KNOWLEDGE, DEMOCRACY, AND THE INTERNET

Nicola Mößner and Philip Kitcher

Abstract

The internet has considerably changed epistemic practices in science as well as in everyday life. Apparently, this technology allows more and more people to get access to a huge amount of information. Some people even claim that the internet leads to a *democratization of knowledge*. In the following text, we will analyze this statement. In particular, we will focus on a potential change in epistemic structure. Does the internet change our common epistemic practice to rely on expert opinions? Does it alter or even undermine the division of epistemic labor?

The epistemological framework of our investigation is a naturalist-pragmatist approach to knowledge. We take it that the internet generates a new environment to which people seeking information must adapt. How can they, and how should they, expand their repertory of social markers to continue the venture of filtering, and so make use of the possibilities the internet apparently provides?

To find answers to these questions we will take a closer look at two case studies. The first example is about the internet platform *WikiLeaks* that allows so-called whistle-blowers to anonymously distribute their information. The second case study is about the search engine *Google* and the problem of personalized searches. Both instances confront a knowledge-seeking individual with particular difficulties which are based on the apparent anonymity of

the information distributor. Are there ways for the individual to cope with this problem and to make use of her social markers in this context nonetheless?

Introduction

It is sometimes claimed that the internet will lead to a "democratization of knowledge" (see e.g. Pfister 2011, 220) and that such "democratization" is a welcome development (see e.g. Coady 2012, ch. 6 and Munn 2012 – both with special focus on the blogosphere).[1] The thesis can be read in several different ways: one can focus on the agents who produce and consume information, on the processes through which information is distributed, or on the epistemological structure. So it is possible to understand democratization as bringing a wider range of people into the exchange of ideas, or as introducing new processes of information dissemination, or as changing the social-epistemic structure, for example by altering the character of the division of epistemic labor.

The internet not only allows for information consumption but also for the presentation of claims to knowledge by anybody having access and minimal technological skills. A common thought about internet democratization is that the inclusive technology of the web not only increases the amount of information available, but also allows claims to knowledge to emanate from a more heterogeneous collection of sources than those represented by traditional mass media.

---

[1] Daniel Jacob and Manuel Thomas (2014) critically discuss the internet's potential as a medium of communication on political democratization (see also the APUZ 62(7) 2012 and Ferdinand 2000).

A more radical reading of the thesis of democratization suggests a change in structure. In the age of the internet, people with access to huge amounts of information become cognitively more autonomous. They are no longer forced to rely on the opinions of a select group of experts.

In what follows, we aim to evaluate the claim of internet democratization of knowledge by analyzing which way(s) of reading the thesis can be maintained and which one(s) have to be rejected. In particular we focus on elucidating claims concerning a potential change in epistemic structure. Does the internet actually empower its users in becoming more autonomous epistemic agents? Does it change our common epistemic attitude of relying only (or primarily) on expert opinions? Do people instead rely on "algorithmic authority" (as defined by Clay Shirky), i.e. using search tools and websites recommending and distributing content that no human editor has suggested but was automatically filtered and put forward?[2] Does the internet alter or even undermine the division of epistemic labor? Our goal is to assess whether enthusiasm for "democratizing knowledge" is warranted.

Our approach takes the thesis as proposing an analogy between the political and the epistemic spheres: what once occurred in the history of today's democratic societies foreshadows similar processes in the structure of human knowledge today.[3] We call this the *analogy approach*. We distinguish three varieties of the analogy approach. (1) The internet does away with any conception of differential expertise: we are all equally authoritative with respect to any issue. (2) The internet reforms the prior division of epistemic labor (see Kitcher 2011), by giving a greater voice to constituencies that have previously not been

_____

[2] He discusses this in his blog on http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority/, accessed September 28 2015.

[3] Our focus is on epistemological claims only. Analysing the political dimension would deserve an article of its own. Readers interested in the latter topic will find a good starting point for further investigations in Cass R. Sunstein's work (see e.g. 2007, 2008) and Marianne Kneuer's edited volume (2013) on the topic.

heard. (3) The internet provides greater access to knowledge for many people (including a large number for whom access to knowledge has been limited).

The bulk of our attention will be given to (2) and (3). Before considering these theses we shall take a brief look at the more radical option, (1).

