

Responsible epistemic technologies: A social-epistemological analysis of autocompleted web search

Boaz Miller

Bar-Ilan University, Israel

Isaac Record

Michigan State University, USA

Abstract

Information providing and gathering increasingly involve technologies like search engines, which actively shape their epistemic surroundings. Yet, a satisfying account of the epistemic responsibilities associated with them does not exist. We analyze automatically generated search suggestions from the perspective of social epistemology to illustrate how epistemic responsibilities associated with a technology can be derived and assigned. Drawing on our previously developed theoretical framework that connects responsible epistemic behavior to practicability, we address two questions: first, given the different technological possibilities available to searchers, the search technology, and search providers, who should bear which responsibilities? Second, given the technology's epistemically relevant features and potential harms, how should search terms be autocompleted? Our analysis reveals that epistemic responsibility lies mostly with search providers, which should eliminate three categories of autosuggestions: those that result from organized attacks, those that perpetuate damaging stereotypes, and those that associate negative characteristics with specific individuals.

Keywords

Autocomplete, autosuggestions, epistemic responsibility, Internet, knowledge, search engines, social epistemology

In 2012, Bettina Wulff, former German President's ex-wife, sued Google for defamation. When searching for "Bettina Wulff," Google automatically suggested words such as "prostitute" and "escort" (Niggemeier, 2012).¹ Some of Google's autosuggestions also reflect prejudices and stereotypes, for example, "Why do gay men ... get AIDS" or "Why do Jews ... have big noses" (Baker and Potts, 2013). A 2013

ad campaign for *UN Women*, which aims at demonstrating the pervasiveness of discrimination against women, features women's mouths covered with real Google autosuggestions, such as "Women shouldn't ... have rights" and "Women need to ... be disciplined" (Figure 1) (Mahdawi, 2013).

In legal cases such as Wulff's, Google has claimed its autosuggestions merely objectively reflect its users' interests (Jordans, 2013; Niggemeier, 2012). Nevertheless, Google officially refrains from auto-completing queries "related to pornography, violence, hate speech, and copyright infringement" (Google, 2014). In 2013, Google eliminated the autosuggestion "autistic people should ... die" (O'Brien, 2013); in 2011, Google was criticized for caving in to pressure from the entertainment industry when it stopped autosuggesting torrent-download sites (Ernesto, 2011).

Search auto-complete raises normative questions—ethical, political, and legal.² We address its *epistemic* dimensions. We ask how search queries can be responsibly auto-completed from an epistemic perspective. Our analysis draws on our previously developed account of the relations between technological possibilities and responsible epistemic conduct (Miller and Record, 2013; Record, 2013).

Theoretical framework

Our analysis is from the theoretical perspective of social epistemology, which is the branch of philosophy that studies the social dimensions of knowledge. While *sociologists* of mass media knowledge study how knowledge claims in the media gain social and political legitimacy (Ettema and Glasser, 1987; Ekström, 2002; Parasie and Dagiral, 2013; Rogers and Ben-David, 2008), *philosophers*—specifically, social epistemologists—are interested in their epistemic validity, analyzed in terms of rationality, evidence, truth, and so on. Social epistemologists widely acknowledge that epistemic validity and socio-political legitimacy are closely intertwined (Fricker, 2007; Longino, 2002; Solomon, 2001).

Social epistemologists of digital and social media engage in normative epistemic evaluation of Internet platforms and practices, such as blogging (Coady, 2011; Goldman, 2008; Munn, 2012; Turner, 2013), googling (Simpson, 2012), ranking (Origi, 2012; Simpson, 2011), user-anonymity (Frost-Arnold, 2014), crowdsourced GPS (Global Positioning System) navigation (Record and Miller, 2016), and Wikipedia (De Laat, 2011; Fallis, 2011; Magnus, 2009; Wray, 2009). A related project analyzes Internet phenomena to re-examine traditional epistemic notions such as knowledge, justification, trust, and testimony, with an eye to epistemically bettering current online practices (Freiman, 2014; Lynch, 2016; Miller and Record, 2013; Record and Miller, 2016; Simon, 2010; Tollefsen, 2009; Weinberger, 2011).

This article's treatment of auto-complete illustrates a principled theoretical basis for allocating epistemic responsibilities. We increasingly embed knowledge practices in technological infrastructures that, together with our competencies to manipulate them, constrain what is practicable for us to know and do. Our theoretical framework



Figure 1. An ad from the United Nations (UN) Women campaign, which features real autosuggestions.

(Miller and Record, 2013; Record, 2013) acknowledges and accounts for these technologies and the social structures that form around them. It states that what we can know depends on what we can do in practice, which depends on the available material and conceptual resources. New technology makes certain previously impracticable actions practicable. Such new possibilities effectively change existing standards of epistemic responsibility, which in turn change existing standards of epistemic justification. For example, it is impossible to know that the surface of the moon is not smooth without observing it with a telescope. It was technologically possible to build a telescope in 1610, but not in 1210; thus, it was possible to have that knowledge in 1610, but not in 1210.

An epistemically responsible subject does, inasmuch as she can, what is required of her to bring about true and rational beliefs. Responsibility is delimited in part by role-expectations. Practicability is delimited by the subject's competencies and her technological, ethical, and economic circumstances. Practicability sets an upper limit on responsibility in that a responsible subject need not do more than what is practicable. It sets a lower limit on responsibility, because when an inquiry or validation becomes practicable, performing it tends to become a minimal requirement for forming responsible beliefs on relevant matters. The more practicable a putative epistemic justification-increasing activity is, the likelier that it would be included among the subject's responsibilities.

Previously, we focused on search-engine users' epistemic responsibilities, while here we extend the analysis to information providers. We argue that autocomplete inevitably and irreparably induces changes in users' epistemic actions, particularly their inquiry and belief formation. These changes have potential epistemic benefits and harms to individual users and society. We identify these changes and explain how autocomplete should prevent or minimize harms. Our methodology can be easily extrapolated to other technologies.

Technical overview

We first examine the workings of autocomplete technology. From a user's perspective, autosuggestions spontaneously appear adjacent to the typed contents of the search box, for example, as a dropdown list. As she types, the suggestions update accordingly. At any time, she can select a suggestion and is immediately transferred to a new page with relevant search hits. In some implementations, the entire hits page corresponding to the "best-guess" autosuggestion is displayed and updated as the user types.

