to point to the available computer databases that existed and to identify the content that resided in those databases. The first indexes on the Internet were fairly primitive, and as Halavais (2009) points out, “had to be created by hand.”

can explain the importance of gatekeepers of the web

With TCP/IP now in place, privately owned computer networks—including LANs (local area networks) and WANs (wide area networks)—were able to communicate with one another and, in principle at least, also able to exchange vast amounts of information over the network. However, another protocol—one that would be layered on top of TCP/IP—was needed to accomplish this objective. So, FTP (File Transfer Protocol), a client/server-based system, was developed and implemented in response to this need. To exchange or share files with a fellow Internet user in this scheme, one first had to set up an FTP server.^[1] Users could then upload files to and retrieve them from an FTP server, via an FTP client. Perhaps more importantly, they could also now effectively *search* for files with one of the newly developed search engines, the first of which was called ARCHIE.

The ARCHIE search engine enabled users to enter queries based on a limited set of features—mainly “file names.” ARCHIE’s searchable database of file names was comprised of the file directory listings of hundreds of systems available to public FTP servers (and eventually to “anonymous” FTP servers as well). In the early 1990s, two other search engines were also fairly prominent: VERONICA (Very Easy Rodent-Oriented Net-Wide Index to Computer Archives) and JUGHEAD (Jonzy’s Universal Gopher Hierarchy Excavation and Display). Both VERONICA and JUGHEAD had an advantage over the ARCHIE search engine in that they were able to search for plain-text files, in addition to searching for file names. These two search engines also worked in connection with a system called GOPHER. According to Halavais (2009, 22), GOPHER’s “menu-driven approach” to search helped to bring “order to the Internet,” since users “could now navigate through menus that organized documents.”

Some of the technological breakthroughs that occurred during the early-Internet era of search engine development exacerbated a privacy-related ethical issue examined in [Section 3](#). Specifically, Internet-wide search functions, enabled by compliance with the TCP/IP and FTP protocols, dramatically increased the scope of the privacy-and-monitoring concerns (initially generated in the pre-Internet era via applications such as the UNIX “Finger” utility) that we examine in [Sections 3.2](#) and [3.3](#). Additionally, “anonymous” FTP servers, also developed in this period, made it possible for technically-savvy users to upload proprietary files, such as copyrighted software applications, on to the Internet (with anonymity). And the indexing schemes supported by the ARCHIE and GOPHER search systems enabled users to search for and download/share those proprietary files with relative anonymity. Although intellectual property issues are not included among the ethical concerns examined in [Section 3](#), it is worth noting that the development of some search-engine-related applications during this era paved the way for the kinds of illegal file-sharing practices involving copyrighted music that arose in connection with the Napster site in the late 1990s. (The controversial Napster web site was one of the first, as well as the most popular, sites used by many young people at that time to exchange proprietary music online with their friends).

2.3 The (Early) Web Era (1990s)

The first Web site was developed in 1991 (at the CERN European laboratory for particle physics) by Tim Berners-Lee, who also founded the World Wide Web Consortium (W3C) at MIT in 1994. The World Wide Web was based on the Hyper Text Transfer Protocol (HTTP) and used a format called the Hyper Text Markup Language (HTML) for designing and delivering documents; many non-technical users found navigating the Web to be much more friendly and versatile than using GOPHER and FTP to exchange files. For the (HTTP-based) Web to realize its full potential and to become attractive to non-technical users, however, a more intuitive user interface was needed. The Mosaic Web browser (later called Netscape Navigator) became available in 1993 and was the first Internet application to include a graphical user interface (GUI); this interface, with its intuitive features that enabled users to click on hyperlinks, made navigating the Web much easier for non-technical users. Although Netscape Navigator was a Web browser, and not a search engine, it provided a forum in which many specialized Web search engine companies were able to flourish. A host of search engines, most of which were dedicated to specific areas or specific kinds of searches, soon became available. Some search engines that were especially popular during this period were Excite (introduced in 1993) and Lycos and Infoseek (both of which were available in 1994). Others included Looksmart and Alta Vista, introduced in 1995, and Ask.com (originally called AskJeeves) in 1997 (Wall 2011).

Although the internal structure of a search engine is fairly complex—comprising, among other components, programs called “spiders” that “crawl” the Web—the user-interface portion of the search process is quite

straightforward and can be summarized in terms of two steps: (1) a user enters search term/phrase or “keyword” in a “search box”; and (2) the search engine returns a list of relevant Web “pages” that typically include hyperlinks to the pages listed. Many of the early Web search engines were highly specialized and thus could be viewed as “vertical” (i.e., in current technical parlance regarding search engine technology) in terms of their scope. For example, Ask.com was designed to accept queries in the form of specific questions and thus could be viewed as a vertical search engine. Halavais defines a vertical search engine as one that limits itself “in terms of topic, medium, region, language, or some other set of constraints, covering that area in greater depth.” (In this sense, vertical search engines are far more capable of drilling down into particular topics than expanding out into associated subjects.) A few of the popular search engines that flourished during the early Web period, however, were more general, or “horizontal,” in nature. Alta Vista, for instance, was one of the first search engines to fit into this category. Today, most of the major search engines are horizontal, and Google is arguably the best known horizontal search engine. We should note, however, that vertical search engines still play an important role today. Consider an example where one uses Google, or an alternative horizontal search engine such as Yahoo! or (Microsoft’s) Bing, to locate the Web site for [Bates College](#). Once the user has successfully accessed the main page on the Bates site she can then use Bates’ local search facility, a vertical search engine, to retrieve information about faculty and staff who work at that college, or retrieve information about various academic programs and co-curricular activities sponsored by that college, and so forth. Within that vertical search engine, however, the user cannot retrieve broader information about faculty and academic programs at related colleges and universities or about related topics in general (as they could when using a horizontal search engine).

Another type of Web search engine is a *meta search engine*, which, as its name suggests, draws from the results of multiple (specialized) search engines and then combines and re-ranks the results. One of the first, and perhaps most popular, meta search engines in the mid-to-late 1990s was HotBot (Wall 2011). Meta search engines had a much more important role to play during the early years of the Web. As search engines improved and became more sophisticated, the need for meta search dramatically declined. Today, most general purpose (horizontal) search engines, such as Google and Bing, are able to return the same level of ranked results (as meta search engines once did), via their aggregation schemes. In fact, the founders of Google have described their search engine as an “aggregator of information” (Brin and Page 1998).

Some of the technological breakthroughs that occurred during the “early Web” era of search engine development helped to make possible two kinds of privacy-related ethical issues examined in [Section 3](#). First, the vast amount of online information about ordinary people that became accessible to Web-based search engines during this era made it possible for those people to become the “targets” of online searches conducted by anyone who had access to the Internet; this concern is examined in [Section 3.2](#). Second, the practice of aggregating personal information, which was being routinely collected by major search engine companies and their advertisers, contributed significantly to the data-mining-related privacy issues that are examined in [Section 3.3](#).

Some privacy issues introduced w/ technological breakthroughs in the early web

2.4 The “Web 2.0” Era (2000–Present)

Although the expression “Web 2.0” is now commonly used to differentiate the current Web environment from the early Web, critics point out that this expression is somewhat vague or imprecise. Whereas the early Web (sometimes referred to as “Web 1.0”) has been described as an online environment that was essentially passive or static, in so far as one could simply view the contents of a Web site that had been set up by an organization or an individual (e.g., when one visited someone’s “home page”), Web 2.0 is more dynamic in that it supports many interactive or “participatory” features. In a Web 2.0 environment, for example, users

↓

can interact and collaborate with others in ways that were not previously possible. These collaborative features include [wikis](#) (with Wikipedia being the best known example), as well as [blogs](#) and [social networking](#) applications (such as Facebook and Twitter). Of course, the relevant question for us to consider is whether the Web 2.0 environment itself either changes or significantly affects the functions of search engines and, more importantly, the ethical issues they generate.

It is not clear whether we can accurately label current search engines as “Web 2.0 search engines,” even though they operate in a Web 2.0 environment. For example, many of the participatory tools and functions that apply to applications such as social networks, blogs, and wikis do not necessarily apply to contemporary search engines. So, it may be more appropriate to use Hinman’s phrase “second-generation search engines.” However, O’Reilly (2005) suggests that Google’s practice of incorporating user-generated content to provide a “better” Web search environment for users is compatible with interactive dimensions and objectives of Web 2.0. But despite O’Reilly’s interpretation of Google’s practices, one might still question whether the phrase “Web 2.0 search engines” is warranted; so, we will instead refer to contemporary (or second-generation) search engines as “Web 2.0-era search engines.”

What, exactly, distinguishes a Web 2.0-era search engine from the earlier ones? Hinman notes that the traditional criteria Web search engine companies used to rank sites was based on two factors: (1) the number of visits to a page (i.e., “popularity”), and (2) the “number of other pages that link to a given page.” With respect to the second criterion, Diaz (2008) and others point to an analogy used in ranking the importance of academic papers. They note, for example, that an academic paper is generally viewed to be important if it is cited by many other papers. And that paper is perhaps viewed as even more important if it is cited by highly cited works. Hinman believes that [the shift to \(what we call\) Web 2.0-era search engines occurred when companies, such as Google, “looked more closely at what users wanted to find”](#) (which, as he also points out, is not always the most popular site). He notes, for example, that Google’s formula employs the following strategy: “Users’ needs → Search terms → Desired site” (Hinman 2005, 22). He also notes that in this scheme,

web 2.0-era search: towards personalization

[what the user wants becomes an integral part of the formula](#), as does the set of search terms commonly used to express what the user wants.

Hinman and others credit Google’s success as the premier contemporary Web search engine to the company’s proprietary algorithm, called PageRank.

Zimmer (2008, 77) believes that Google’s ultimate goal is to “create ‘the perfect search engine’ that will provide only intuitive, personalized, and relevant results.” Halpern (2011) points out that the search process has already “become personalized”—i.e., “instead of being universal, it is idiosyncratic and oddly peremptory.” And Pariser (2011), who asserts that “there is no standard Google anymore,” also notes that with “personalized search,” the result(s) suggested by Google’s algorithm is probably the best match for the search query. Some ethical implications affecting the personalization of search algorithms are examined in [Section 3.4](#).

Before concluding this section, we should briefly mention two additional kinds of contemporary search engines. First, [DuckDuckGo](#) (an alternative to Google and Bing) which prides itself on not profiling its users, claims that it does not personalize search results. So, this search engine’s users will typically get the same search results, as well as the same ranking of results, for a given query. Second, a new type of search engine, capable of “discovering” objects or “things” on the Internet is now also available. Two examples of this kind of “discoverable search engine,” Shodan and Thingful, are examined in the final section of this

entry.

3. Ethical Implications and Core Ethical Issues

Most Internet users are well aware of the virtues of search engines. As we noted earlier, many of us now depend on them to direct us to information that affects nearly all facets of our day-to-day lives—information about work, travel, recreation, entertainment, finances, politics, news, sports, music, education, and so forth. However, as we also noted earlier, the use of search engines has generated a cluster of ethical concerns. In Section 3, we organize these concerns into five broad categories: (i) search-engine bias and the problem of opacity/non-transparency, (ii) personal privacy and informed consent, (iii) monitoring and surveillance, (iv) censorship and democracy, (v) (cyber)security issues, “discoverable search engines,” and the Internet of Things. A different, but also related, cluster of ethical issues — viz, questions concerning moral accountability and social responsibilities for search engine companies — are examined in Section 4.

5 categories of ethical issues

3.1 Search Engine Bias and the Problem of Opacity/Nontransparency

What is search-engine bias, and why is it controversial? In reviewing the literature on this topic, it would seem that the phrase “search-engine bias” has been used to describe at least three distinct, albeit sometimes overlapping, concerns: (1) search-engine technology is not neutral, but instead has embedded features in its design that favor some values over others; (2) major search engines systematically favor some sites (and some kinds of sites) over others in the lists of results they return in response to user search queries; and (3) search algorithms do not use objective criteria in generating their lists of results for search queries.