Radical Democracy

One way to develop the analogy approach is to start with the *theory* of forms of government. Blurring distinctions that have been made for two millennia, we can suppose that authoritarian governments are those in which political decisions are made by a proper subset of the mentally competent, mature members of a society without any consultation of those who do not belong to that subset: the extreme case is that of a tyrant who resolves issues as he sees fit. Democracy by contrast insists that the views of each citizen must be heard and must count. The analogy then proceeds by taking experts to be the counterparts of the privileged group of deciders.[4] Epistemic authority is the analogue of tyranny. To reject tyranny/authority in favor of democracy is to identify every mature, mentally competent person's judgment on any issue as being equal in status. The idea of a division of epistemic labor, according to which there's a partitioning of issues and the assignment of expertise to particular subgroups for particular types of questions, is to be abandoned.

Positions akin to this have been championed by some scholars who have protested the "hegemony of western science," perhaps with most flair by Paul Feyerabend (see Feyerabend 1978, Feyerabend 1987). Yet, as we have formulated it, the thesis faces severe objections.

---

[4] Here we are thinking in terms of direct democracy. The analogy breaks down with respect to representative democracy, since the experts might serve as the representatives who make the decisions. We are grateful to a referee for pointing this out to us.

The division of epistemic labor pervades all known societies and probably extends deep into our human past. People visit different places, undergo different experiences, and bring back information useful to others. To suppose that all opinions are on a par is to assign equal authority, even with respect to details of local geography, to those who live in a particular locale and those who spend their lives on a distant continent. Your judgment about the layout of the place you inhabit is, we believe, rather more likely to be correct than that of people who spend their entire lives in some antipodal region.

A more serious example extends the idea of differential access to different parts of nature by focusing on differences in training. When you have unusual stomach pains, you are probably inclined to consult a doctor. Doing so is typically more likely to bring accurate diagnosis and consequent relief than simply asking a random stranger, or even someone whose judgment you trust but who has had no medical training. Similarly, if your Google search throws up sites associated with the Mayo Clinic[5], you are probably going to do better than if it generates sources with no medical connections. The idea of equal epistemic status, across people and internet sites alike, is a myth. The positions advocated by Feyerabend, and articulated more recently by some feminist epistemologists (see Longino 2001) are far more sophisticated. These thinkers are not so much concerned to abolish the division of epistemic labor as to reconfigure it in a radical way, rescuing *some* sources that have previously been marginalized. The sphere of experts is to be expanded, so that on medical questions we do not simply focus on a particular form of training: western medicine is not to be the only game in town, but that doesn't entail including those who are clearly not playing. The

---

[5] The USA allows internet sites to offer medical advice, and some sites are associated with prominent hospitals and medical schools (the Mayo Clinic, Johns Hopkins, and so forth); other sites have a less distinguished background. In other countries, the law discourages physicians and hospitals from offering diagnostic suggestions, since if a diagnosis leads to unfortunate results, the site responsible for it would be legally liable.

proposed expansion might be challenged by appealing to the track records of various potential "experts", arguing that the newly included sources succeed at far lower rates. Yet that challenge leads into complex issues we shall not consider here. Besides the obvious need to attend to the details of the success rates, there is the tricky matter of probing the notion of success. Even if western rates of cure are far higher, that need not clinch the issue. Measuring "success" by comparing rates of recovery might fail to recognize important values forfeited in adopting the form of life in which western medicine is embedded. When a broader conception of the valuable life is adopted, the western conception of the body and its maladies might be seen to bring losses as well as gains. Or so the champions of expansion might maintain.

For our present purposes, we do not need to confront these complex issues. The simpler arguments about differences in epistemic access suffice to doom the version of the analogy approach that abolishes the division of epistemic labor. That sort of "democratization of knowledge" is not to be welcomed.

An Epistemic Framework

Before we proceed with the much more plausible versions, (2) and (3), it will be valuable for us to outline the epistemic framework that guides our inquiry and to introduce some terminology. As we see it, contemporary epistemology divides into a number of projects with different aims. A significant part of analytic epistemology, for example, is taken up with trying to specify the conditions under which individuals or groups have some epistemically important property (knowledge, justification etc.). In the context of the present discussion,

for example, one might seek a precise analysis of when a subject can be said to know something on the basis of an internet search.