Well-functioning autocomplete features rely on several technical components that have only recently come together. Algorithms for quick fetching of spelling corrections have been used since the 1960s. So have data structures for quickly predicting terms from an index (Knuth, 1998: Chapter 6). But *search engine* autocomplete also requires responding to users' input in real time. Real-time communication became practical only after certain browser, server, and communication technologies became widely available. Traditionally, a web user would work on a browser on the client side and have to initiate sending information to a remote server by pressing a "send" button, clicking a hyperlink, or typing a URL. The client would then wait for the server to respond. This synchronous communication architecture suited slow networks, which could not efficiently handle continuous rapid back-and-forth transmissions.³ Only around 2004, when high-speed infrastructure improved and reached the common household, was interactive real-time communication developed (Garrett, 2005).

Autocomplete hinges on the ability to prompt the user with the most relevant suggestions, which depends on the quality of the index. Current indexing methods use a vector-space model, in which sets of keywords are represented as vectors in higher dimensional space. Relevant suggestions are taken from similar vectors as measured by the "distance" between vectors (Hiemstra, 2009). Statistical analyses are also heavily used. These techniques require brute computing power and vast memory space, which only became practical around 2000.

Autosuggestions depend on many *ad hoc* implementation decisions. We can make informed guesses about the factors influencing autosuggestions. The number of queries for a term is a major factor and so are the number of click-throughs, the user's geographic location, and her personal use history. Other relevant factors include the number of hits and their quality, diversity across domain names, search language, explicit occurrence of terms in search hits, and trends among users (BBC News, 2012; Sullivan, 2011; Wiideman, 2011).⁴

Autocomplete is susceptible to manipulation. Manipulability raises concerns about the objectivity of autosuggestions. There are several recorded successful attempts to manipulate Google's autocomplete, for example, using Amazon's *Mechanical Turk* crowdsourcing system, by having many users on different machines search the same terms (Sullivan, 2011; Wiideman, 2011). In 2013, an Israeli plaintiff won a defamation case against Google for making defamatory autosuggestions when searching for his name. His lawyers manipulated the Google interface to suggest "takes chocolate milk" when searching the judge's name. They claimed they could have easily manipulated it to autosuggest "takes bribe" (Anonymous, 2013).

Epistemically relevant features of autocomplete

Autocomplete introduces novel distinctive technological possibilities, which according to our theoretical framework impose novel epistemic responsibilities. First, autosuggestions result from many individuals' collective search behavior. This may help users draw attention to aspects they might otherwise have neglected, although the mere fact that something is popular does not guarantee its correctness or usefulness. In popular searches, a user's impact on autosuggestions other users see is negligible, but in unpopular searches, such as manipulation cases, an individual or a small group may have substantial influence.

Second, autosuggestions are generated automatically, without a human being's direct involvement. This distinguishes autocomplete from human testimony and other practices that social epistemologists usually study. Automaticity and lack of supervision raise questions about the system's reliability and capacity to correct mistakes.

Third, autosuggestions are subject to a feedback loop—the more people choose a suggestion, the likelier it is to reappear. Popular autosuggestions become more popular, while unpopular autosuggestions wither. This ampliative effect has epistemic implications, insofar as searchers deem the results representative. Searchers seeking information on unpopular topics may have to enter very precise search terms without the aid of autosuggestions before they find the desired results. Meanwhile, sensational details may propagate unchecked. Autocomplete may operate more like a rumor mill than a reference librarian.

Fourth, users' exposure to autosuggestions is involuntary.⁵ Users cannot type a search without encountering autosuggestions. Once seen, they cannot “unsee” the results. A searcher seeking the airtime of her favorite television program, for example, might see an autosuggestion that “spoils” the next episode by revealing crucial plot information.

Psychology suggests two likely consequences of involuntary exposure. *First*, initially disregarded associations sometimes transform into beliefs because humans are prone to source-monitoring errors: subjects mistake the original information source and may put more or less credence in the information than they would have given the correct source (e.g. Johnson et al., 1993).⁶ Someone might read “Bettina Wulff prostitute” and initially disregard it, but later, having forgotten the source, recall having read it. This is supported by the sleeper effect, according to which when people receive a message with a discounting cue, they are less persuaded by it immediately than later in time (Kumkale and Albarracín, 2004).⁷ *Second*, involuntary exposure to certain autosuggestions may reinforce unwanted beliefs. Humans are bad at identifying and rooting out their implicit biases (Kenyon, 2014). Because exposure is involuntary, even subjects hygienic in their epistemic practices may be negatively affected.

Fifth, autosuggestions interactively affect a user's inquiry, leading to paths she might not have pursued otherwise. Effectively, if a user looks at the screen, she can't help but see the autosuggestions, and these impressions can affect her inquiry. Autosuggestions may seem to a user to delimit the possible options or represent what most people find relevant, either of which may change her search behavior. She may change her search terms for one of the suggestions, add or subtract additional terms to rule out or in suggested results. She may abandon her search altogether because the autosuggestions seem to provide the answer or indicate that there is no answer to be found that is she may assume that because nothing is being

suggested, no results for the query exist. Furthermore, because the displayed information may be incomplete or out of context, she might reach a different conclusion on the basis of autosuggestions than if she actually visited the linked page.

Altering a user's path of inquiry can have positive effects, as when he is exposed to relevant information he might not have encountered given his chosen search terms. But the effects may also be negative. For example, the first woman in space was Soviet cosmonaut Valentina Tereshkova, but when a user types "first woman in space," Google autosuggests "Ride"—a reference to Sally Ride—the first *American* woman in space. Some users follow this autosuggestion into a wrong path of inquiry (White and Marchionini, 2007: 701). Such derails in inquiry may be deleterious—as when a searcher seeks information about a vaccination's side-effects and is led to persuasive anti-vaccination websites promoting the now-discredited idea that vaccines cause autism.

Finally, autosuggestions affect users' belief formation process *in a real-time interactive and responsive manner*. "It helps to complete a thought," as one user put this (Ward et al., 2012: 12). They may thus generate beliefs the user might not have had. Based on autosuggestions, I might erroneously believe that President Obama is a Muslim, Bettina Wulff was a prostitute, vaccines cause autism, women should stay out of the workforce, or Sally Ride was the first-ever woman in space. Alternatively, I might come to believe that these things are possible, where before I held no beliefs about them, or I might give these propositions more credence than I would otherwise. Autocomplete is like talking with someone constantly cutting you off trying to finish your sentences. This can be annoying when the person is way off base or pleasant when he seems like your mind-reading soulmate. Either way, it has a distracting, attention-shifting effect that other interactive interface technologies lack.