3.1.1 The Non-Neutrality of Search Engines

While some users may assume that search engines are “neutral” or value-free, critics argue that search engine technology, as well as computer technology in general, is value-laden and thus biased because of the kinds of features typically included in their design. For example, Brey (1998, 2004) and others (see, for instance, Friedman and Nissenbaum 1996) have argued that computer technology has certain built-in features that tend to favor some values over others. Brey worries that some of these technological features have embedded values that are “morally opaque.” Because the values embedded in these features are not always apparent to the technical experts who develop computer systems, Brey believes that a methodological framework, which expands upon what he calls the “standard” applied-ethics model typically used in “mainstream computer ethics,” is needed to identify or disclose the “hidden” values at the design stage. He refers to this model as “disclosive computer ethics” (Brey 2004, 55–56).

Identifying the human values embedded in technological design and development has been the main objective of a movement called *Value Sensitive Design* or *VSD*, which Friedman, Kahn and Borning (2008, 70) define as a

theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process.

Friedman et al. illustrate their model using the example of Internet cookies—i.e., small text files that a Web browser places on a user’s computer system for the purposes of tracking and recording that user’s activities on a Web site. In particular, they examine the design of cookies in connection with the informed-consent

process vis-à-vis Web browsers. They further argue that embedded features in the design of cookies challenge **the value of informed consent** and that this value is important because **it protects other values such as privacy, autonomy, and trust**.

Cookies technology is not only embedded in the design of contemporary Web browsers, it is also used by major search engine companies to acquire information about users. In so far as these companies place cookies on users' computer systems, without first getting their consent, they also seem to contribute to, and perhaps even exacerbate, at least one kind of technology-related bias—i.e., one that threatens values such as privacy and autonomy, while favoring values associated with surveillance and monitoring. However, since this kind of bias also applies to design issues affecting Web browsers, it is not peculiar to search engines per se.

3.1.2 The Manipulation of Search Results

Some critics, including Introna and Nissenbaum (2000), tend to view the schemes used to manipulate search results as the paradigm case of bias in the context of search engines. In their highly influential paper on this topic, Introna and Nissenbaum argued that search engines

systematically exclude certain sites and certain types of sites, in favor of others, systematically giving prominence to some at the expense of others.

There has been considerable speculation as to why this is the case, but we briefly examine two reasons that have been prominent in discussions in the literature: (a) the interests of advertisers who sponsor search engines; and (b) schemes used by technically-savvy individuals and organizations to manipulate the ordered ranking of sites returned by search engines. (A third reason that is also sometimes suggested has to do with the nature of the algorithms used by major search engine companies; that issue is examined in this [Section 3.1.3](#)).

3.1.2.1 Online Advertising Strategies and Search Bias

Some critics, including Hinman (2005) and Noble (2018), point out that search engine companies are “answerable” to the paid advertisers who sponsor them. So, for many of these critics, bias-related concerns affecting the inclusion/exclusion of certain sites can be attributable mainly to the interests of paid advertisers. Google founders Brin and Page (1998, 18), who initially opposed the idea of paid advertisement on search engines, noted that it would seem reasonable to

expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of consumers... Since it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious...[and] less blatant bias are likely to be tolerated by the market.

It is worth noting that advertising schemes used by search engines have evolved over time. For example, Diaz (2008, 21) points out that banner ads, which were common on the Internet during the Web 1.0 era, have been replaced by “paid placement of ads.” He also notes that “paid listings” (unlike the earlier banner ads) do not always look very different from the normal results returned to users. Elgesem (2008) points out that search engines such as GoTo, whose lists of search results were based entirely on “paid hits” from advertisers, allegedly failed because of user dissatisfaction with the results they received. Eventually,

however, GoTo was taken over by Google, which continued to use GoTo's method for generating paid-ad-based search results but physically separated those results from the "organic" results that appear on the center of its pages (Elgesem 2008); this scheme, which contrasts the two different kinds of results, seems to have been accepted by Google users.

Diaz describes two other kinds of bias-related concerns that affect advertising schemes used by search engine companies: (i) the arbitrary (and seemingly inconsistent) criteria used by these companies in accepting advertisements, and (ii) the imprecise (and sometimes confusing) criteria used to separate editorials (approved by a search engine company) from their paid advertisements. (Although the kind of discrimination with regard to the ads that Google accepts or rejects may seem arbitrary in an innocuous sense, Diaz notes that Google has also refused to accept ads from some organizations that have been critical of corporations that were already sponsors for Google.) Regarding (ii), Diaz explains how bias-related concerns affecting paid advertisements can sometimes be confused with editorials that are also displayed by search engine companies on their Web pages. Whereas editorials and ads may look very different in newspapers, Diaz notes that this is not always the case in search engines. (We also examine some aspects of search engine bias in the context of online advertisements in our analysis of moral-accountability issues for search engine companies in Section 4.

Some critics assume that as conflicts affecting online advertisements in the context of search engines are eventually resolved, the amount of bias in search engine results will also decline or perhaps disappear altogether. However, other schemes can also be used to influence the privileging of some sites over others, in terms of both their inclusion (vs. exclusion) and their ranking.

3.1.2.2 Technological Schemes Used to Manipulate Search Results

We have already noted that some technically-savvy individuals and organizations have figured out various strategies for "gaming the system"—i.e., positioning their sites higher in the schemes used by search engine companies to generate results (see, for example, Goodwin 2018). These schemes are commonly referred to by "insiders" as instances of SEO, or Search Engine Optimization, and we briefly consider what can now be regarded as a classic SEO ploy. Some organizations and individuals had effectively used HTML meta tags and keywords (embedded in HTML source code) to influence search engine companies to give their sites higher rankings in the ordering schemes for their respective categories of search. Eventually, however, search engine companies recognized the manipulative aspects of these HTML features and began to disregard them in their ranking algorithms (Goldman 2008).

Many organizations now use a different kind of strategy to achieve a higher ranking for their sites—one that takes advantage of the (general) formulas currently used by major search engine companies. Blanke (2005, 34) notes that in Google's algorithm, Web pages "achieve a better ranking if they optimize their relationship within the system of hubs and authorities." Whereas "authorities" are Web pages that are "linked by many others," hubs "link themselves to many pages." Diaz (2008) points out that highly referenced hubs will have the highest Page Ranks. So Web site owners and designers who know how to exploit these factors (as well as how to use various SEO-related schemes) to manipulate ranking in search engine results will have the highest ranked sites. Diaz also notes that these sites "tend to belong to large, well-known technology companies such as Amazon and eBay," while "millions of typical pages... will have the lowest ranking." Thus, it would seem that the kind of search engine bias identified by Introna and Nissenbaum that, whereby certain Web sites are systematically included/excluded in favor of others, will not necessarily be eliminated simply by resolving conflicts related to paid advertising in the context of search engine companies. In all

likelihood, organizations will continue to figure out ways to use SEO-related techniques to achieve a higher ranking for their sites and thus gain better exposure on search engines, especially on Google. As Hinman (2005) puts it, “Esse est indicato in Google (to be is to be indexed on Google).”

3.1.3 The Problem of Objectivity

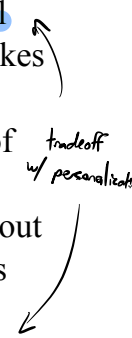
Concerns affecting search engine bias vis-à-vis questions having to do with “objectivity” have two distinct aspects: (A) objectivity regarding criteria used in search algorithms; and (B) objectivity with respect to the results returned by a particular search engine in its responses to multiple users entering the same search query. With regard to (A), we saw that the traditional criteria used by Web search engine companies to rank sites was based on two factors: the number of visits to a page, and the number of other pages that link to a given page; Hinman (2005) believes that this technique would seem to give the user “some semblance of objective criteria.” He points out, for example, that even if search engines were to “get it wrong” in returning the best results for a particular search query, there was an “objective fact of the matter to be gotten wrong.” And even though the early search engines ranked their sites in terms of “popularity,” there was, as Hinman (2005, 22) puts it, a “technical and objective meaning” of popularity. But we also saw that this formula has changed dramatically in the case of Web 2.0-era search engines, where increasingly “personalized algorithms” tend to tailor search results to fit the profile of the user entering the query. i.e., non-customized search

This trend toward the “personalizing” of algorithms feeds directly into concerns affecting (B). Even though many sophisticated users might suspect that the lists of returns for their search queries are biased, for any number of possible reasons—e.g., paid advertising, unfair influence by large corporations, editorial control, and so forth—many search engine users still tend to assume, perhaps naïvely, that when any two users enter the exact same search query in a major search engine such as Google, they would receive identical lists of responses. In other words, even if the formula used is skewed, or biased in a way that favors some sites over others, the search algorithm would nonetheless return results based on a formula that is internally consistent, and thus standard or “objective” in some sense. But, this is no longer the case in an era where formulas based on “personalization” generate search results tailored to a user’s profile. For example, if I enter the term “eagles” in a search box, the list and the order of returns that I receive will likely depend on the profile that the search engine company has constructed about me. If the company determines that I am interested in biology, for instance, I may be directed to a site sponsored by the Audubon Society. But, if instead, it determines that I am a sports enthusiast living in the Philadelphia area, I may be directed first to the Philadelphia Eagles Web site. On the contrary, if my profile suggests that I like rock/pop music, I may be directed first to the site for the Eagles music group. So there would not appear to be any overall “objective” formula used by the search engine in question. interesting example of search engine customization providing utility

Some question whether a lack of objectivity with respect to the results returned to users in response to their search queries is necessarily a problem. For example, Blanke (2005, 34) believes that we should “neither demand nor expect to receive information from search engines that is objective” (i.e., information that is either “neutral or complete”). So, he argues that any critique of search engines as being biased “misses its target,” because one should not expect search engines to deliver only neutral and objective results. His rationale for this claim, however, seems to be based on the view that search engine technology “was not designed to do this.” But this analysis seems to beg the question, unless, of course, Blanke means that search engine technology could not, in principle, be designed to deliver neutral and objective results.

Questions concerning objectivity in the context of search engines are also examined by Goldman (2008), who seems to defend—and, at times, perhaps even applaud—search engine bias in that respect. First, he

notes that search engine companies “make editorial judgments about what data to collect and how to present that data” (2008, 122). However, he also believes that search engine bias is “necessary and desirable”—i.e., it is “the unavoidable consequence of search engines exercising editorial control over their databases” (p. 127). So he is willing to concede that “search engine companies, like all other media companies, skew results.” But while many assume that search engine bias is undesirable, Goldman sees it as a “beneficial consequence of search engines optimizing content for their users.” He also believes that “the ‘winner takes all’ effect caused by top placement in search results will be mooted by emerging personalized search technology” (p. 121). He further argues that “personalized ranking algorithms” will “reduce the effect of search engine bias because there will likely be multiple ‘top’ search results of a particular search term instead of a single winner [and] personalized algorithms will eliminate many of the current concerns about search engine bias” (p. 130). Thus Goldman seems to suggest, paradoxically perhaps, that any problems affecting objectivity will be solved by increased *subjectivity* in the form of personalization of results achieved by personalized search algorithms. However, this direction in search engine evolution can have serious negative effects for democracy and democratic ideals (discussed further in [Section 3.4](#)).



Concerns affecting objectivity and bias in the context of search engines are also closely related to controversies pertaining to the lack of “openness” or “transparency” (see, for example, Noble 2018). Some critics point out that search engine companies are not fully open, or transparent, both with respect to *why* they (a) include some sites and not others (in their lists of results for users’ queries), and (b) rank some pages in their list of search results higher than others. These kinds of opacity/non-transparency-related concerns tend to fall under a description that Hinman (2005) calls “the problem of algorithm.” Hinman notes that the algorithms that govern searches are well-kept secrets, which he also believes is appropriate. Because Google’s PageRank algorithm is “a patented and closely guarded piece of intellectual property” (Halpern 2011, 4), we don’t know the algorithm’s formulas. And this factor, of course, makes it difficult to comment on an algorithm’s objectivity or lack thereof.

Another set of worries affecting opacity/non-transparency arise because search engine companies do not always disclose, either fully or clearly, their practices with respect to two important points: (i) whether (and to what extent) they collect information about users; and (ii) what they do with that information once it has been collected. These kinds of opacity/non-transparency concerns involving search engine companies, which are also related to privacy issues affecting monitoring and surveillance, are examined in detail in [Section 3.3](#).