The framework we adopt here aligns with a different tradition, the naturalist-pragmatist approach to knowledge. The primary purpose of that approach is to propose ways of refining the epistemic methods and strategies people and groups of people deploy, and exploring ways of adapting their knowledge-seeking to new environments. Following Peirce and Dewey the focus is on change of belief, rather than on belief, and the task is to extend the resources for successful change in view.

We suppose that children acquire, from their earliest years on, a complex of concepts, beliefs, values, and learning strategies. Initially, they are like sponges, sopping up whatever comes their way. Once they realize that apparently reliable sources of information sometimes disagree, they are forced to begin the enterprise of *filtering* those sources – an enterprise that will continue for the rest of their conscious lives.[6]

By the time people arrive at maturity, they have revised the initial mix of concepts, beliefs, values and strategies through a constant process of interaction with potential informants (including non-personal sources, texts, news media, and, these days, internet sites). They have far more sophisticated ways of posing questions, a far more extensive collection of beliefs, refined strategies for investigation and a rich set of markers for identifying the sources to which it's appropriate to turn in resolving various types of issues, together with a set of values that inspire them to direct their inquiries along particular lines.[7] Philosophers

---

[6] For an illuminating account of children's learning from others, see (Harris 2012). Our treatment of this topic has been aided by conversations between one of us (Kitcher) and Dorothy Chen.

[7] This process of knowledge acquisition is explained in detail by Edward Craig (1990). He emphasises that an inquirer or "seeker", as we shall call her, is in need of some *detectable indicator property* which tells her that a potential informant will tell her the truth as to whether that p (see ibid., 26f.). This kind of indicator is exactly what our "markers" are supposed to provide.

of science have long emphasized the theory-ladenness of observation (see Sellars 1953,

Hanson 1958, Kuhn 1962). Less noted has been the process through which people learn to

see, hear, touch and taste[8] – how those around us guide us in adjusting our bodily

orientation and motions, in focusing our attention differently, as we engage in more complex

forms of perception (for those learning about music: "Try not to concentrate on the melodic

line, but find a moment at which an inner voice first emerges, and then follow it"; or for

aspiring wine tasters: "Let the wine linger on your tongue, swish it around a bit, and pause

before swallowing").[9] Adult observation is framed by a thoroughly social ontogeny that blurs

any distinction between what a subject knows on the basis of testimony and what she learns

"on her own." Moreover, an important part of the mix is the output of the ventures in

filtering. Once epistemic development is well under way, subjects are set up for further

learning from others (knowledge on the basis of testimony: see Mößner 2010) by the *social*

structuring of the channels through which they will accept information. Their socialization

has directed them to look in certain places for the answers to certain questions. The social

markers they now use to rate prospective sources have emerged from the early efforts at

filtering, when they discovered that not every informant is reliable. Socially shaped

observations of the environment play a role in calibrating others, but at each stage the

maturing subject depends on the concepts, beliefs, values, and strategies in place at that

point. Moreover, large parts of the ambient consensus about the division of epistemic labor

---

[8] A crucial exception in this context is Ludwik Fleck (1979, orig. 1935). He has pointed out the relevance of practical training in science to teach students how to perceive correctly, i.e. in accordance with the prevalent thought style (see also his paper "To Look, to See, to Know" 1986, orig. 1947).
[9] Lorraine Daston and Peter Galison offer an exceptionally illuminating overview of how scientists learn how to observe (see Daston and Galison 2007).

will typically go unquestioned: particular kinds of educational institutions, media sources, books, and experts will be endorsed because society has given them a privileged status.[10]

Approaching knowledge in the way just sketched often generates skepticism – or at least unease. Might it not all have started so badly that people are condemned to be mired in error forever? The picture we've painted of the developing cognitive subject has the same structure as a plausible picture of the history of inquiry. Original chaos gives way to more successful constellations of beliefs; more adequate ways of categorizing the world of experience become entrenched (in the sense of Goodman 1953); observation is refined and expanded; the social structure of inquiry divides epistemic labor at a finer grain. The sense of all this as a progressive venture is grounded in an increasing ability to intervene in nature: inquiry (across a broad spectrum of fields including the humanities and social science as well as the natural sciences) enables us to do things that our predecessors would have thought of as miraculous.[11] Just the same pattern of advance in interventional success is discernible in the development of the individual cognitive subject, as she grows from infancy to maturity.

Our outline of a naturalist-pragmatist epistemology is intended to set the stage for the epistemic questions we aim to consider. The internet generates a new environment to which people seeking information must adapt. How can they, and how should they, expand their repertoire of social markers to continue the venture of filtering, and so make use of the possibilities the internet apparently provides?