Other technologies share some features of autocomplete. We are involuntarily exposed to advertisements; non-autocompleted search is subject to a feedback loop; and autocomplete interfaces in applications other than web search, for example, autocompleted email addresses in an email client, may also alter their users' belief formation processes in real time. But no other technology has *all* six features. From the perspective of our theoretical framework, this means that autocompleted search introduces new technological possibilities, which may entail new epistemic responsibilities.

Epistemic virtues and harms of autocomplete

Autocomplete has potential virtues. Empirical studies (Kato et al., 2013; Ward et al., 2012; White and Marchionini, 2007) find that it increases the speed of human-computer interactions and improves the quality of results, especially for slow typists or poor spellers (spelling correction is one of its most common uses); autocomplete disambiguates search terms, alerts users to possibilities beyond the ones they envisioned, or, conversely, narrows down their search and makes it more specific; it makes some users more confident that what they are looking for is out there because they are not the first to look for it. In addition, the technology has a didactic effect: it teaches users how to search by exemplifying common search constructions. Finally, autocomplete enhances users' meta-knowledge about the views and interests in society, especially politically incorrect views, which may be

underrepresented in mainstream channels. For example, without autocomplete, one may not know that others wonder whether Obama is Muslim.

The effects of autocomplete, though, vary for different users. For example, autocomplete helps dyslectic users search more effectively and reach performance levels equivalent to those of non-dyslectic users (Berget and Sandnes, 2015), while it cognitively overloads elderly users, thus derogating from the effectiveness of their search (Doubé and Beh, 2012).

Autocomplete also has drawbacks. Following Fricker's (2007: 9–29) analysis of testimonial injustice, we identify three potential harms of autocomplete. *First*, the *search target's* dignity and autonomy may be harmed due to the spread of misinformation about them. Such misinformation may be about a specific individual, as in Wullf's case, or about minority or disempowered-group members, as in the prejudicial suggestions about gays, Jews, or the Māori (Baker and Potts, 2013; Elers, 2014). *Second*, the *searcher* is harmed when she develops false, biased, or skewed beliefs. As noted, autocomplete may interfere in the midst of a searcher's stream of thought or process of inquiry and divert it toward forming biased beliefs. *Third*, *society as a whole* is harmed when prejudicial, biased, or false beliefs about members of different groups in society lead to the incorrect assessment of the trustworthiness of informants who belong to these groups and, consequently, to blocking the circulation of critical ideas in society.

Autocomplete perpetuates prejudicial beliefs in subtler ways than merely making autosuggestions that state prejudicial stereotypes. A study of Google autosuggestions found that American parents are 2.5 times more likely to Google "Is my son gifted?" than "Is my daughter gifted?" and twice as likely to Google "Is my daughter overweight?" than "Is my son overweight?" Parents are generally more concerned with their daughters' looks and their sons' intellectual abilities. Not only do these queries reflect parents' gender-differentiated treatment of their children, their rates are misaligned with an American reality in which boys are 9% more likely to be overweight than girls, while girls are 11% more likely to be in a gifted program (Stephens-Davidowitz, 2014). Not only does autocomplete perpetuate googling parents' stereotypical beliefs about boys and girls due to its feedback loop, but because of its involuntariness and real-time interference in its users' belief formation processes, it can also plant ideas such as that their daughter might be overweight in parents who did not entertain such thoughts before.

Assigning responsibility in the autocomplete case

How can and should the harms of autocomplete be rectified? According to our theoretical framework, technological possibility constitutes a salient factor in determining lower and upper bounds on epistemic responsibility. We therefore need to answer two questions (which we do in this section and in the "Epistemically responsible autocomplete" section, respectively). First, given the different technological possibilities available to the different parties (searchers, the search technology itself, and search providers), who should bear which responsibilities? Second, given the epistemically relevant features of the technology and its harms, how should search terms be responsibly autocompleted? We argue that the epistemic responsibility lies mostly with the search provider rather than searchers or the technology itself and that the prospects of devising a purely technological solution are grim due to inherent properties of the technology.

Who should bear epistemic responsibility for the harms of autocomplete? There are four options: holding the searchers collectively responsible, holding the searchers individually responsible, holding the technology responsible, and holding the search provider responsible. We will consider each candidate in turn.

Can we hold searchers collectively responsible? Assuming that corporate bodies may be held responsible *as collectives*, independently of their individual members,⁸ we may seemingly assign collective epistemic responsibility to a group of people who make the same problematic searches or click on the same problematic autosuggestions, and thus influence the suggestions that others see. However, all accounts of collective responsibility require that a group has minimal internal structure and cohesion for bearing responsibility as a group. The collective typically has neither the information nor the organizational structure to consider the effect of its aggregate behavior or produce a coordinated response to such predictions. The only thing that unites members of the group is that they have used the same search terms. The group thus cannot assume collective responsibility.⁹ But there are exceptions. Concerted efforts to manipulate search results by crowdsourcing appear to meet the cohesion requirement because participants explicitly elect to participate. Likewise, actions by a hacker activist group probably qualify. For the purposes of this article, however, we focus on typical individual searchers' aggregate behavior.

What about holding searchers individually responsible? This is tricky because not every individual performing a problematic search acts in an epistemically irresponsible way. A searcher need not be biased to make a problematic search. Suppose the view that the Ferengi are greedy reflects a prejudicial stereotype. Typing the query "Why are the Ferengi greedy?" is not always epistemically irresponsible because one may type such a query without believing it, for example, to identify who holds this view. Even if someone types a query that reflects a prejudicial belief, she still might not be epistemically irresponsible. Rather, typing the query and reading the search results may help her encounter alternative views, correct her errant belief, and become aware of its prejudicial nature. Epistemic responsibilities might demand that individuals explore additional search terms rather than avoiding particular ones. Indeed, the recommendation to seek out additional sources accords with our previous findings regarding search engines generally (Miller and Record, 2013).

One may respond that typing a problematic query is not epistemically irresponsible for its epistemic effects on the searcher, but its side-effects for other searchers. This is similar to an individual's moral responsibility to refrain from unnecessarily driving her car or consuming too much because of the negative environmental side-effects of such actions. Thicke (2013) dubs this the "Internet Observer Effect," which is that "any time you do anything on the internet; it affects the epistemic landscape of not just yourself, but everyone else who uses the internet." Thicke gives the example of searching for an individual who has offensive views about rape. Merely searching for this author and "rape" could cause Google's algorithms to associate the author more closely with "rape," causing his website to be ranked higher when others search for "rape," ultimately influencing people's beliefs about what counts as rape.