3.2 Privacy, Consent, and Non-voluntary Disclosure of Personal Information

At least two distinct kinds of privacy concerns arise in the context of search engines. One set of privacy issues emerges because search engine companies can collect personal information about search engine users; in this scheme, the users are, in effect, “data subjects” for the search engine companies and their advertisers. However, search engine users themselves—whether acting on their own behalf or on behalf of organizations that hire them—can use the technology to conduct online searches about people. In this case, the targeted people (some of whom may never have used or possibly have never even heard of a search engine) are the subjects of search engine users. In both cases, privacy concerns arise in connection with questions about fairness for the data subjects involved. Consider that many of those who become the subjects of the search queries have not explicitly consented either to having certain kinds of personal information about them collected or having personal information about them (that has been collected in some other context) also being made available on the Web, or both.

In this section, we examine search-engine-related privacy concerns affecting people who have become the

subjects, or “targets,” of queries by search engine users. This kind of privacy concern is exacerbated by the ever expanding amount of personal information about ordinary people that is currently discoverable by search engines and thus accessible to Internet users. But why, exactly, is this problematic from the vantage point of privacy? For one thing, it is not clear that most people have voluntarily consented to having information about them placed in databases or in online forums that are accessible to search engines (and thus potentially available to any Internet user). And we noted that search engine users, whether they are acting simply on their own, or as representatives of business and corporations, can and often do access a wealth of information about many of us via search engines. Privacy advocates question whether this practice is fair, especially to people who have not explicitly consented to having personal information about them included in the online forums and databases that are now so easily searchable because of sophisticated search engine technology.

issue to do w/ ease of access

*Privacy
people don't
want consent
to this data
being made
available
online*

Privacy concerns that arise in contexts in which people are the subjects of search queries can be further differentiated in terms of two separate categories, i.e., where search engines are used to: (i) track the location of individuals, sometimes for the purpose of harassing or stalking them; and (ii) acquire personal information about people. We briefly examine each practice.

Regarding (i), one might ask why using search engines to track and locate persons is controversial from a privacy perspective. First, consider that some organizations have developed specialized search engines for questionable purposes, such as stalking people. For example, one search facility (Gawker-Stalker Maps, introduced in 2006) was designed specifically for the purpose of stalking famous people, including celebrities. Imagine a case in which a celebrity has been spotted while dining at an up-scale restaurant in San Francisco. The person who spots the celebrity can send a “tip” via text message or e-mail to Gawker-Stalker, informing the site’s users of her whereabouts. The Gawker site then provides its users, via precise GPS software, with information about exactly where, and at what time, she was last seen. Users interested in stalking that celebrity can then follow her movements electronically, via the Gawker site, or they can locate and follow her in physical space, if they happen to be in the same geographical vicinity as the celebrity at that time (Tavani 2016). Currently, Gawker-Stalker seems to be in a state of transition. Although gawker.com, a subsidiary of Gawker Media, ceased operations following a bankruptcy suit in 2016, the controversial Gawker site was acquired by the Bustle Digital Group in 2018 (Kelly 2019).

Second, we note that it is not only celebrities and “high-profile” public figures that are vulnerable to being stalked, as well as to having personal information about them accessed, via search engines. Consider the case of Amy Boyer, a twenty-year old resident of New Hampshire, who was stalked online by a former “admirer” named Liam Youens and who was eventually murdered by him in 1999. Using standard Internet search facilities, Youens was able to get all of the information about Boyer that he needed to stalk her—i.e., information about where she lived, worked, and so forth (see, for example, Tavani and Gridzinsky, 2002). Incidents such as the Boyer case invite us to question current policies—or, in lieu of clear and explicit policies, our default positions and assumptions—with regard to the amount and the kind of personal information about ordinary persons that is currently accessible to search engine users. It now appears likely that Amy Boyer had no idea that so much personal information about her was so easily accessible online via search engines.

We next examine (ii), the use of search engines to find information about people—not about their location, but about their activities, interests, and backgrounds. As in the Amy Boyer case, these search-engine-related privacy issues also arise when ordinary people become the subjects of search queries. Consider that, increasingly, employers use online search techniques to acquire information about prospective and current

employees. It is well known that many employers try to access information on the Facebook, Twitter, and Instagram accounts of job applicants they are considering. This kind of information has, in certain instances, been used by employers in determining whether or not to hire particular applicants (and possibly also used in deciding whether or not to promote current employees). So, for example, a college student who posts on Facebook one or more pictures of himself drinking alcoholic beverages, or perhaps behaving wildly at a party, can potentially jeopardize his future employment opportunity with a company that might otherwise hire him upon graduating from college. In defense of this kind of “screening” practice used by companies in hiring employees, one could argue that the company has merely elected to use currently available tools to search for information about persons who *voluntarily* posted material (e.g., in the form of photos, etc.) about themselves on Facebook.

Our primary concern here is with personal information that has not been voluntarily disclosed by persons, but is nonetheless accessible online via search engines. This kind of concern involving access to personal information in online forums is by no means new or recent. Consider that in the decade preceding Facebook, employers had been able to access information about job applicants via online search tools—e.g., they could (and did) use search engines to accomplish this task, simply by entering the name of the individual in a search engine box. Imagine a hypothetical scenario in which a person, Lee, applies for a full-time employment position at Corporation *X*. Also, imagine that someone on the corporation’s search committee for this position decides to conduct an online search about Lee, shortly after receiving her application. Further imagine that in response to the query about Lee, three results are returned by the search engine. One result includes a link to a gay/lesbian organization in which Lee is identified as someone who contributed to a recent event hosted by that organization. Next, imagine that Lee is turned down for the job at Corporation *X*. Further imagine that Lee becomes curious as to why she might not have been selected for that job and she decides to do an Internet search of her name for the first time. Lee then discovers the search result linking her to the gay/lesbian organization (Tavani 1998). Should Lee infer that she was denied the job because of her apparent association with this organization? Is that a reasonable inference? Maybe not. Nevertheless, an important question arises: Is it fair that someone has posted this information about Lee online, without her consent, to a source that is accessible to one or more search engines? Is that information about Lee now “fair game,” and should it be viewed simply as information that is “up for grabs” (Nissenbaum 2004) and thus appropriate for use by prospective employers?

How is the scenario involving Lee different from a case in which an employer uses information on a Facebook account to screen job applicants? For one thing, Facebook users typically post information about themselves that can be seen by others and thus have voluntarily consented to have that information available for others to access (assuming that they have not specified their Facebook privacy settings). Furthermore, they are also aware that such information about them exists on that online forum. But what about job applicants who do not have accounts on Facebook, or on any other social networking site? Are they less vulnerable to online scrutiny by potential employers? Many people may have no idea about either the kind or amount of online personal information about them that is accessible to an employer or to anyone using a search engine.

Next consider a hypothetical scenario similar to Lee’s, where Phil, who recently earned a Ph.D., is applying for a faculty position at University *X*. But suppose that a few disgruntled former students have posted some highly critical and negative remarks about Phil on RateMyProfessor.com. Next, suppose that a member of the faculty search committee at University *X* conducts an online search on Phil and discovers the disparaging remarks made by the students. Finally, Phil is informed that he has not been selected for the faculty position. Shortly after receiving his letter of rejection, Phil happens to discover the comments about him made by the

disgruntled students, by conducting an online search of his name. Would it be unreasonable for Phil to infer that the remarks by these students on RateMyProfessor.com influenced the hiring committee's decision not to select him?

In one sense, Phil's predicament is very similar to Lee's—viz., neither job applicant had any kind of control over what people other than themselves had posted about them in online forums accessible to search engines. However, the negative information posted about Phil was directly related to the kind of criteria that typically would be used in considering an applicant for a faculty position. The information posted about Lee, while not directly related to the job for which she was applying, could nonetheless also harm her chances of gaining that position. In neither case, however, did Lee or Phil have any say about the kind of information about them, or about the accuracy of that information, that could be so easily retrieved online and used by a prospective employer in making a decision about whether or not to hire them.

the factor of
'relevant criteria' in this
thought experiment

On the one hand, we can ask what kind of recourse people like Phil and Lee could expect to have in situations such as this—e.g., can they reasonably expect to have control over any kind of information about them that is currently accessible to search engines? But, on the other hand, it may not seem totally unreasonable for them to have some expectation of limited control over their personal information, even if only to be able to challenge the legitimacy of inaccurate information, especially when they had not consented to having it included in online forums and databases accessible to search engines.^[2] There is also an aspect of this kind of personal information that overlaps with “public” information. So, perhaps the tension that arises in these scenarios can be viewed as a contemporary variation of the age-old debate about the private vs. public nature of personal information. This tension is further complicated by the fact that in the U.S. most people, as the subjects of online searches, enjoy little, if any, normative protection regarding personal information about them that is now available online—mainly because of the presumed “public nature” of this personal information involved (Tavani 2005). (As we will see in Section 4.2.1, however, citizens in EU countries enjoy much more normative protection of their personal data in online contexts than U.S. citizens.)

Some forms of personal information enjoy normative protection via specific privacy policies and laws, because they qualify as information about persons that is considered either sensitive or intimate, or both. We can refer to this kind of personal information as Non-Public Personal Information (or NPI). However, many privacy analysts now worry about the ways in which a different kind of personal information—Public Personal Information (or PPI), which is non-confidential and non-intimate in character—is easily collected and exchanged over the Internet. How can PPI and NPI be distinguished? NPI, which as noted above, is viewed as information about persons that is essentially confidential or sensitive in nature, includes information about a person's finances and medical history. PPI, although also understood as information that is personal in nature, is different from NPI in at least one important respect: it is neither sensitive nor confidential. For example, information about where an individual works or attends school, as well as what kind of automobile he or she owns, can be considered personal information in the sense that it is information about *some individual as a particular person*. However, this kind of personal information typically does not enjoy the same kinds of privacy protection that has been granted to NPI (Tavani 2016).

Initially, concerns about personal information that can be gathered and exchanged electronically focused mainly on NPI. In response to these concerns, some specific privacy laws and policies were established to protect NPI. But many privacy advocates now also worry about the ways in which PPI is routinely collected and analyzed via digital technologies. They have argued that PPI deserves greater legal and normative protection than it currently has. Nissenbaum (1997, 1998) has referred to the challenge that we face with

regard to protecting (the kind of information that we refer to as) PPI as the “problem of protecting privacy in public.” Some privacy advocates argue that our earlier assumptions about what kinds of publicly available information about us need explicit legal protection (or perhaps some kind of less formal “normative” protection, at the level of policies) are no longer adequate because of the way much of that information can now be processed via digital technologies, especially in the commercial sphere. For example, seemingly innocuous information about persons, based on their activities in the public sphere, can be “mined” to create user profiles based on implicit patterns in the data and those profiles (whether accurate or not) can be used to make important decisions affecting people.

old assumptions & norms
don't hold up in digital context

3.3 Monitoring and Surveillance

We next examine privacy concerns in which search engine users themselves are the data subjects (i.e., for search engine companies). Zimmer (2008, 83) notes that personal information about users is “routinely collected” when they use search engines for their “information-seeking activities.” But why, exactly, is this problematic from the perspective of privacy? For one thing, search engine companies such as Google create a record of every search made by users, and these records are also archived. The topic searched for, as well as the date and time the specific search request is made by a user, are included in the record. Until recently, many people had been unaware that their search queries were being recorded and tracked.