---

[10] Miranda Fricker's work on epistemic injustice (2010) elaborately shows how these mechanisms of the acquisition of social markers can also lead to the marginalization of particular groups of people as epistemic agents.

[11] Although we allude here to the famous (notorious?) "No Miracles" argument for scientific realism, we are not committing ourselves to any realist stance. Successful intervention is simply seen as a measure of progress, without supposing progress to be connected with the discovery of (correspondence) truth. Our approach is thus neutral between the positions adopted in (Laudan 1981), (Kitcher 1993) and (Mößner 2012, 2013).

The predicaments we want to consider involve individuals who wish to know the answers to particular questions: we shall refer to them as *seekers*. A *source* for a seeker is something (or someone) that delivers potential answers to the focal questions. Seekers select sources, modifying their beliefs in accordance with those sources, by applying *markers*, properties they take to be indices of reliability. Prior to the internet age, seekers, having acquired during their development a socially-shaped menu of markers, used those markers to differentiate sources. What happens when new voices are added? Can seekers use their acquired epistemic resources to validate some internet sources or to generate new markers that would distinguish among internet sites? How should the next stages of epistemic development go forward?

It is important to recognize that misidentification of sources is not the only potential problem with a division of epistemic labor. Seekers might be unable to use the markers at their disposal to find any sources at all, or, if they found them, to extract from them the information they need. It's easy to forget that ignorance can be a predicament deeper than simply not being able to generate for yourself the answer to a pressing question. You may not have any clues about where to look for help. Or, if you can identify potential sources, you may find them completely impenetrable. Thus, besides dangers of misidentifying sources there is also a failure of transparency and, correlated with this, a problem of access. So, to summarize, we have:

1. *Misidentification problems*: the socially prevalent division of epistemic labor, to which a seeker might turn for extending her filtering of sources, assigns markers that pick out potential informants who are not reliable sources, or that fail to apply to the reliable sources.

2. *Transparency problems*: the socially prevalent division of epistemic labor only assigns markers that seekers are unable to apply and, thus, the range of potential sources remains opaque to the seeker.

3. *Access problems*: the socially prevalent division of epistemic labor only assigns markers that pick out sources from which seekers cannot extract the information they seek.

We now turn to consider aspects of the internet that may give rise to these kinds of problems.

Sources of Trouble: Inclusion and Epistemic Opacity

The analogy approach to internet democratization focuses on political *practice*, rather than on theory. Democracies change political life by involving a larger number of people in political decision-making. Applying this to the epistemic domain would conceive democratization as increasing inclusiveness. But the participants in epistemic practice might play either of two distinct roles, serving as contributors to knowledge or recipients of information. So the analogy yields two interesting theses, (2) and (3). (2) celebrates the internet for its provision of greater opportunities for *producing* knowledge; (3) welcomes the increased options for knowledge *consumption*. As we'll see shortly, there's an interesting tension between these two effects. The more inclusive web-based technologies (especially Web 2.0) become, the larger the difficulties confronting a potential consumer who hopes to use internet sites, databases or social media platforms as epistemic sources.

Recent developments in internet technology – social media platforms like Google+ or Facebook – provide many new opportunities for people to go beyond passive consumption

of information.[12] Whether in the form of "likes" or by making explicit comments, users of

these platforms can express their views about many different things. Part of this

development is akin to pervasive practices of reporting what has been experienced on one's

daily journeying, and doubts about the informant's competence don't typically arise when

someone is reporting on incidents known to take place in her vicinity. Yet, if the source is not

someone familiar[13], someone with whom there will be ongoing interactions and available

procedures for responding to unmasked deception, questions of sincerity do arise. How

does a seeker discern those posts that misinform – either because they are intended to

mislead, or because they are made in a spirit of fun[14], on the assumption that nobody would

ever take them seriously[15]?

There are ways of overcoming the problem, at least in some instances. If you simply claim to

have observed a rare meteorological phenomenon or a bird of a species thought to be on the

verge of extinction, those who do not know you may wonder if you are to be trusted, even if

they have no doubts about your observational competence. To allay their doubts, you can

post a photograph – or, even better, a series of photographs, for that would allay worries that

a previously published image has been embedded in a newly photographed background.[16]

The strategy here extends one that has been familiar to scientists (or "natural philosophers")

since the early days of the Royal Society: the phenomena are presented so as to make virtual

---

[12] In our usage, "information" marks out items of *potential* knowledge. It may contain opinions about certain topics as well as neutral data.