Nevertheless, assigning individual responsibility to searchers, or at least the bulk of responsibility, is wrong. Many searchers are unaware of the negative epistemic implications of their actions. Even when they are, they do not know how their

actions affect the autosuggestions others encounter. Even if a searcher knew how her search affected others' autosuggestions, she had no control over this process. She could refrain from searching at all, but this all-or-nothing alternative seems disproportionate to the epistemic risks, especially considering the potential rewards of autocomplete.

One may object that ignorance of how technology works is not always a valid defense for epistemically irresponsible action. Indeed, we advocate this view elsewhere (Miller and Record, 2013). A responsible subject should typically do what she can to ensure the information upon which she forms her beliefs is unbiased and complete enough to underwrite her judgments. Elsewhere, we argued that this entails a duty to ensure that her information sources are unbiased, especially when she does not know how they work. But the current worry is not about her own beliefs, but the potential harm to others' beliefs, whom she may not even know. Extending a subject's epistemic responsibilities to consider indirect harm to others' beliefs goes too far. Searching is distinct from cases of testimony or expert testimony, where extending epistemic responsibility to consider indirect harms may be appropriate.

A second reason why an individual searcher should not bear (much) responsibility is that the effect of her actions is typically minuscule. The infinitesimal effect of a single search suggests that the searcher's share of responsibility is similarly infinitesimal. Individual searches influence the behavior of the algorithm, but this influence is typically so dilute as to be almost irrelevant in practice.

The third candidate for having epistemic responsibility for autocomplete is the technology itself. One might immediately object that responsibility is inherently tied to human judgment and cannot be ascribed to technology. However, some argue that technologies could be bearers of responsibility. According to Latour (2002), traditional ethics unjustifiably distinguishes judgments about responsibilities, which it deems moral, from judgments about which means best serve which ends, which it deems morally neutral. But available means and the certainty of our judgments about them delimit the space of possible ends, which in turn delimits the space of responsibilities and obligations. This makes means themselves moral entities because they perform moral roles and constitute moral ends (cf. Adam, 2008).

Latour further describes how technological agents perform their social function in a "relentlessly moral" fashion. The hydraulic door closer vigorously shuts the door on those laden with packages or infirm with age. They plead or curse in vain, for the door closer cannot be convinced or bribed (Latour, 1988). Similarly, the speed bump equally slows down the reckless speeding driver and the ambulance pressed for time (Latour, 1996: 9).¹⁰ Autocomplete has similar moralizing effects. By suggesting certain search terms, autocomplete encourages searchers to trust them to yield helpful results. It didactically leads users to enter different, potentially better search terms.

Relentless, moral algorithms cannot be convinced or bribed to show different results. But does that make them responsible for their behavior? We avoid the principled ontological question of whether artifacts can be proper moral agents and pursue a targeted inquiry into the technological possibilities presently available. We ask whether autocomplete can *de facto* reliably perform the epistemic

responsibilities accompanying autocomplete. Insofar as it cannot, the responsibility *must* lie at a qualified human agent.

Regardless of whether technology can be a proper moral agent, there is a prevalent tendency, especially among technologists, of thinking that the solutions to the sort of problems we raised should be primarily technological or algorithmic. Implementing this thought would amount in practice to letting technology resolve questions of appropriate autocomplete. We argue against this option and insist that discharging technology alone with responsible autocomplete is not practicable.

So why can't an algorithm make epistemically responsible autosuggestions avoiding the harms discussed in the "Epistemic virtues and harms of autocomplete" section? Algorithms cannot autonomously identify which alternatives might harm subjects' belief formation processes. Moreover, if we take additional inspiration from Fricker's (2007) analysis of testimonial injustice, to correct for the harms of epistemic injustice, a biased subject should reflectively revise his prejudices over time, when the evidence he encounters does not support them. An algorithm seems incapable of this. The algorithm would need to identify which suggestions are offensive or prejudicial and know the actual state of affairs. For example, in the "is my son ... gifted" and "is my daughter ... overweight" case, discussed in the "Epistemic virtues and harms of autocomplete" section, the algorithm would need to know the actual rate of giftedness in boys and obesity in girls to know how frequently to make these suggestions. And even this would be insufficient for responsibly autocompleting *individual* queries, unless the search provider had data about the specific user's children to know which suggestions are most appropriate.

Judgments regarding harmfulness are context-sensitive, while current text and data analysis algorithms are not. Similar or even identical texts, keywords, and pictures may be offensive in some contexts and not others. Morozov (2013: 140–143) describes two examples where Google's algorithms made a wrong automatic judgment. In the first, they banned a respectable highbrow magazine from Google's AdSense program for publishing a story by a renowned author about a man's sexual experiences because they deemed it pornographic. In the second case, they banned a Nigerian political activist's site for posting photos protesting against police brutality in Nigeria because they deemed them as too violent. Society has different standards for treating nudity, violence, and so on in different contexts. It matters whether they appear in a museum, an encyclopedia, or a top-shelf magazine. Current technology faces insurmountable difficulties implementing subtle context-sensitive standards required for responsible autocomplete.

Kuflik (1999) lists several additional limitations of computing technology, which militate against discharging responsibilities to unsupervised computers. These limitations are principled in that they are inherent to the operating principles of code-running digital machines; thus, they cannot be overcome by mere improvements of computational power. First, testing large software is difficult because there are many more possible than testable inputs. Thus, software reliability greatly depends on the inputs its testers could predict in advance. Additionally, correcting one error can introduce new errors. While developers try to design software modularly, there can be unanticipated dependencies between modules. Moreover, small errors in the input or code can lead to large and significant errors in the output. Recent examples include a person charged US\$81 billion on his credit card for filling a gas tank (Trumbo, 2009) and a used book posted on Amazon for US\$23 million by automatic pricing algorithms (Eisen,

2011). These stories are newsworthy for their piquancy, but similar errors are common, and fixing them is part of many programmers' routine. For these reasons, responsibility must ultimately lie in human hands.

Two practicable options for making sensitive judgments are allowing users to flag harmful autosuggestions and blacklisting certain terms. But in each case, responsibility clearly lies with the users or programmers—not the algorithm. Indeed, programmers intervene just this way to alter autosuggestions. There is a blacklist of phrases that Google would not autosuggest. All major search engines refrain from autocompleting pornographic queries. This is evident in Figure 2, where only the small search engine Babylon makes sexually explicit suggestions.

Blacklisting policies constitute a tacit admission that (1) there are undesirable responses, (2) the algorithms do not weed them out, and (3) programmers can intervene effectively to correct flaws in the algorithm. Search providers hold themselves responsible for some results of unsupervised algorithms. It may be argued that these interventions are relatively narrow: they only intend to filter out “bad” words, sexually explicit suggestions, and some Internet back alleys. But the scope of intervention is irrelevant. Providers can and do alter the algorithm to remove certain autosuggestions.