In January 2006, many Google users learned that the search engine company had kept a log of all of their previous searches. At least four major search engine companies had been subpoenaed by the Bush Administration in 2005 for search records based on one week’s of searches during the summer of 2005 (see, for example, Nissenbaum 2010). They were Google, Yahoo, AOL, and Microsoft (MSN). Google refused to turn over the information. The other search engine companies would not say how they responded; many assumed, however, that those companies complied with the government’s subpoena. Google was requested to turn over two items: (1) the results of search queries/requests it received during a one-week period, and (2) a random list of approximately one million URLs searched. Google argued that turning over this information would: (a) violate the privacy of its users (and undermine their trust in the search engine company), and (b) reveal information about the (proprietary) algorithm and the processes that Google uses and that this would potentially harm its competitive edge as a search service. A court ruled that Google did not have to comply with (1), but it reached a compromise regarding (2), ruling that Google turn over 50,000 URLs to the government (Nissenbaum 2010, 29–30).

not the same as trust in the engine itself

The information collected about a user’s search queries might seem relatively innocuous—after all, who would be interested in knowing about the kinds of searches we conduct on the Internet, and who would want to use this information against us? On the other hand, however, seemingly innocuous personal information can be mined by information merchants and used to construct personal profiles about us, and that these profiles, in turn, can be based on information that is not accurate and can be used to make decisions about us that are not fair. For example, imagine a case in which a student happens to be writing a paper on Internet pornography and uses a search engine to acquire some references for her research. Records of this user’s search requests could reveal several queries that individual made about pornographic Web sites, which in turn might suggest that this user is interested in viewing pornography. So individual searches made by a particular user could theoretically be analyzed in ways to construct a profile of that user that is inaccurate. And, records of the searches made by this and other users could later be subpoenaed in court cases (Tavani 2016).

As already noted, information about a user’s search queries is collected by search engine companies as well

as by many different kinds of “information merchants” in the commercial sphere. Halpern (2011, 8) notes that there are approximately 500 companies that are able to track all of our online movements, thereby “mining the raw material of the Web and selling it to...data mining companies.” Pariser (2011) points out that in tracking our activities, Google’s uses fifty-seven *signals*—

everything from where you were logged in, from what browser you were using, to what you had searched for before to make queries about, to who you were and what kinds of sites you’d like.

And Zimmer (2008, 77) notes that Google integrates information gathered from

Web cookies, detailed server logs, and user accounts... [from Google applications such as Gmail, Google +, and Google Chrome]... which provides a powerful infrastructure of dataveillance to monitor, record, and aggregate users’ online activities.

Furthermore, Pariser notes that Google and other major search engine companies use “prediction engines” to construct and refine theories about who we are (and what we want to do next). He also notes that many information merchants regard every “click signal” a user creates as a “commodity” that can be “auctioned off within microseconds to the highest bidding consumer.” Pariser points out that one information merchant, a company called Acxiom, has

accumulated an average of 1500 pieces of data on every person in its database—personal data that ranges from credit scores to medications used.

Of course, some users might respond that they do not feel threatened by this practice; for example, they might be inclined to feel safe from a loss of personal privacy because they assume that the data collected about them is anonymous in the sense that it is identifiable only as an IP address, as opposed to a person’s name. However, Zimmer (2008, 83) notes that in 2005, one AOL user was able to be identified by name “because the Web searches she performed on various topics were recorded and later released by AOL.” It turns out that AOL Research had released over three months worth of personal search data involving 650,000 users (Wall 2011, 18). Nissenbaum (2010, 30) points out that in this case, AOL used a process in which

certain identities could be extracted from massive records of anonymized search-query data that AOL regularly posted on the Internet for use by the scientific research community.

Zimmer believes that the incident involving AOL is not unique, but is instead one more use of data surveillance or “dataveillance” (a term coined by Roger Clarke in 1988)—i.e., one applied in the context of search queries.

difference b/w surveillance and monitoring

It is also important to consider whether a meaningful distinction can be drawn between monitoring and surveillance in this context. Noting that the two terms are often used interchangeably, Nissenbaum (2010, 22) differentiates between them in the following way. Whereas **surveillance** is a “form of monitoring ‘from above’ by political regimes and those in authority,” **monitoring** is used in broader social and “socio-technical” contexts. In Nissenbaum’s scheme, both monitoring and surveillance are examples of what she calls “socio-technical contexts,” but they are usually put to different uses. For example, Nissenbaum points out that monitoring can be done by systems “whose explicit purpose is to monitor” (e.g., CCTVs). But she also notes that **information itself can constitute a “modality of monitoring.”** For example, she points out that Clarke’s notion of “dataveillance” includes monitoring practices that involve both interactions and

transactions. However, we can generally regard the kinds of practices carried out by information merchants in the consumer sphere as instances of monitoring (rather than surveillance) in Nissenbaum's sense of that term.

We next shift our focus away from privacy concerns about monitoring in the commercial sector to worries about *surveillance* by government actors with respect to information acquired as a result of users' search queries. Earlier in this section we noted that in 2005, the Bush Administration informed Google that it must turn over a list of all users' queries entered into its search engine during a one week period (the exact dates were not specified by Google). The Bush Administration's decision to seek information about the search requests of ordinary users triggered significant criticism from many privacy advocates. Although the Bush Administration claimed that it had the authority to seek electronic information in order to fight the "war on terror" and to prevent another September 11-like attack, some critics worried that the government was trying to use the subpoenaed information, not for national defense or anti-terrorism purposes, but rather to gain data to support its stance on the Child Online Protection Act, which had been challenged in a U.S. District Court and was being revisited by Congress (Nissenbaum 2010, 29). These critics also worried about the implications this has for privacy (as an important human value) in the ongoing tension involving security vs. privacy interests. And even if privacy is not an absolute value but is sometimes outweighed by security concerns, as Himma (2007) argues, some critics question the rationale used for obtaining records of search requests made by ordinary citizens.

Hinman (2005) notes that the Patriot Act, passed in the aftermath of 9/11, allowed U.S. government officials to get information from libraries about which books members had borrowed. He then shows how the reasoning used in the case of libraries could easily be extended to search engines—for example,

if the government could see which books someone was taking out of a library, why couldn't it also see which searches we made on search engines?

Hinman also points out that there are several other ways in which a user's search requests can be disclosed because of practices used by major search engine companies such as Google. He also worries that such practices could eventually lead to surveillance and to suppressing political dissent (as is it has in China). Hinman questions whether Google might have been under political pressure from outside interests (e.g., the Bush Administration) to take down photographs of tortured prisoners in the controversial Abu Ghraib detention center (used by the U.S. during the Iraq War in 2004), which were posted but then soon removed with no apparent explanation by that search engine company. (Some related, as well as additional, questions concerning moral-responsibility-related issues for search engine companies are examined in detail in Section 4.)

3.4 Democracy, Censorship, and the Threat to Liberty and Freedom

In this section, we consider some implications that the surveillance of users' queries by search-engine companies can have for a free and open society. In the early days of the Internet, many people assumed that search engine technology favored democracy and democratic ideals. For example, Introna and Nissenbaum (2000, 169) note that search engines were viewed as a technology that would

...give voice to diverse social, economic, and cultural groups, to members of society not frequently heard in the public sphere [and] empower the traditionally disempowered, giving them access both to typically unreachable modes of power and to previously unavailable troves

of information.

However, Introna and Nissenbaum also describe what can be viewed as an “anti-democratic” aspect of contemporary search technology when they note that search engines “systematically exclude” certain Web sites, as well as “certain types of sites,” over others. And Diaz (2008, 11) echoes this concern when he notes that that major search engine companies such as Google direct “hundreds of millions of users towards some content and not others, towards some sources and not others.” So, following Diaz (p. 15), we can ask whether the kinds of “independent voices and diverse viewpoints” that are essential for a democracy are capable of being “heard through the filter of search engines.” *big question about essential properties of democracy in a digital context* *and, why search engines in particular are important here*

Search engines have often been described as (the “gatekeepers of cyberspace,” and some critics note that this has significant implications for democracy. For example, Diaz (2008, 11) points out that

if we believe in the principles of deliberative democracy—and especially if we believe that that the Web is an open ‘democratic’ medium—then we should expect our search engines to disseminate a broad spectrum of information on any given topic.

Hinman (2005, 25) makes a similar point, when he notes that “the flourishing of deliberative democracy is dependent on the free and undistorted access to information.” And because search engines are “increasingly the principal gatekeepers of knowledge,” Hinman argues that “we find ourselves moving in a philosophically dangerous position.” (We briefly return to Hinman’s point in the [concluding section](#) of this entry.)

Morozov (2011) also describes some concerns for democracy vis-à-vis contemporary search engines by calling attention to the filtering of information that search engines make possible. For one thing, he agrees with Sunstein (2001) who worries that the kind of selectivity made possible by Internet filtering can easily trap us inside our “information cocoons.” And Lessig (2000) suggests that any kind of filtering on the Internet is equivalent to censorship because it blocks out some forms of expression. Morozov points out that whereas Sunstein worries that people *could* use Internet technology to “overly customize what they read,” the reality is that contemporary search engine companies have already silently done this for them. Morozov’s concerns about what search engine companies are now doing through filtering and customization schemes, and why this is problematic for a democracy, are echoed by Pariser (2011, 13) who points out that “personalization filters serve up a kind of invisible autopropaganda, indoctrinating us with our own ideas, amplifying our desire for things that are familiar and leaving us oblivious to the dangers lurking in the dark territory of the unknown”. *personalization as censorship/propaganda* *→ and thus anti-democratic*

Pariser notes that while democracy “requires citizens to see things from one another’s point of view,” we are instead increasingly “more enclosed in our own bubbles.” He goes on to note that democracy also “requires a reliance on shared facts,” but instead we are being presented with “parallel but separate universes.” To illustrate how this trend away from citizens having shared facts can be so dangerous for a democracy, Pariser uses the example of the debate about climate change in the U.S during the past decade. He points out that studies have shown that between 2001 and 2010 many people’s beliefs about whether the climate was warming shifted significantly, based on one’s affiliation with a major political party. Pariser notes that a Web search for “climate change” will yield very different results for a person whom the search algorithm determines to be a Democrat than for someone it determines to be a Republican. He also notes that the search algorithm will generate different results for someone it determines to be an oil company executive vs. an environmental activist.

Along lines similar to Pariser’s, Halpern (2011, 5–6) notes that search engines like Google direct us to

material that is most likely to reinforce our own “worldview, ideology, and assumptions” and thus “cut us off from dissenting opinion and conflicting points of view.” Pariser points out that with Google, a user now gets the results that the search engine company believes are best for that particular user. He describes a case where two people entered in the keyword “BP” (for British Petroleum) during the time period of the accident involving the Deep Water Horizon oil rig in the Gulf of Mexico. In response to one user’s query, investment information about BP was returned as the lead result, while the other user received information about the oil spill.

Not only do some practices by search engine companies pose a threat for democracy and democratic ideals, other practices (in which search engine companies are arguably complicit) reinforce censorship schemes currently used by non-democratic nations. For example, it is well known that China has succeeded in blocking access to political sites that it regards as threatening. Consider that it blocks ordinary Chinese users from access to sites such as “Tiananmen Square,” “Free Tibet,” and “Dalai Lama.” Critics note that Google agreed to comply with China’s censorship laws when the search engine company entered the Chinese market in 2006. Spinello (2012) believes that this agreement violated Google’s “don’t be evil” principle—a core principle of the search engine company—because Google “facilitated and supported” China’s censorship regime. But some of Google’s defenders have argued that the search engine company’s entry into China made it possible for China’s residents to have greater access to information, overall (i.e., beyond what would otherwise be accessible through Baidu (www.Baidu.com), a Chinese-owned search engine service founded in 2000). Other defenders of Google point out that the search engine giant did not act alone because major U.S. companies, such as Yahoo and MSN, also complied with China’s censorship laws. And Hinman notes that the Chinese government also received cooperation from other American companies, such as Cisco Systems, in establishing the infrastructure or backbone components of its Internet firewall.

In 2010, Google changed its policy for operating in China and directed its Google.cn users to a site in Hong Kong that was then uncensored. However, Hinman believes that we should still worry about Google’s willingness to comply with the Chinese government’s strict censorship laws when initially setting up its business operations in China. He reasons that if this search engine giant could be so easily influenced by a government that has a relatively low economic impact on its business overall, it could be much more influenced by the U.S. Government, where the political and economic impact would be far more significant. Some worry that this factor has also given Google considerable power over companies that rely on it for their Internet traffic (Spinello 2012). They also worry that Google, in an effort to retain that economic power in the U.S. in the future, could conceivably cave into pressure to comply with government policies (in the U.S. and possibly other democratic nations as well) that might support censorship at some level.