[13] Just consider how many "friends" people claim to have on Facebook and, accordingly, how the concept of friendship has shifted in these contexts.

[14] At first glance it might seem indifferent in epistemological terms whether we receive wrong information by people who deceive intentionally or just for the fun of it. Nonetheless, it seems worthwhile noticing that it is only through the advent of internet technology that attempts of the latter kind increased in number as to become worthy of epistemic consideration.

[15] Take as an example the boy who used to regularly write Wikipedia entries to train his fictional writing capabilities (see Magnus 2009, 77).

[16] On photographs as scientific evidence see also Mößner (2013).

witnesses of those who did not actually attend an experimental demonstration (see Shapin and Schaffer 1985).

Would-be informants succeed only if they can overcome the doubts of those who seek information of the sort they offer. Seekers only increase their knowledge if they can find the genuine sources. Initially, it appears that the new voices heard on the internet cannot be assessed by the markers seekers have learned to use. For their credentials, as well as their characters, are unknown. Potential sources, it is often said, are "anonymous."[17]

Anonymity isn't the pertinent notion. What matters is a related concept we'll dub *epistemic opacity*. We often trust a source of information, even when we don't know the name(s) of whoever is responsible for the words we read or hear. Much of the time a source can be recognized without attribution of authorship. It's enough that the document was produced in what we take to be an appropriate way: it comes out of a trusted institution or a process we deem reliable, it's published in a reputable journal or written by people with impressive credentials. Names aren't needed.

What is at stake in worries about the internet is epistemic opacity. A source is epistemically opaque for a seeker when the seeker cannot apply the markers available so as to vouch for the reliability of that source. You can be fully aware of the name on the internet post about the health effects of drinking a glass of red wine a day, but that doesn't help in deciding whether to lay in a case of pinot noir. What you need to know is whether this source has any access to relevant evidence and whether the source can be trusted to offer an impartial perspective.

---

[17] Karen Frost-Arnold (2014) offers an interesting discussion of anonymity on the internet. She does not, however, make the distinction between anonymity and epistemic opacity that we consider important here. We are grateful to a referee who urged us to be clear about this distinction.

Our imagined informants about the unusual weather event or the rare bird used an obvious strategy for overcoming the problem of epistemic opacity.   They attempted to turn a case of testimony into one in which people can "judge for themselves." We shall be principally concerned with instances in which this cannot be done, cases in which there's no serious possibility of "providing the evidence." First, however, we want to note some more mundane issues about thesis (3), questions that are prior to any process for evaluating potential internet sources.

Residual Inequalities

How exactly might the internet enable people to have greater opportunities of "consuming" knowledge? Before the digital age, access to information was obviously limited by the economic status of the seeker, her being a member of an organization (e.g. a club, a political party, or a company) and due to the level of her membership as well as to the organization of the ambient society. The poor cannot buy books, and sometimes must make sacrifices to purchase even the least expensive printed media. In some societies public libraries are inadequate, and, even within affluent societies, there are communities within which the library collections are extremely limited. Within some domains, what is readily available in print is highly technical.

People who can afford computers and who can connect to the internet are now able to sample from a vastly larger menu of potential sources.[18] Some of those sources are deliberately designed to present complex topics in ways that do not presuppose any

---

[18] Although it has to be added that it still depends on your local background how much information is really available to you. People in China, for example, still have to face massive governmental censorship and, as a more general difficulty, particular language skills are required to access international websites.

extensive education, for example by using not only static visual representations but also animation, video clips and the like. In this sense, they continue Otto Neurath's project of education with the aid of visual means (see Neurath 1991). Problems of making visual contact with pages and being able to understand what those pages have to say have been substantially ameliorated.[19]

Plainly, however, democratization of this sort is imperfect. Not everyone can afford a computer (or tablet). Not everyone can go to a place in which free internet service is provided. There are residual inequalities excluding some of those who were shut out in the pre-digital era. Yet it would be wrong to underrate the magnitude of the change. As devices for using the internet become ever cheaper, opportunities have expanded for a large population that, even a decade ago, would have considered the technology of that time beyond their budgets. Communities that provide their schoolchildren with computers for classroom use, cities that set up a free public wi-fi system, societies that offer public spaces with computers for anyone to use and that provide free courses in how to make the best use of them, all point towards a future in which the democratization of consumption can be further advanced – perhaps even toward a final state in which the residual inequalities vanish, one where every human being who could engage in an internet search has the capacity to do so.