The final site for autocomplete responsibility is the search company. Because we arrived here by elimination, let us ensure we have not been hasty. One may raise several reasons why creators of complex algorithms should not be held (fully) responsible. First is the “many hands” problem, which suggests that programmers are not individually responsible because software projects are typically created by large dispersed teams where the causal connection between any programmer and the searcher may be negligible (Friedman, 1990; Jonas, 1984; Nissenbaum, 1994). Furthermore, software projects combine off-the-shelf code packages with their own, which the programmers cannot test. We think this insulates individual programmers from responsibility for parts of the system, but does not insulate the corporate entity controlling the algorithm from responsibility.

The temporal and physical separation between the provider and individual searches militates against holding the provider directly responsible for a specific unanticipated harm. We think this argument holds some weight. Providers should not be responsible for every epistemic harm caused by their algorithm. But in certain well-described cases, like the Bettina Wulff case or the reinforcement of negative stereotypes about disenfranchised groups, search providers can intervene to curtail harmful suggestions once identified.

Another reason to doubt that search providers bear much responsibility is that user behavior is unpredictable. Unanticipated epistemic practices regarding autocomplete would seem to relieve the provider of responsibility. Search providers, however, already monitor user behavior, and even if they do not anticipate every new use of autocomplete, they can identify trending topics, review them for potential harms, and remove them from autosuggestions before they cause widespread harm.

Taken together, features of autocomplete like temporal and physical distance between the originator and user, combined with unpredictable user behavior, relieve the provider of some responsibility for some epistemic harms. But they do not shield them. Although muted by time and distance, providers are in causal contact with each individual search. Providers cannot typically intervene in an individual search, but providers may intervene to affect aggregate search behavior,

as they already do by blacklisting certain autosuggestions, either through user settings like Google’s SafeSearch or by fiat. Given the kind of intervention technologically possible, search providers have a responsibility to the aggregate, not to each individual.

Finally, only search providers have the autonomy to significantly change the algorithm functioning. Typical users change the algorithm only incrementally. The algorithm itself includes self-modifying code, but its scope is narrow and unreflective to be left operating unsupervised. Only search providers have the autonomy necessary to bear the bulk of epistemic responsibility.

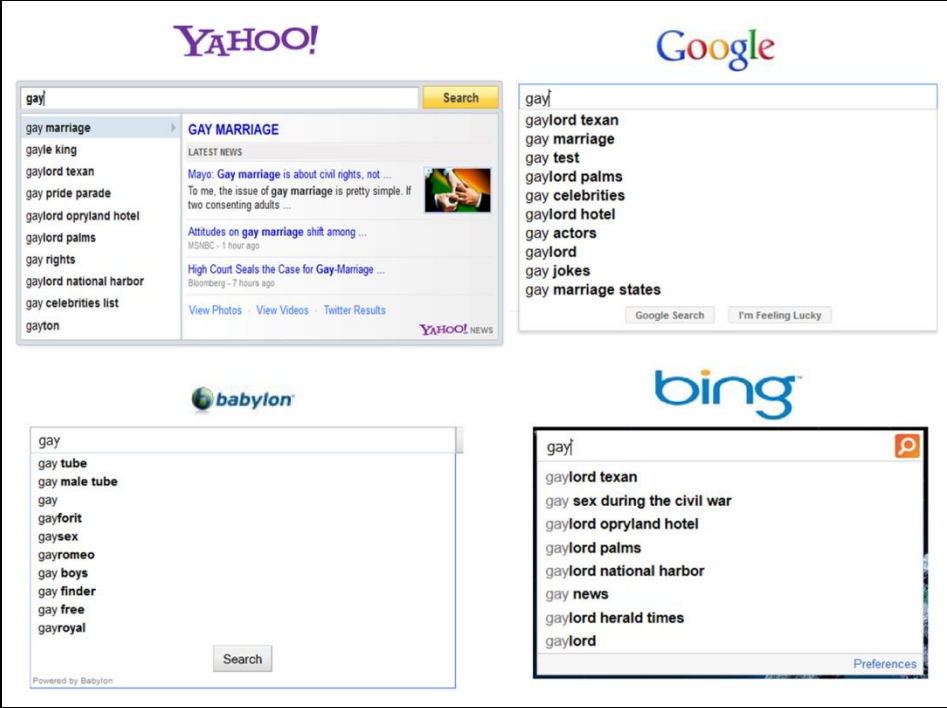


Figure 2. Autosuggestions for “gay” in four search engines (2 April 2013).

Epistemically responsible autocomplete

Which autosuggestions should be monitored? We identified three harms of irresponsible autocomplete. First, a search target’s dignity can be harmed due to the spread of misinformation about them. Second, autosuggestions may contribute to false, biased, or skewed beliefs in searchers. Third, extant biases can lead to incorrectly assessing the trustworthiness of informants who belong to certain groups and hinder the circulation of critical ideas in society. The first harm is moral, but the others are epistemic. We will focus on them. Our analysis entails that providers should eliminate at least three broad categories of autosuggestions: first, those that result from organized attacks to game the system; second, those that perpetuate damaging stereotypes about socially disadvantaged groups; and third, those that libel specific individuals.

Search providers already invest heavily in preventing attempts by “search engine optimizers” to artificially associate a website with popular search terms or

using robots or confederates to inflate its popularity directly. Whenever someone “games” the system, the overall quality of the search is damaged, so providers have selfish motives to monitor these abuses. But such efforts also have epistemic effects: whenever artificially inflated results are rejected, what remains is a better reflection of the Web makeup. Of course, the Web itself is full of misinformation and often biased toward privileged groups (Introna and Nissenbaum, 1999; Rogers, 2004: 3–9), but considering whether search providers are responsible to correct these features exceeds the scope of this article.

Second, search providers are responsible to counter damaging stereotypes about socially disadvantaged groups because they are likely to reinforce users’ biases. Stereotypes are especially damaging to groups without online bandwidth to counter them. Smaller, less privileged, poorer groups have a quieter voice on the Web, and the automaticity and feedback of autosuggestions can amplify these differences—so that common stereotypes appear frequently, while the contrary perspective shrinks disproportionately from view. Our prescription allows search providers to avoid compounding the propagation of such stereotypes by autosuggesting them without actually correcting the search results themselves.

Finally, individuals face major difficulties battling slanderous autosuggestions. Such autosuggestions should be avoided. Anyone interested in Wulff’s alleged salacious life could type “Bettina Wulff prostitute” and find the same information they would if the search term were suggested. But defamatory autosuggestions should not spontaneously appear to whoever simply types “Bettina Wulff.”