While many initially believed that that the Internet would promote democracy by weeding out totalitarian societies, because they are “inefficient and wasteful” (Chorost 2011), Berners-Lee (2010) believes that the Web “we have come to know” is now threatened because both totalitarian and democratic governments alike are “monitoring people’s online habits, endangering important human rights.” Perhaps Hinman (2005, 25) best sums up this worry, when he remarks,

We risk having our access to information controlled by ever-powerful, increasingly opaque, and almost completely unregulated search engines that could shape and distort our future largely without our knowledge. For the sake of a free society, we must pursue the development of structures of accountability for search engines.

constellation of dangerous properties

* tangential *: trust as the unquestioning attitude (or maybe generally) important for securing freedom to operate w/in opaque systems...?

If Berners-Lee, Hinman, and others are correct, it would seem that we have much to worry about indeed as we go forward trying to preserve our basic freedoms in a democracy, while at the same time taking

advantage of many of the personalizing- and customizing-based features that search engines such as Google now offer us. It would also seem that increased transparency on the part of search engine companies could be an important step in helping us to alleviate some of these concerns. But **who, exactly, should be responsible for regulating search engine companies and for holding them accountable?** We examine these and related questions in Section 4. First, however, we briefly consider some ethical issues, particularly cyber-security related concerns, that arise because of a relatively new kind of search engine that is capable of “discovering” the location of “things” (objects) via the Internet.

3.5 (Cyber)Security and the Internet of Things

In Sections 3.1 through 3.4, we examined a wide range of ethical concerns affecting the use of search engines, focusing mainly on what can be viewed as the “standard” or received ethical concerns that have come to be associated with search engines during the past three decades. One topic that arguably has been neglected, or perhaps considered only indirectly (e.g., in our earlier discussion of privacy-related issues in Sections 3.2 and 3.3), has to do with **ethical concerns affecting (cyber)security**. In addition to data security, which is closely related to data privacy (described earlier), issues associated with system and network security, as well as with national and international security, also now arise in light of a specific kind of search engine that is capable of locating “things.” In this section, we briefly consider some emerging security-related challenges generated by the use of this type of search engine in the context of the Internet of Things (IoT).

What is IoT, and how does it raise security-related ethical concerns affecting search engines? IoT generally refers to the interconnection of “objects” on the Internet; these (networked and “intelligent”) objects can include webcams, printers, and other devices that are not easily identifiable, locatable, or accessible to users via conventional search engines (see, for example, Burgess 2018). So users might assume that since information about these objects would not be accessible via typical Internet searches, the objects would be fairly secure from network hackers. Many users might also be enamored by the kinds of conveniences they could enjoy if their “intelligent” objects were networked together. Consider, for example, a “smart home” where these objects could communicate with one another as well as with that homeowner’s mobile devices (and possibly also communicate directly with networked devices and applications embedded in the homeowner’s automobile). On the one hand, we can imagine that the owner of this home would be delighted if its “intelligent” refrigerator communicated with her while driving home to alert her that the refrigerator’s supply of milk is low. The refrigerator might also communicate directly with the homeowner’s smart phone to trigger an app designed to display the lowest prices for milk at nearby grocery stores. So, ordinary users could easily come to enjoy the kinds of conveniences made possible via their interconnected “intelligent” objects. On the downside, however, it is likely that these objects may not have an appropriate level of network security. Thus users could be vulnerable to having their objects hacked and also be subject to having personal information about them (vis-à-vis their interactions with these objects) acquired by unauthorized individuals and organizations, if search engines were indeed capable of discovering and locating their (non-sufficiently-secure) objects and devices.

Whereas traditional search engines have assisted users in locating online information pertaining to the names of people and places, commercial/governmental/educational organizations and institutions, etc. (via a standard protocol involving HTTP addresses and HTML programming code), some search engines are now also capable of searching the Internet for certain kinds of “things” (or objects). Two of these search engines are *Thingful* and *Shodan*. Thingful describes itself as a “discoverable search engine” that provides users with “a unique geographical index of connected objects around the world” (<https://thingful.net/>). As such,

Thingful boasts that it can index across multiple IoT networks and infrastructures. Because this search engine can locate the geographical position of objects and devices, some critics worry that Thingful can also easily access personal data regarding the ways in which users interact and communicate with their connected devices and objects (as in the case of a homeowner communicating with the “intelligent” objects in her smart home).

Shodan, a controversial search engine created by John Matherly in 2009, enables users to find specific types of (Internet-connected) computers, devices, and systems (including servers and routers), via globally located servers that index the internet continuously. Shodan has been described both as a “dark Google” and the “scariest search engine ever” because it could be used to locate components in a nation’s critical infrastructure as well as its military defense system, including the “command and control systems for nuclear power plants and particle-accelerating cyclotrons” (Goldman 2013). Some have even suggested that Shodan was used to locate and monitor major components in Iran’s controversial nuclear program (Charette 2012). This has also led to speculation that Shodan may have assisted the (then secret) “Olympic Games Operation” allegedly carried out by the U.S. and Israel to cause Iran’s centrifuges—i.e., fast-spinning machines that enrich uranium—to spin out of control (O’Harrow 2012). While some might be inclined to argue that nations such as the U.S and Israel would have been justified in using whatever cyber-related means they had at their disposal to disrupt the progress of Iran’s nuclear program, we can clearly see the downside of such a practice; i.e., it could also be adopted by so-called “rogue nations” and used against “legitimate” nation states.

So it would seem that “discoverable search engines” like Shodan and Thingful pose threats not only to our personal privacy and security, but also to the security of a nation’s critical infrastructure and military defense systems as well. It will be interesting to see whether this relatively new kind of search engine will be used by nation states, and possibly even by terrorist groups as well, in their future strategies and practices affecting cyberwarfare, cyberterrorism, and cyberespionage. It will also be interesting to see whether these search engines, and the companies and organizations that develop them, will need to be closely regulated because of the significant societal threats they now pose. We next examine a related question as to whether search engine companies in general have some special social/moral responsibilities, in light of their important societal roles.

4. Moral Accountability and Social-Responsibility Issues for Search Engine Companies

Thus far, issues from the perspectives of business ethics and professional responsibility have not been directly addressed in this entry; our main focus in Section 3 has been on the kinds of moral impacts that search engines have for ordinary users, especially with respect to privacy, surveillance, and freedom. In Section 3.1, however, we hinted that search engine companies (SECs) might be held morally accountable for their practices involving bias in the ranking of search results. And in concluding Section 3.4, we briefly entertained the notion that major SECs might be held accountable for policies and practices that either favor or directly support censorship. In this section, we consider whether SECs have any special social responsibilities and moral obligations because of their “privileged place” in our society. We begin by briefly describing a key societal role that search engines play in providing access to information and knowledge.

Elgesem (2008, 241) notes that search engines have an important societal role as “contributors to the public use of reason”, and Noble (2018, 148) argues that search results can actually “structure knowledge” for the

*part of the ‘radically non-personal epistemology’
Nguyen examines towards the end of ‘Unquestioning Attitude’?*

societies within which they operate. As we already noted in Section 3, Hinman (2005, 2008) and Diaz (2008) view search engines as “gatekeepers” of knowledge on the Web. Taddeo and Floridi (2016, 1575), who also describe the important “gatekeeping function” that SECs like Google now have because of their “centrality” in information societies, worry about the lack of consensus thus far with regard to which “principles should shape...[an SEC’s]...moral responsibilities and practices.” The authors argue that we need an “ethical framework” both to “define” an SEC’s responsibilities and “provide the fundamental sharable principles” necessary to guide an SEC’s conduct “within the multicultural and international context in which they operate.” But Taddeo and Floridi also believe that this problem applies to other kinds of major “online service providers,” which the authors refer to as “OSPs”; they include Facebook and Twitter as well as Google under that category. However, we limit our focus on moral-responsibility issues in this section to SECs per se.

Does the central role that major SECs like Google and Bing now play as “gatekeepers” of knowledge in our information society entail some special responsibilities for them? Hinman (2005, 21) lists four reasons why these companies should shoulder significant social responsibility. First, he notes that search engines “play an absolutely crucial role” in accessing information and that without them, the Web would “simply be inaccessible to us” and thus “almost useless.” Second, he points out that **access to information is “crucial for responsible citizenship,”** also noting that “citizens in a democracy cannot make informed decisions without access to accurate information.” Third, Hinman notes that search engines have become “central to education,” and he points out that students now search on Google and other major search engines more frequently than they visit libraries. Fourth, he points out that **major search engines are owned by private corporations** – i.e., by businesses that are mainly, and “quite properly,” interested in making a profit. Regarding Hinman’s fourth point, it would seem that conflicts can easily arise because of an SEC’s mission to be profitable and its broader societal role in providing access to information in a manner that is fair, accurate, and unbiased. *& this democracy*

4.1 Commerce-Related Conflicts for SECs

Consider one kind of conflict involving bias and profit for SECs. Nicas (2011, 1) points out that while many SECs were initially “content to simply produce search results,” some are now very much involved in diverse markets, “offering everything from online music to local coupons to mobile phones.” Nicas also describes a bias-related concern affecting this trend by noting that when Google entered into the online travel business, it “began placing its new flight-search service atop general search results” – i.e., above those of other major players in the online travel business such as Orbitz and Expedia. Because Google also engages in these kinds of practices in EU countries, where an estimated 85% of users select Google either as their main or sole search engine, it has and continues to face charges of **anti-trust violations.** In Europe, Google has been formally charged with “systematically favoring its own comparison shopping product” by prominently displaying that product/service in its search returns, “irrespective of its merits” and at the expense of its competitors (Chappell 2016). If the accounts provided by Nicas, Chappell, and others are correct, it would seem that there are good reasons to be concerned about commercial conflicts involving major SECs such as Google.

A relatively new bias-related conflict for SECs has recently surfaced in connection with “voice search,” which is becoming increasingly popular with users of mobile devices. Barysevich (2016) notes that voice search is the “fastest growing type of search,” pointing to statistics in the U.S. showing that 55% of teenagers and approximately 40% of adults use this mode of search on a daily basis. Why are users’ searches trending in the direction of voice versus (“traditional”) text? Hawkins (2017, Other Internet Resources) offers two

reasons: (a) voice search is faster than text, since most people also tend to speak much more quickly than they type; and (b) voice search is hands-free, which is convenient for people using mobile devices. Given Google's preeminence – and some might even say its “dominance” – in the online “search industry” to date, one might assume that Google has also been the lead player in the emerging area of voice search. But Hawkins points out that Bing has had an edge over Google in the voice-search market thus far. He attributes this to the fact that Apple's Siri, Microsoft's Cortana, and Amazon's Alexa – three of the four major “virtual assistants” on the market – use Bing for their searches. Of course, one might well expect that Microsoft's voice-assistant applications would direct its users' searches to Bing, and that Google's voice assistants would do likewise in directing users to its search engine. But one could also question whether voice-search users, especially of products/services other than Microsoft's and Google's, should be given an explicit option to set default voice-search requests on their devices to a specific search engine of their choice? Otherwise, companies like Amazon and Apple could be perceived as biased because of their tilting, even if unintentionally, the voice-search market toward Bing (and away from Google)? However, since Google has already cornered so much of the existing search and search-related markets, it might be difficult for us to view that SEC as somehow itself a victim of (search) bias.

Some critics have pointed out that conflicts of interest involving SECs are not limited to bias and unfair business practices in the commercial sector, but can also affect the free flow of knowledge in society. For example, Carr (2011, 163) worries that the commercial control that Google and other major SECs now have over the “distribution of digital information” could ultimately lead to “restrictions on the flow of knowledge.” And Hinman (2008, 67) believes that the control of knowledge that SECs have is “in a very fundamental sense, a public trust, yet it remains firmly ensconced in private hands and behind a veil of corporate secrecy.” Much of this secrecy, as we have seen, is closely tied to the proprietary search algorithms that major search engines use, which also raises the question of whether aspects of these algorithms should be more transparent to the general public. But Elgesem (p. 241) believes that SECs should not be required to disclose information about their proprietary search algorithms, arguing instead that they should be required to (a) make their policies known to users and (b) follow those policies as closely as possible.

We conclude this section by noting that SECs continue to face some critical challenges with respect to fulfilling their “gatekeeper role” in a socially responsible way, while simultaneously protecting both their proprietary algorithms and the interests of their shareholders to whom they also have legal and moral obligations. In the remainder of Section 4, we focus our attention on two different kinds of accountability-related ethical concerns affecting SECs: (i) *moral responsibility/legal liability* issues arising in cases where search engines have provided links to Web sites whose content can be considered morally controversial; and (ii) questions involving trust. We begin with an analysis of (i).

4.2 Legal Liability, a “Right to Erasure,” and Fake News in the Digital Era

Should SECs be held accountable for directing users, even if unintentionally, to some Web sites whose content is either illegal or morally controversial, or both? We should note that American SECs, including Google, claim that they provide only a service, and thus are not “content providers.” These companies believe that it is not reasonable to hold them responsible/liable for the content on sites that they merely identify or provide links to in their lists of search returns. They also correctly point out that in the U.S., providers of online services – including major ISPs like Verizon and Comcast – are not held legally liable for the online content accessible through their services, provided they comply with official legal requests to remove content that explicitly violate U.S. laws. Analogously, SECs believe that if they comply with official legal requests to “de-index,” or remove links to, sites that willingly and intentionally violate U.S. laws – e.g.,

laws violating copyright, child pornography, and so forth – they too should not be held legally liable for any content that users happen to access via their services. However, many American-owned SECs operate internationally, where laws and regulatory schemes differ from those in the U.S. (as we saw in the case of Google in China).

It is worth noting that in Europe, SECs are viewed as “controllers of personal data,” as opposed to mere providers of a service (see, for example, Google Spain SL, Google Inc. 2013); as such, these companies, as well as ISPs that operate in EU countries, can be held responsible/liable for the content that is accessible through their services. One major difference regarding how SECs are viewed in the U.S. versus Europe came to the fore in the relatively recent debate about whether users should have the “Right to Be Forgotten,” now commonly referred to as the Right to Erasure.

4.2.1 The Right to Erasure (RtE)

What is RtE, and why has it been controversial from the point of view of moral accountability issues for SECs? We can trace the origin of what eventually became the RtE debate to 2010, when Mario Costeja González requested that Google remove a link included in its list of returns for online searches of his name (see, for example, Google Spain SL, Google Inc. 2014). The link in question was to an article in a Spanish newspaper (*La Vanguardia*) about a home foreclosure that occurred 16 years earlier. González, a Spanish citizen, petitioned Spain’s National Data Protection Agency, to have the link removed. He argued that the information about his foreclosure, which was still prominently displayed in Google’s list of search returns for “Mario González,” was no longer “relevant.” Although the Spanish court ruled in González’s favor in 2010, many critics were unsure that this court’s ruling would also hold in other EU countries. So, these critics were not surprised when Google appealed the Spanish court’s ruling.

As the debate in Europe escalated, many RtE advocates argued that the (former) EU Directive on Data Protection (Directive 95/46/EC), which protects the privacy rights citizens of all EU countries, includes language that indirectly supported RtE. For example, they noted that Article 12 (“Right of Access”) of that directive grants “the rectification, erasure, or blocking of data the processing of which does not comply with the provisions of the Directive because the information is...inaccurate” (see the European Commission’s “Fact Sheet on ‘The Right to Be Forgotten’ Ruling”). So, citizens of EU countries already had an explicit legal right to have *inaccurate* personal information about them removed/erased from the Internet, or at least de-linked or de-indexed from search engines. Some RtE supporters also believed that Article 12 could be interpreted to imply a right to the erasure/de-linking of online personal information that is no longer deemed to be *relevant*.

SECs and other RtE opponents, including publishers and journalists, have viewed RtE from a very different perspective. They argued that RtE would censor information, thereby limiting freedom of expression, and would degrade the overall quality of the Internet by limiting the amount of online information available to users (see, for example, Tavani 2018). Because of these and related concerns, they concluded that the public would be *harmed* by RtE. However, some RtE supporters, including Bottis (2014, 2), suggested that RtE-related data protection was needed precisely because it would help to eliminate harm – viz., “psychological harm” (and possibly even physical harm in some cases) – that users might otherwise experience. Consider, for example, that victims of “revenge porn” sites would have explicit legal recourse regarding their requests to remove/de-index links that associated their names with those controversial sites (see the account of revenge porn vis-à-vis the Right to Be Forgotten, now RtE, in Kritikos 2018).

terms to implicate erasure

In 2012, the European Commission (EC) proposed a revised Data-Protection Regulation (Article 17) for the (former) EU Directive, which included specific provisions for the removal (i.e., the “rectification, erasure, or blocking”) of personal data that was *irrelevant* (as well as personal data that was “inaccurate,” “inadequate,” or “excessive”). The EU Parliament approved the EC’s recommendations in 2013, after making a few minor modifications to Article 17 and renaming the “right” in question the “Right to Erasure.” Perhaps not surprisingly, the newly revised RtE regulation was challenged by Google and other corporations operating in Europe. However, in May 2014, the Court of Justice of the European Union (CJEU) upheld the Spanish court’s (2010) ruling in the González case, also confirming Article 17 of the updated EU Directive on Data Protection. The high court’s ruling confirmed that citizens residing in all EU nations have the right, under certain conditions, to request that search engine operators remove links to some kinds of personal information about them that are “no longer relevant.” The CJEU also noted, however, that *RtE is not an absolute right*; for example, this right would have to be balanced against other rights, including “freedom of expression” (see the EU’s “Fact Sheet on the ‘Right to Be Forgotten’ Ruling”).

Although Google announced that it would comply with the CJEU’s ruling, it also worried that it would not be able to respond to all of the users’ requests that it would likely receive to remove/de-index links; doing so, it suggested, would not only be an onerous task but might also be a *practical impossibility*. Kelion (2019) notes that in the time period between the CJEU’s May 2014 ruling and September 2019, Google had received “more than 845,000 requests to remove a total of 3.3 million web addresses,” which resulted in approximately “45% of the links ultimately getting delisted.” So it would seem that Google and its supporters indeed have a reasonable concern, given the magnitude of removal requests it has received. But one can still reasonably ask, as Bottis has, whether the sheer volume of these requests constitutes sufficient grounds for repealing RtE.

response to the “impossibility”
argument against RtE

Bottis points out *we do not cease to enact and to comply with laws that apply in physical space simply because their enforcement could not possibly eliminate crimes* such as prostitution, drug dealing, etc. She further notes that *in the digital world, protecting privacy and copyright has sometimes seemed impossible to achieve, but we still propose, enact, and enforce laws to protect both as best we can*. So on this view, we could not justify “de-legislating” RtE solely on the grounds that it is difficult to enforce. But it is also worth noting that as in the case of Google’s arguments for repealing RtE, many of the arguments advanced by RtE’s supporters also include one or more logical fallacies; an analysis of some of those fallacious arguments is included in Tavani (2018).

In one sense, the debate about RtE (formerly referred to as the RTBF debate) would now seem moot, in Europe at least, following the CJEU’s 2014 ruling and its later approval of the General Data Protection Regulation (GDPR), which was adopted by the EU in April 2016 and went into effect in May 2018. Yet, a new and different kind of debate concerning RtE has since replaced the original – viz., a dispute about (i) how RtE should be implemented in the EU, especially in light of what some have criticized as the “vague” guidelines given to SECs; and (ii) whether RtE needs to be universally applied to be effective. With respect to (i), Google and other SECs claimed that the CJEU did not provide sufficiently clear and precise criteria for guiding them in determining which requests for removal/deletion merit serious consideration and which do not. So, Google and other SECs continue to examine RtE requests on a case-by-case basis.

Regarding (ii), many privacy advocates argue that while RtE applies in all EU countries, it cannot be an effective law without strong international support; so, many members of the EC, as well as the privacy regulators in the 28 EU nations (including the UK which will soon exit the EU but will likely continue to support the principles comprising RtE), have argued that RtE should apply beyond Europe. Even though

SECs operating in Europe are required to comply with GDPR (and its RtE provisions) in all EU countries, RtE can be very easily circumvented both within and outside Europe. For example, a European user who enters the keyword “Mario González” on Google.com (or on non-local (European) versions of Bing, Yahoo, DuckDuckGo, etc.) will still find the link to the story about that person’s home foreclosure prominently listed, even though users of Google.es (in Spain), or Google users in other EU countries such as Germany (Google.de), will not see that link in their list of search returns. So RtE supporters have petitioned for broader application of the new law, while Google and other SECs have challenged that view in court. In September 2019, the EU’s highest court ruled that Google did not have to apply RtE globally (Kelion 2019). So, many privacy advocates in Europe and elsewhere believe that a universal or global version of RtE is still needed for it to be an effective law (see, for example, Global Partners Digital 2018).

4.2.2 The Challenge of Fake News in the Digital Era

Not only are SECs now required to comply with requests (from European citizens living in EU countries) to de-index and remove links to personal data that is deemed to be “irrelevant,” they currently face a somewhat similar challenge with respect to what to do about indexing and providing links to false and/or misleading information included in online forums. This is especially apparent in the case of Fake News (FN). What is FN, and why is it a challenge for SECs? Although there is no single, universally agreed upon definition of FN, it is generally understood to mean “fabricated news,” “false news stories,” and so forth. According to the *Cambridge English Dictionary*, FN is defined as “false stories that appear to be news, spread on the internet or other media, usually created to influence political views” (see the link in Other Internet Resources). The prevalence of FN in online forums, and especially on social media sites and blogs, raises questions about whether SECs are complicit in furthering the dissemination of this false information.

In the U.S., concerns about FN received considerable media attention following the 2016 presidential election, when it was reported that actors in nations outside the US were deliberately posting false and misleading information on social media sites such as Facebook to influence the outcome of the election against then-candidate Hilary Clinton. In one sense, it would seem that the specific details of this controversy affect mainly, or possibly even solely, social media sites such as Facebook and Twitter as well as blogs, but not necessarily SECs. But we can also question whether SECs might share some responsibility for exacerbating concerns about FN by providing users with search results that – even if unintentionally – link to various FN stories included on social media sites and elsewhere on the Internet.

One might be inclined to argue that since FN deliberately promotes erroneous and inaccurate information as truthful content, it should be more closely regulated by governmental agencies. But as Lipinski (2018, 69) points out, “while all Fake News possesses an element of untruth, not all Fake News is defamatory.” He goes on to note that in the U.S., the purveyor of non-defamatory FN is protected by the First Amendment, so there is no legal recourse against the “speaker” of non-defamatory FN or against the media outlets that post it. Thus, Lipinski’s description of the law could be interpreted to suggest that SECs, along with social media sites, should be exonerated from any legal liability when it comes to the dissemination of FN. Yet, even if social media sites continue to permit the posting of FN on their platforms, we can still ask whether SECs might be held to a higher standard, given their privileged role as gatekeepers of knowledge on the Web to ensure that search engine users are directed to information that is “accurate”.

Should SECs be responsible for de-indexing links that happen to direct their users to some of the original sources of FN posts. An SEC, in its defense, could argue that it would be unfair to be held responsible for pointing users to information that is legal and that is already freely available on the Internet. However, a

critic might respond that SECs operating in Europe have already agreed – albeit under the threat of law – to remove links to information that has been deemed “irrelevant,” even though that information’s content is truthful. So, *a fortiori*, the critic might ask: why shouldn’t those companies also be required to remove links to information that is blatantly false and/or deliberately misleading? Perhaps the EU countries – building on the rationale used in RtE (and also included in the broader GDPR) for allowing the erasure of personal information that is “inaccurate,” as well as no longer “relevant” – will take the lead in the future in requiring SECs to respond to requests to remove links to at least some kinds of FN.

sort of as extension to the language in the RtE

It is also worth noting that some major SECs have codes of ethics or codes of conduct containing specific principles that can be interpreted as conflicting with practices that contribute to or perpetuate the dissemination of FN, or any kind of information that has been shown to be false or misleading. For example, Google’s Code of Conduct (see link in Other Internet Resources), Section I (titled “Serve Our Users”) states that because Google delivers “great products and services,” it holds itself to “a higher standard.” And in Section I.1 (titled “Integrity”), the Google Code states:

Our reputation as a company that our users can trust is our most valuable asset, and it is up to all of us to make sure that we continually earn that trust. All of our communications and other interactions with our users should increase their trust in us.