Objections

Let us last consider potential objections. First, speech should be protected on the Internet. Censorship is bad. We agree. But omitting autosuggestions is not censoring search results. The distinction is between finding information one is seeking and being involuntarily informed that certain information exists. Typing a full query invoking a stereotype would still yield the same hits as if it were autosuggested. Even searchers who do not explicitly invoke stereotypes in their queries will likely encounter them in the search results. But under our policy, these individuals would not have the stereotype thrust in their faces without context. Someone searching for information about, say, the history of an ethnic group would not be exposed to autosuggestions denigrating it. We do acknowledge a risk. Eliminating some autosuggestions could limit the “coverage” search engines provide and lead users to believe they have the whole story when some aspects of the story are withheld.

A related objection is that search providers should represent the Web as accurately and with little interference as possible. According to Rogers (2004: 28) and Rieder (2009: 143, 148), the Web is and ought to remain “a valuable collision space between official and unofficial accounts of reality.” They claim that the Web forms a non-hierarchical structure, where each page can link to other pages. A hyperlink is “an acknowledgement by one organization of another organization’s relevance to the discourse, based on some appreciation for that latter organization’s knowledge and reputation” (Rogers, 2002: 203). This makes the Web a neutral playground for different accounts of reality to compete for credibility by trying to get as many incoming hyperlinks as possible. The most hyperlinked webpages are perceived as most credible. Conflicting accounts of reality may gain

credibility within different clusters of webpages. Any interference by search providers in this naturally occurring hyperlinking structure would wrongly infringe an organic web-neutrality.

We question the assumptions underpinning this argument. Even if the Web once had a non-hierarchical structure, most traffic today occurs in a few mega-sites, such as Facebook and YouTube, which only partly implement a traditional static hyperlinks structure. Furthermore, are aggregated hyperlinks a good indicator of perceived credibility? Can we legitimately infer actual credibility from perceived credibility? We cannot delve into these issues here.¹¹ In any case, this objection applies only to *search results*, not *autosuggestions*. Autosuggestions are not intended to merely reflect an allegedly objective structure of the Web, but also users' search interests and patterns. Searches do not have a hyperlinked structure—they do not link to one another—therefore, this objection is irrelevant to them.

It might be argued that autocomplete should only neutrally and objectively represent searchers' interests because this is the best way to serve individual searchers' needs. Any other agenda, including avoiding epistemic harms, exceeds search providers' mandate. As noted, this is Google's official line, but in practice, Google and other providers do not follow it to the letter. They refrain from autocompleting searches for pornography, torrent sites, and some offensive content. Their *de facto* policies seem neither consistent nor defensible. Their pretense to neutrality and objectivity looks more like a façade for avoiding liability than a genuine commitment to users' best interests.

But suppose we take their claim to neutrality at face value. Is such a policy epistemically responsible? Suppose, for the sake of the argument, that when a father who worries that his daughter is overweight starts typing a query "is my daughter," it should be autocompleted with "overweight?" because this reflects his actual concern (even if unsubstantiated or prejudicial). Even so, such autosuggestions are epistemically irresponsible in aggregate because the search interface does *not* and *cannot* make such autosuggestions *only* to parents already concerned about their daughter's weight. It makes the same autosuggestion to nearly all users who start typing the same query, including people who have not had this thought. Such autosuggestions epistemically harm many users by reinforcing or planting prejudicial unsubstantiated thoughts. Even if autocomplete benefits users who are explicitly searching for this information, its potential harms to users looking for other information outweigh any alleged benefits.

Finally, search providers may object that rooting out prejudices is too burdensome. They would need to hire and train human minders to evaluate trending topics for potentially epistemically harmful content. Indeed, this may be overly burdensome for search providers to prevent all such harms. But the responsibility is to respond to aggregate behavior. The idea is to prevent the algorithm's damaging ampliative effects from swamping the unbiased or counter-biased alternative.

Conclusion

Autocompleted web search presents novel epistemic opportunities and challenges. Autocomplete inevitably and irreparably induces changes in users' epistemic actions, particularly their inquiry and belief formation. While these changes have

potential epistemic benefits for searchers and society, they also have harms, such as generating false, biased, or skewed beliefs about individuals or members of disempowered groups. Our analysis constitutes an example of assigning epistemic responsibilities associated with an epistemic technology on a principled theoretical basis. Using our theoretical framework that connects responsible epistemic behavior to practicability, we considered what corrective measures are technologically possible and desirable for each of the candidate responsible parties: searchers, the search technology and the search provider. Search providers have the bulk of required features and technological possibilities to identify and prevent harms. They therefore bear the bulk of responsibility to mitigate them and should eliminate autosuggestions that result from organized attacks to game the system, those that perpetuate damaging stereotypes about socially disadvantaged groups, and those that associate negative characteristics with specific individuals.

Author note

Miller and Record shared equally in the creation of this paper. Miller presented this work at a special workshop on "Crafting Ignorance: Secrecy and Suggestion in Science and Technology" organized in 2014 by the Science Technology and Society (STS) Cluster at the National University of Singapore. He thanks the organizers, especially Axel Gelfert, for their generous hospitality. Record presented this work at the "Philosophy of Information" conference at Duke University in 2014. He thanks the audience, especially Kay Mathiesen and Don Fallis, for valuable feedback and conversation.

Acknowledgements

We would like to thank Eleanor Louson for copyediting an earlier version of the paper and helping us to avoid several blunders, and David Enoch for his illuminating comments on an earlier draft.

Funding

Record is grateful to the University of Toronto's Critical Making Lab and AEGIS Ontario for institutional support. Miller is grateful to the Edmond J. Safra Center for Ethics, Tel Aviv University, the Dan David Foundation, Tel Aviv University, the Cohn Institute for History and Philosophy of Science and Ideas, Tel Aviv University, and the Sidney M. Edelstein Center for the History and Philosophy of Science, Technology, and Medicine, the Hebrew University of Jerusalem, for postdoctoral fellowships.

Notes

1. In January 2015, Wulff and Google settled the case before trial. Google removed the offensive autosuggestions (Huber, 2015).
2. Karapapa and Borghi (2015) list multiple legal cases around the world, in which Google was sued over harms allegedly caused by its autosuggestions. The outcomes of these cases have been diverse, and they have not followed a consistent legal logic (Karapapa and Borghi, 2015: 263).
3. Asynchronous communication may be implemented with client-side scripting, Active-X, Java, and Flash technology, but these methods are cumbersome and inefficient.
4. It is safe to assume that search providers other than Google use similar methods.
5. Google's Autocomplete cannot be turned off (Google, 2014).
6. Thanks to Adam Dubé for this point.
7. Thanks to an anonymous reviewer for this journal for this point.
8. For an overview of the debate about group responsibility, see Tollefsen (2004).