Does Google’s commitment to hold itself to a higher standard – i.e., to more than what is legally required of that corporation in a minimalist sense – imply that it has an obligation to avoid directing its users to sites that traffic in FN? And if Google is seen as enabling increased access to false and misleading information, could that lead to an erosion of the “trust” that this SEC now seems to enjoy? In its code of conduct, Google claims that it strives “continually” to “earn” the trust of its users. But in an era of FN, we can wonder whether the sense of trust that Google, or any other major SEC, seeks to increase might instead begin to erode.

In Section 4.3 we further examine the concept of trust in the context of SECs. First, however, we conclude this section by noting that as of this date, no clear and explicit policies have been adopted by major SECs regarding FN. In fact, there is no evidence to suggest that SECs even see themselves as bearing any level of responsibility for the rapid spread of FN. Many SECs would likely argue that if any entity ought to be held responsible for the widespread FN on the Internet, it should be the social media sites and the blogs on which most of that misinformation is originally posted. Unfortunately, however, an adequate analysis of the role of social media sites and their specific responsibility in the FN controversy is beyond the scope of this entry; for more information about ethical issues affecting social media sites per se, see Vallor (2016).

4.3 Some Questions Affecting Trust

In the preceding section, we saw how the rapid spread of Fake News online raises one kind of trust-related concern for SECs: the accuracy of the information on some sites to which search engine users are directed. We next briefly examine the concept of trust in more general terms, before considering whether and how it might be possible for users to enter into trust relationships with SECs. As in the case of other key ethical issues affecting SECs, it is not possible to discuss here the many aspects of trust in the detail that they would otherwise warrant. For an excellent overview of trust (in general), see McLeod (2015). We limit our analysis to some current trust-related issues involving SECs, and suggest that the way SECs respond to these issues could affect the sense of trust that users will have in the future for major SECs such as Google and Bing.

What is trust, and why is it so important in the context of SECs? Following Baier (1986), many philosophers

view trust as an “attitude,” while others see trust either in “relational terms” (Robbins 2016) or as a kind of “expectation” (Walker 2006). And as Govier (1997, 35) notes, one reason why trust is important is because “attitudes of trust and distrust affect the nature and quality of our social reality.” When discussing the topic of trust, it is important to make some key distinctions at the outset. For example, we need to differentiate trust from trustworthiness, which is often viewed either as a *property* of the trustee (rather than as an attitude of the trustor) or a *relation* between trustor and trustee. Additionally, it is important to distinguish between ethical trust and epistemic trust (both of which are described below), in the context of SECs.

Some literature-wide distinctions

Are contemporary SECs trustworthy organizations? If so, why is it that we claim to trust them? Is it because (a) we perceive them as a reliable resource for directing us to the most relevant and accurate online information, in response to our search queries (i.e., epistemic trust)? Or is it because (b) SEC’s policies regarding what they do with our personal information, once collected, are perceived as transparent and fair (ethical trust)? Or is it because of both (a) and (b)? We can also ask whether there might be some alternative/additional reasons why one could claim to trust or distrust an SEC.

Of course, one could reasonably question whether it is possible for humans to enter into a trust relationship with an SEC, as a corporate entity, or even with a non-human entity of any kind. For example, some might assume that a genuine trust relationship can only exist between humans (or perhaps more specifically, between “human agents”). Others, however, believe that trust relationships can be extended to include some non-human, or “artificial,” agents—i.e., not only corporations, which are sometimes viewed as aggregates of human agents, but also artificial electronic agents such as “bots.” Because issues affecting trust and responsibility for SECs are distributed over a vast and diffuse network, many diverse kinds of agents can be involved—e.g., from corporate CEOs and executives/boards, to customer-service representatives, to engineers and software developers, to the many (non-human) artificial (AI) agents or bots that can also “make some limited decisions” on behalf of an SEC. So what, exactly, is it that we trust/distrust when we say we trust/distrust an SEC? In claiming to trust Google, for example, we may, in effect, be saying that we trust all of the component elements comprising that SEC.

the non-human and
diffusive problems for
trust (interrelated)

quasi-agency for bots

I don't think this is
true (e.g. w/ algorithmic recourse)

One model that can help us to think about trust in contexts as vast and complex as major SECs is Margaret Urban Walker’s framework of “diffuse default trust” (Walker 2006). In Walker’s scheme, trust relationships occur in environments (i.e., “spaces and circumstances”) that she calls “zones of default trust.” Walker differentiates trust from mere reliance by noting that in a trust relationship between A and B, A not only expects B to do X, but A expects it of B (Walker, 79). And because trust “links reliance to responsibility,” Walker (p. 80) believes that A has a “normative expectation” of B. But she also notes that we are not always conscious of these normative expectations, because they are typically “unreflective and often nonspecific” expectations where “strangers or unknown others may be relied upon to behave in acceptable ways.” Our reliance on the “good and tolerable behavior of others” is at the core of Walker’s notion of “default trust” (p. 85).

hence “diffuse”

how the “default
trust” view allows for
more flexibility in trust
relations

An important feature of Walker’s framework is that it allows for trust relations between individuals who have not met, and who will possibly never meet, in person. And in her scheme, trust relationships can occur not only between humans who are unknown to one another, but also between humans and non-human entities such as corporations. Walker illustrates the latter point in an example where someone encounters particularly bad service on an airplane operated by a major commercial airline. She points out that when we have such an experience of bad service, it is appropriate to feel resentment, not necessarily toward specific individuals who work for the airline, but toward the airline itself. So the entities comprising a zone of trust need not be limited to known (human) individuals.

Walker (p. 85) extends her notion of default trust to include zones of “diffuse default trust,” as also illustrated in her example involving the major airline corporation. Consider the many ways that responsibility can become “diffuse” when distributed over a vast zone of trust such as a commercial airline (Walker), or a major SEC like Google which also qualifies as a zone of diffuse-default trust comprising a diverse range of agents (both human and non-human). Among the non-human entities in that zone are artificial electronic agents (bots) who “operate behind the scenes,” both with humans and other artificial agents. The latter kinds of entities also include multi-agent systems that can “act” in limited ways on behalf of an SEC. For a detailed description of how Walker’s model can be applied in digital contexts, including zones of diffuse default trust comprising both humans and multi-agent systems, see the analysis in Buechner and Tavani (2011).

there's quasi-agency again

As noted above, it is important to distinguish between epistemic trust and ethical trust, both of which are essential for trusting SECs. Whereas concerns regarding the former sense of trust overlap with issues affecting an SEC’s reliability, as well as with the accuracy of information contained on the Web sites it indexes and links to, ethical trust in the case of SECs can overlap with questions/concerns pertaining to a search organization’s user policies and whether those policies are open, fair, and trustworthy. Recognizing and complying with the requirements of both aspects of trust is crucial for an SEC like Google to realize a key objective stated in its code of conduct—viz., to “increase” and “continually earn...trust” with its users (The Google Code of Conduct, 2018).

We conclude this section by acknowledging that much more could be said about the important role that trust plays in the context of SECs. But we have at least identified some of the reasons why trust is very important for SECs and also showed how those companies can qualify as “zones of diffuse default trust” (Walker). Although we have examined only one model of trust – Walker’s framework – in analyzing some of the many dimensions of trust affecting SECs, we have not argued that Walker’s is the only, or even the best, theory for understanding trust issues affecting SECs. Nor have we argued that Walker’s theory is without internal problems. Instead, we put aside those kinds of questions in favor of using a specific model to analyze trust-related questions affecting SECs in a systematic way.

We began Section 4 by asking whether SECs have any special moral obligations because of their “privileged place” in society (i.e., in light of their role as “gatekeepers” of knowledge on the Web), and in Section 4.1 we examined Hinman’s concern about the kind of control of knowledge that SECs currently have. In Section 4.2, we identified a relatively recent trust-related challenge that SECs now face because of the proliferation of Fake News on the Web. And based on our further analysis of trust in Section 4.3, it would seem that an SEC’s ability to establish and maintain trust with its users will be crucial for that organization to thrive in the future.

trust as essential to the success of SECs

5. Conclusion

In this entry, we have seen how various ethical issues arose during key developmental stages of search technology that eventually led to contemporary “Web 2.0-era search engines.” Search technology itself has evolved over the past 75 years—i.e., from pre-computer-based techniques (such as memex) intended to help scientists and professional researchers locate and retrieve important information in a timely manner, to an Internet-based technology that assisted ordinary users in locating (and linking directly to) relevant Web sites in response to their manifold search queries, to a highly sophisticated technology that, in its current form, has become increasingly commodified and personalized to the point that it might now be regarded as a threat to some of our basic freedoms and democratic ideals.

The ethical issues examined in this entry are mainly from a deontological perspective, thus reflecting the published work on this topic to date. However, in no way is our analysis of ethical issues affecting search engines intended to be exhaustive; rather it merely reflects the standard or “mainstream” approach that applied ethicists and other scholars have taken thus far in their analyses of search engine controversies. One could easily imagine questions arising from other ethical perspectives as well. From the vantage point of social justice, for example, one could reasonably ask whether search engine companies in the U.S. and other developed nations are morally obligated to help bridge the “information divide” by providing easier and more ubiquitous access (via search technologies) to users in developing nations, especially to non-English speaking users. Also, from a utilitarian or consequentialist perspective, one might ask whether the overall social consequences would be more beneficial if search engine users were legally permitted to retrieve some forms of proprietary information (e.g. proprietary online information concerning health and public-health studies that were made possible by tax-payer funding) for personal use.

Additionally, some alternative ethical frameworks could also be used in analyzing ethical issues affecting search engines. These include, but are not limited to, Hans Jonas’ “Imperative of Responsibility” (Jonas 1984), Bernard Gert’s “Common Morality” (Gert 2007), and John Rawls’ “justice as fairness” (Rawls 1999). For a very brief description of how ethical aspects of search engines might also be approached from the perspectives of “rights,” “fairness,” “common good” and “virtue,” see the Markkula Center for Applied Ethics (in the Other Internet Resources).

Even though the primary focus has been on identifying and analyzing *ethical* issues affecting search engines, in closing it is worth reiterating a point made in the introductory section—viz., that search engine technology also has implications for some other kinds of philosophical issues. This is especially apparent in the area of epistemology, where some critics raise concerns related to our received notions of the nature and justification of knowledge claims, in an era of widespread use and dependence on search engines. For example, Hinman (2008, 75) argues that search engines “contribute significantly to the social construction of knowledge”—i.e., not only do they provide access to knowledge, but they also increasingly “play a crucial role in the constitution of knowledge itself.” An analysis of this claim, however, as well as other epistemological and (broader) philosophical aspects of search engines, is beyond the scope of the present entry.