9. For an entity to be epistemically responsible for its effect on another's belief formation processes, it arguably has to have such an effect, be able to identify potential harms and alternative actions, and alter its behavior to avoid such harms. See Jonas (1984), Floridi and Sanders (2004) and Wallach and Allen (2009).
10. We thank Allan Olley for this reference.
11. For a discussion, see De Maeyer (2013).

References

- Adam A (2008) Ethics for things. *Ethics and Information Technology* 10: 149–154.
- Anonymous (2013) Google settles auto-complete defamation suit in Israel. *law.co.il*, 23 November. Available at: www.law.co.il/en/news/israeli_internet_law_update/2013/11/28/settlement-reached-in-google-auto-complete-defamation-suit/
- Baker P and Potts A (2013) “Why do white people have thin lips?” Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies* 10: 187–204.
- BBC News (2012) Google autocomplete: the internet where you are. *BBC News*, 7 September. Available at: www.bbc.co.uk/news/world-us-canada-19258912
- Berget G and Sandnes FE (2015) Do autocomplete functions reduce the impact of dyslexia on information-searching behavior? The case of Google. *Journal of the Association for Information Science and Technology*. Epub ahead of print 22 October. DOI: 10.1002/asi.23572. Available at: [dx.doi.org/10.1002/asi.23572](https://doi.org/10.1002/asi.23572)
- Coady D (2011) An epistemic defence of the blogosphere. *Journal of Applied Philosophy* 28: 277–294.
- De Laat P (2011) Open source production of encyclopedias: editorial policies at the intersection of organizational and epistemological trust. *Social Epistemology* 26(1): 71–103.
- De Maeyer J (2013) Towards a hyperlinked society: a critical review of link studies. *New Media & Society* 15(5): 737–751.
- Doubé W and Beh J (2012) Typing over autocomplete: cognitive load in website use by older adults. In: *Proceedings of the 24th Australian computer-human interaction conference (OzCHI '12)* (ed V Farrell, G Farrell, C Chua, et al.), Melbourne, VIC, Australia, 26–30 November, pp. 97–106. New York: ACM.
- Eisen M (2011) Amazon's \$23,698,655.93 book about flies. *It is Not Junk*, 22 April. Available at: www.michaeleisen.org/blog/?p=358
- Ekström M (2002) Epistemologies of TV journalism: a theoretical framework. *Journalism* 3(3): 259–282.
- Elers S (2014) Maori are scum, stupid, lazy: maori according to Google. *Te Kaharoa* 7(1): 16–24.
- Ernesto (2011) Google starts censoring BitTorrent, RapidShare and more. *Torrent Freak* 26 January. Available at: torrentfreak.com/google-starts-censoring-bittorrent-rapidshare-and-more-110126/
- Ettema J and Glasser T (1987) On the epistemology of investigative journalism. In: Gurevitch M and Levy MR (eds) *Mass Communication Review Yearbook*, vol. 6. London: SAGE, pp. 338–361.
- Fallis D (2011) Wikipistemology. In: Goldman AI and Whitcomb D (eds) *Social Epistemology: Essential Readings*. Oxford: Oxford University Press, pp. 297–312.
- Floridi L and Sanders JW (2004) On the morality of artificial agents. *Minds and Machines* 14(3): 349–379.
- Freiman O (2014) Towards the epistemology of the internet of things: techno-epistemology and ethical considerations through the prism of trust. *International Review of Information Ethics* 22: 6–22.
- Fricker M (2007) *Testimonial Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

- Friedman B (1990) Moral responsibility and computer technology. *Paper presented at the Annual Meeting of the American Educational Research Association*, Boston, MA. Available at: eric.ed.gov/?id=ED321737
- Frost-Arnold K (2014) Trustworthiness and truth: the epistemic pitfalls of Internet accountability. *Episteme* 11: 63–81.
- Garrett JJ (2005) Ajax: a new approach to web applications. *Adaptive Path*, 18 February. Available at: www.adaptivepath.com/ideas/ajax-new-approach-web-applications
- Goldman AI (2008) The social epistemology of blogging. In: van den Hoven J and Weckert J (eds) *Information Technology and Moral Philosophy*. Cambridge: Cambridge University Press, pp. 111–122.
- Google (2014) Autocomplete. Available at: support.google.com/websearch/bin/.py?hl=en&answer=106230
- Hiemstra D (2009) Information retrieval models. In: Goker A and Davies J (eds.) *Information Retrieval: Searching in the 21st Century*. London: Wiley, pp. 1–20.
- Huber J (2015) Bettina Wulff und Google streiten sich nicht mehr. *Der Tagesspiegel*, 15 January. Available at: www.tagesspiegel.de/medien/vergleich-noch-vor-dem-prozesstermin-bettina-wulff-und-google-streiten-sich-nicht-mehr/11236566.html
- Introna LD and Nissenbaum H (1999) Shaping the Web: why the politics of search engines matters. *The Information Society* 16(3): 1–17.
- Johnson MK, Hashtroudi S and Stephen LD (1993) Source monitoring. *Psychological Bulletin* 114(1): 3–28.
- Jonas H (1984) *The Imperative of Responsibility: In search of an Ethics for the Technological Age*. Chicago, IL: University of Chicago Press.
- Jordans F (2013) German court orders limits on Google hints. *The Boston Globe*, 15 May. Available at: www.bostonglobe.com/business/2013/05/14/german-court-orders-changes-google-search-hints/o7LdJiDAbdaVmQrOyrLdvL/story.html
- Karapapa S and Borghi M (2015) Search engine liability for autocomplete suggestions: personality, privacy and the power of the algorithm. *International Journal of Law and Information Technology* 23(3): 261–289.
- Kato MP, Sakai T and Tanaka K (2013) When do people use query suggestion? A query suggestion log analysis. *Information Retrieval* 16(6): 725–746.
- Kenyon T (2014) False polarization: debiasing as applied social epistemology. *Synthese* 191(11): 2529–2547.
- Knuth DE (1998) *The Art of Computer Programming*, vol. 3. 3rd ed. Reading, MA: Addison-Wesley.
- Kuflik A (1999) Computers in control: rational transfer of authority or irresponsible abdication of autonomy? *Ethics and Information Technology* 1: 173–184.
- Kumkale GT and Albarracín D (2004) The sleeper effect in persuasion: a meta-analytic review. *Psychological Bulletin* 130(1): 143–172.
- Latour B (1988) (aka Johnson J) Mixing humans and nonhumans together: the sociology of a door-closer. *Social Problems* 35(3): 298–310.
- Latour B (1996) *Petites leçons de sociologie des sciences*. Paris: Le Seuil.
- Latour B (2002) Morality and technology: the end of the means. *Theory, Culture & Society* 19(5–6): 247–260.
- Longino H (2002) *The Fate of Knowledge*. Princeton, NJ: Princeton UP.
- Lynch MP (2016) *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York: Liveright.
- Magnus PD (2009) On trusting Wikipedia. *Episteme* 6(1): 74–90.
- Mahdawi A (2013) Google's autocomplete spells out our darkest thoughts. *The Guardian*, 22 October. Available at: www.theguardian.com/commentisfree/2013/oct/22/google-autocomplete-un-women-ad-discrimination-algorithms
- Miller B and Record I (2013) Justified belief in a digital age: on the epistemic implications of secret internet technologies. *Episteme* 10(2): 101–118.
- Morozov E (2013) *To Save Everything, Click Here: Technology, Solutionism and the Urge to Fix Problems that Don't Exist*. London: Penguin.