Bibliography

- Abbate, J., 1999. *Inventing the Internet*, Cambridge, MA: MIT Press.
- Baier, A. C., 1986. “Trust and Antitrust,” *Ethics*, 96: 231–260.
- Barysevich, A., 2016. “How Voice Search Will Forever Change SEO,” *Search Engine Journal*, June 14. [available online](#).
- Berners-Lee, T., 2010. “Long Live the Web: A Call for Continued Open Standards and Neutrality,” *Scientific American*, November. [available online](#).
- Blanke, T., 2005. “Ethical Subjectification and Search Engines: Ethics Reconsidered,” *International Review of Information Ethics*, 3: 33–38.
- Bottis, M., 2014. “Allow Me to Live the Good Life, Let Me Forget: Legal and Psychological Foundations of the Right to Be Forgotten and the New Developments in the European Union Laws,” in *Well-Being, Flourishing, and ICTs: Proceedings of the Eleventh International Conference on Computer Ethics—Philosophical Enquiry*, Menomonie, WI: INSEIT, Article 10.
- Brey, P., 1998. “The Politics of Computer Systems and the Ethics of Design,” in *Computer Ethics:*

- Philosophical Enquiry*, M. J. van den Hoven (ed.), Rotterdam: Erasmus University Press, pp. 64–75.
- , 2004. “Disclosive Computer Ethics,” in *Readings in CyberEthics*, 2nd edition, R. A. Spinello and H. T. Tavani (eds.), Sudbury, MA: Jones and Bartlett, pp. 55–66.
- Brin, S. and Page, L., 1998. “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” in *Seventh International World-Wide Web Conference (WWW 7)*, Amsterdam: Elsevier.
- Buechner, J. and H. T. Tavani, 2011. “Trust and Multi-Agent Systems: Applying the ‘Diffuse, Default Model’ of Trust to Experiments Involving Artificial Agents,” *Ethics and Information Technology*, 13(1): 39–51.
- Burgess, M., 2018. “What is the Internet of Things?” *Wired*, February 16. [available online](#)
- Bush, V., 1945. “As We May Think,” *Atlantic Monthly*, July. [available online](#).
- Carr, N., 2011. *The Shallows: What the Internet is Doing to Our Brain*, New York: Norton.
- Chappell, B., 2016. “EU Charges Google with Antitrust Violations, Will Also Look at Android,” *National Public Radio*, April 15, [available online](#).
- Charette, R. N., 2012. “Gone Missing: The Public Policy Debate on Unleashing the Dogs of Cyberwar,” *IEEE Spectrum*, June 4. [available online](#).
- Chorost, M., 2011. *World Wide Mind: The Coming Integration of Humanity, Machines, and the Internet*, New York: Free Press.
- Diaz, A., 2008. “Through the Google Goggles: Sociopolitical Bias in Search Engine Design,” in *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, pp. 11–34.
- Elgesem, D., 2008. “Search Engines and the Public Use of Reason,” *Ethics and Information Technology*, 10(4): 233–242.
- European Commission, 2014, “Fact Sheet on ‘The Right to Be Forgotten’ Ruling C-131/12.” [available online](#).
- Floridi, L., 2014. “The Right to Be Forgotten—The Road Ahead,” *The Guardian*, October 8. [available online](#).
- Friedman, B., P. Kahn, and A. Borning, 2008. “Value Sensitive Design and Information Systems,” in *The Handbook of Information and Computer Ethics*, K. E. Himma and H. T. Tavani (eds.), Hoboken, NJ: John Wiley and Sons, pp. 69–101.
- Friedman, B. and H. Nissenbaum, 1996 “Bias in Computer Systems,” *ACM Transactions on Computer Systems*, 14(3): 330–347.
- Gert, B., 2007. *Common Morality: Deciding What to Do*, New York: Oxford University Press.
- Global Partners Digital, 2018. *Travel Guide to the Digital World: Data Protection for Human Rights Defenders*, London: GPD. [available online](#).
- Goldman, D., 2013. “Shodan: The Scariest Search Engine on the Internet,” *CNN Money* (The Cybercrime Economy), April 8. [available online](#).
- Goldman, E., 2008. “Search Engine Bias and the Demise of Search Engine Utopianism,” in *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, pp. 121–134.
- Goodwin, D., 2018. “What is SEO? Here’s Search Engine Optimization Defined by 60 Experts,” *Search Engine Journal*, January 2, [available online](#).
- Google Spain SL, 2013. Google Inc. v. Agencia Española de Protección de Datos, Mario Costeja González. Case C 131/12. Opinion of Advocate General, June 25. [available online](#).
- , 2014. Google Inc. v. Agencia Española de Protección de Datos, Mario Costeja González. Case C 131/12. Judgment of the Court (Grand Chamber), May 13. [available online](#).
- Govier, T., 1997. *Social Trust and Human Communities*, Montreal and Kingston: McGill-Queen’s University Press.
- Halavais, A., 2009. *Search Engine Society*, Malden, MA: Polity.

- Halpern, S., 2011. "Mind Control and the Internet," *New York Review of Books*, June 23. [available online](#)
- Himma, K. E., 2007. "Privacy vs. Security: Why Privacy is Not an Absolute Value or Right," *University of San Diego Law Review* (Fourth Annual Editors' Symposium), 45: 857–921.
- Hinman, L. M., 2005. "Esse Est Indicato in Google: Ethical and Political Issues in Search Engines," *International Review of Information Ethics*, 3: 19–25.
- , 2008. "Searching Ethics: The Role of Search Engines in the Construction and Distribution of Knowledge," in *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, pp. 67–76.
- Introna, L. and H. Nissenbaum, 2000. "Shaping the Web: Why The Politics of Search Engines Matters," *The Information Society*, 16(3): 169–185.
- Jonas, H., 1984. *The Imperative of Moral Responsibility: In Search of an Ethics for the Technological Age*, Chicago: University of Chicago Press.
- Kelion, L., 2019. "Google Wins Landmark Right to Be Forgotten Case," *BBC News*, September 24. [available online](#).
- Kelly, K.J., 2019. "Gawker Stalker Might Get Another Chance to Acquire Website," *New York Post*, August 6. [available online](#).
- Kritikos, K. C., 2018. "The Right to Forget, Obliterate, Erase: Defending Personal Data Privacy in the Digital Age," *Journal of Information Ethics*, 27(2): 47–65.
- Lessig, L., 2000. *Code and Other Laws of Cyberspace*, New York: Basic Books.
- Levy, D. M., 2008. "Information Overload," in *The Handbook of Information and Computer Ethics*, K. E. Himma and H. T. Tavani (eds.), Hoboken, NJ: John Wiley & Sons, pp. 497–515.
- Lipinski, T., 2018. "The Law and Economics of Recognizing the Right to Be Forgotten in an Era of Fake News," *Journal of Information Ethics*, 27(2): 66–80.
- Markkula Center for Applied Ethics, 2016. "Unavoidable Ethical Questions about Search Engines," Santa Clara University. [available online](#).
- McLeod, C., 2015. "Trust," *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2015/entries/trust/>>.
- Moor, J. H., 1997. "Towards a Theory of Privacy in the Information Age," *Computers and Society*, 27(3): 27–32.
- Morozov, E., 2011. "Your Own Facts," *New York Times Sunday Book Review*, June 10. [available online](#)
- Nagenborg, M. (ed.), 2005. *The Ethics of Search Engines*, Special Issue of *International Review of Information Ethics*, Vol. 3.
- Nicas, J., 2011. "Google Roils Travel," *Wall Street Journal*, 12/27. [available online](#).
- Nissenbaum, H., 1997. "Toward an Approach to Privacy in Public: Challenges of Information Technology," *Ethics and Behavior*, 7(3): 207–219.
- , 1998. "Protecting Privacy in an Information Age," *Law and Philosophy*, 17: 559–596.
- , 2004. "Privacy as Contextual Integrity," *Washington Law Review*, 79(1): 119–157.
- , 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford: Stanford University Press.
- O'Harrow, R., 2012. "Cyber Search Engine Shodan Exposes Industrial Control Systems to New Risks," *The Washington Post*, June 3, [available online](#).
- Noble, S.U., 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: New York University Press.
- O'Reilly, T., 2005. "What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software," *O'Reilly Media*, [available online](#).
- Palfreman, J., and D. Swade, 1991. *The Dream Machine: Exploring the Computer Age*, London: BBC Books.


- Pariser, E., 2011. *The Filter Bubble: What the Internet is Hiding from You*, New York: Penguin.
- Rawls, J., 1999. *A Theory of Justice*, revised edition, New York: Oxford University Press.
- Robbins, B. G., 2016. "What is Trust? A Multidisciplinary Review, Critique, and Synthesis." *Sociology Compass*, 10(10): 972–986. doi:10.1111/soc4.12391
- Scott, M., 2014. "The Right to Be Forgotten Should Apply Worldwide, Panel Says," *New York Times*, November 26. [available online](#).
- , 2016. "Europe Tried to Rein In Google. It Backfired," *New York Times*, April 18, [available online](#).
- Spinello, R. A., 2011. *CyberEthics: Morality and Law in Cyberspace*, 4th edition, Sudbury, MA: Jones and Bartlett.
- , 2012. "Google in China: Corporate Responsibility on a Censored Internet," in *Investigating Cyber Law and Cyber Ethics: Issues, Impacts, Practices*, A. Dudley, J. Braman, and G. Vincenti (eds.), Hershey, PA: IGI Global, pp. 239–253.
- Sunstein, C., 2001. *Republic.com*, Princeton, NJ: Princeton University Press.
- Taddeo, M. and L. Floridi, 2016. "The Debate on the Moral Responsibilities of Online Service Providers," *Science and Engineering Ethics*, 22: 1575–1603. doi:10.1007/s11948-015-9734-1
- Tavani, H. T., 1998. "Internet Search Engines and Personal Privacy," in *Computer Ethics: Philosophical Enquiry*, M. J. van den Hoven (ed.), Rotterdam: Erasmus University Press, pp. 214–223.
- , 2005. "Search Engines, Personal Information, and the Problem of Protecting Privacy in Public," *International Review of Information Ethics*, 3: 39–45.
- , 2007. "Philosophical Theories of Privacy: Implications for an Adequate Online Privacy Policy," *Metaphilosophy*, 38(1): 1–22.
- , 2016. *Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing*, 5th edition, Hoboken, NJ: John Wiley and Sons.
- , 2018. "Should We Have a Right to Be Forgotten? A Critique of Key Arguments Underlying This Question," *Journal of Information Ethics*, 27(2): 26–46.
- Tavani, H. T. and F. S. Grodzinsky, 2002. "Cyberstalking, Personal Privacy, and Moral Responsibility," *Ethics and Information Technology*, 4(2): 123–132.
- Tavani, H. T. and J. H. Moor, 2001. "Privacy Protection, Control of Information, and Privacy-Enhancing Technologies," *Computers and Society*, 31(1): 6–11.
- Vallor, S., 2016. "Social Networking and Ethics," *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2016/entries/ethics-social-networking/>>.
- Van Couvering, E., 2008. "The History of Internet Search Engines: Navigational Media," in *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, pp. 177–206.
- Walker, M. U., 2006. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*, New York: Cambridge University Press.
- Wall, A., 2011. "History of Search Engines: From 1945 to Google Today," *Atlantic Online*, [available online](#).
- Zimmer, M., 2008. "The Gaze of the Perfect Search Engine: Google as an Institution of Dataveillance," in *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer (eds.), Berlin: Springer-Verlag, pp. 77–99.

Academic Tools

🔗 [How to cite this entry](#).

🔗 [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).

 [Look up topics and thinkers related to this entry](#) at the Internet Philosophy Ontology Project (InPhO).

 [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

Other Internet Resources

- Hawkins, J., 2017. “5 Things to Know About Voice Search and Bing,” *SEMrush Blog*, August 4. [available online](#).
- [Fake News](#), entry in the *Cambridge English Dictionary*.
- [Google Code of Conduct](#), Alphabet Investor Relations, last updated July 31, 2018.
- [The Ethics and Politics of Search Engines](#), panel discussion, co-sponsored by Santa Clara University Markkula Center for Applied Ethics and the Santa Clara University Center for Science, Technology, and Society, on February 27, 2006.
- [Unavoidable Ethical Questions about Search Engines](#), Markkula Center for Applied Ethics
- [Code of Ethics and Professional Conduct](#), Association for Computing Machinery.
- [SEO Code of Ethics](#), Bruce Clay, Inc.

Related Entries

[computing: and moral responsibility](#) | [ethics: internet research](#) | [information technology: and moral values](#) | [information technology: and privacy](#) | [information technology: phenomenological approaches to ethics and privacy](#) | [social networking and ethics](#) | [technology, philosophy of](#)

Acknowledgments

I am grateful to the following colleagues for their helpful suggestions on earlier drafts of this entry: Maria Bottis, Jeff Buechner, Lloyd Carr, Jerry Dolan, Fran Grodzinsky, Ken Himma, Larry Hinman, Lundy Lewis, Martin Menke, Richard Spinello, and Michael Zimmer. I would also like to thank the anonymous SEP reviewers for their insightful comments and helpful suggestions.

[Copyright © 2020](#) by
[Herman Tavani](#) <htavani@rivier.edu>

[Open access to the SEP is made possible by a world-wide funding initiative.](#)
[Please Read How You Can Help Keep the Encyclopedia Free](#)

Stanford | Center for the Study of
Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2020](#) by [The Metaphysics Research Lab](#), Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054