- Munn NJ (2012) The new political blogosphere. *Social Epistemology* 26(1): 55–70.
- Niggemeier S (2012) Autocompleting Bettina Wulff: can a Google function be libelous? *Spiegel Online International*, 20 September. Available at: www.spiegel.de//;/google-autocomplete-former-german-first-lady-defamation-case-a-856820.html
- Nissenbaum H (1994) Computing and accountability. *Communications of the Association for Computing Machinery* 37(1): 72–80.
- O'Brien K (2013) Google eliminates “die” search suggestion for autism. *The Star-Ledger*, 27 February. Available at: www.nj.com/parenting/index.ssf/2013/02/google_changes_policy_for_auti.html
- Origi G (2012) Designing wisdom through the web. In: Landemore H and Elster J (eds) *Collective Wisdom: Principles and Mechanisms*. Cambridge: Cambridge University Press, pp. 38–55.
- Parasie S and Dagiral E (2013) Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago. *New Media & Society* 15(6): 853–871.
- Record I (2013) Technology and epistemic possibility. *Journal for General Philosophy of Science* 44(2): 319–336.
- Record I and Miller B (2016, forthcoming) Taking iPhone seriously: epistemic technologies and the extended mind. In: Pritchard D, Clark A, Kallestrup J, et al. (eds) *Extended Epistemology*. Oxford: Oxford University Press.
- Rieder B (2009) Democratizing search? From critique to society-oriented design. In: Stalder F and Becker K (eds) *Deep Search: The Politics of Search Engines*. Edison, NJ: Transaction, pp. 133–151.
- Rogers R (2002) Operating issue networks on the Web. *Science as Culture* 11(2): 191–213.
- Rogers R (2004) *Information Politics on the Web*. Cambridge MA: MIT Press.
- Rogers R and Ben-David A (2008) The Palestinian—Israeli peace process and transnational issue networks: the complicated place of the Israeli NGO. *New Media & Society* 10(3): 497–528.
- Simon J (2010) The entanglement of trust and knowledge on the Web. *Ethics and Information Technology* 12(4): 343–355.
- Simpson TW (2011) e-trust and reputation. *Ethics and Information Technology* 13(1): 29–38.
- Simpson TW (2012) Evaluating Google as an epistemic tool. *Metaphilosophy* 43(4): 426–445.
- Solomon M (2001) *Social Empiricism*. Cambridge, MA: MIT Press.
- Stephens-Davidowitz S (2014) Google, tell me: is my son a genius? *The New York Times*, 19 January, SR6.
- Sullivan D (2011) How Google Instant’s autosuggestions work. *Search Engine Land*, 6 April. Available at: searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592
- Thicke M (2013) The internet observer effect. *The Bubble Chamber*, 15 January. Available at: thebubblechamber.org/2013/01/the-internet-observer-effect/
- Tollefsen DP (2004) Collective intentionality. In: Fieser J and Dowden B (eds) *The Internet Encyclopedia of Philosophy*. Available at: www.iep.utm.edu/coll-int/
- Tollefsen DP (2009) Wikipedia and the epistemology of testimony. *Episteme* 6(1): 8–24.
- Trumbo J (2009) Driver fills tank in Richland, drains bank. *Tri-City Herald*, 26 February. Available at: [web.archive.org/web/20110813043134/http://www.tricityherald.com/2009/02/26/491472/driver-fills-tank-in-richland.html](http://www.tricityherald.com/2009/02/26/491472/driver-fills-tank-in-richland.html)
- Turner S (2013) The blogosphere and its enemies: the case of oophorectomy. *The Sociological Review* 61(S2): 160–179.
- Wallach W and Allen C (2009) *Moral Machines: Teaching Robots Right from Wrong*. Oxford: Oxford University Press.
- Ward D, Hahn J and Feist K (2012) Autocomplete as a research tool: a study on providing search suggestions. *Information Technology and Libraries* 31(4): 6–19.
- Weinberger D (2011) *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York: Basic Books.

White WW and Marchionini G (2007) Examining the effectiveness of real-time query expansion. *Information Processing and Management* 43(3): 685–704.

Wiideman S (2011) Google autocomplete study, 22 February. Available at: www.wiideman.com/research/google-autocomplete/study-results

Wray KB (2009) The Epistemic cultures of science and Wikipedia: a comparison. *Episteme* 6(1): 38–51.

Author biographies

Boaz Miller is a Teaching Fellow in the Graduate Program in Science, Technology and Society at Bar-Ilan University. His research encompasses the social dimensions of knowledge and explores the influence of the social context on the processes and products of inquiry with an eye to developing standards for their normative epistemic appraisal. Boaz holds a PhD and MA from the Institute for the History and Philosophy of Science and Technology (IHPST) at the University of Toronto and a BSc in Computer Science and “Amirim” Interdisciplinary Honors Program in Humanities from the Hebrew University of Jerusalem. His research interests are in social epistemology, new media, and philosophy of science.

Isaac Record is an Assistant Professor of Practice at Lyman Briggs College, Michigan State University, where he teaches courses in philosophy of science, science and technology studies, and critical making. His research seeks to situate our epistemic and ethical circumstances within a network of values, capabilities, and material and social technologies. Isaac holds a PhD and MA from the Institute for the History and Philosophy of Science and Technology (IHPST) at the University of Toronto and a BS in Electrical Engineering and BS in Computer Engineering from the University of Maine.