A Harder Problem

The issues just briefly reviewed are relatively easy ones. More difficult are those arising from epistemic opacity, and from the threat that those newly equipped to search the web for

---

[19] There are, for example, attempts to offer automatic translations (see https://translate.google.com/, accessed December 14 2014) and also keyboards to write and read internet sites in braille.

information will be unable to sort out the wheat from the chaff. How can seekers' stocks of social markers evolve to cope with the highly diverse range of potential internet sources?

Many internet sites provide information from anonymous sources, but they are not epistemically opaque. Consider WikiLeaks. Strict anonymity protects the whistleblower whose report we read. Yet that report comes to us through a filter, and any doubts we might have about its reliability can be assuaged if we know enough to assess the filtering process. Of course, the process of information transmission might take different turns here. A recipient might consult directly the homepage of WikiLeaks or she might, for example, read information leaked on that platform in her daily newspaper. In the latter case, journalists of traditional media will add their professional evaluation of the source's trustworthiness to the epistemic process. In this latter instance, it might thus be argued that it is the journalist's judgment the recipient relies on in her search for information instead of trusting WikiLeaks. Yet, what happens if this professional intermediary is not put to work?

What do we know about the route between the reception of a whistleblowing report by WikiLeaks and the eventual presentation of that report on the site? The following items of information are in the public domain: (1) WikiLeaks staff check incoming reports for plausibility; (2) WikiLeaks staff analyze the document for internal clues to any potential fabrication; (3) WikiLeaks staff include people with significant international knowledge, who have impressive educational credentials; (4) WikiLeaks staff include people with skills in the computer analysis of documents.[20] On the basis of (1) and (3), we might conclude that fake reports might be detected on grounds of implausibility; on the basis of (2) and (4), we might conclude that fake reports might be detected through an analysis that revealed how they had

---

[20] For information on this topic see https://wikileaks.org/About.html, accessed September 27 2015.

been cobbled together. Further, we know that WikiLeaks is only concerned to publish the most significant reports it receives, and that it will give greater prominence to cases in which the exposure promises to cause a large shift in public opinion. The more significant the report, the more likely that those whose conduct is exposed will come forward with documents that contest the claims. So the would-be faker faces a dilemma: if the report has the features that might lead WikiLeaks to give it some prominence, it's likely to attract efforts to rebut it, and to track down the culprit; if the report lacks such features, it may fall below the threshold for WikiLeaks to post it.

The dilemma resembles one that arises in the context of scientific fraud. Scientists who dry-lab predictable extensions of well-established results are likely to escape detection, but their fraudulent papers are unlikely to be published, or, if published, to make much impact. Those who aim for more exciting conclusions will attract attention, and, unless they are very lucky, subsequent experimental work will undermine their claims. The dilemma is not watertight, of course. Familiar examples of relatively unambitious scientific fraud – such as the case of Robert Slutsky (see Engler *et al.* 1987) – reveal how dry-labbing can find just the right level. Perhaps there are analogous instances with respect to fabricated whistle-blowing. Nevertheless, the considerations from significance combine with the conclusions about WikiLeaks filtering to support extending previous markers to validate a new source.

Prior to encountering WikiLeaks, seekers are able to mark some sources as reliable by considering the processes through which those sources arrive at a decision to publicize a putative piece of information. Part of what has been recognized in the ontogeny of their ability to learn from others is that some potential sources attempt to check the statements they pass on, that the checks are likely to be more effective if those who do the work have

particular credentials (advanced degrees, the ability to deal with documents in a variety of languages), that greater publicity increases the costs to those whose shortcomings are exposed and therefore inspires attempts to rebut the charges. The acquired markers can be assembled in a novel constellation to apply to the process of filtering the whistleblowing reports. Furthermore and to come back to the role played by traditional media, WikiLeaks reports are often brought to public attention by journalists. Thus, having learned which sources you can trust in the media sector, markers already available can be extended to those on whom the trusted media sources rely. Although the authors of those initial reports remain strictly anonymous, there is no epistemic opacity – and hence a smooth extension of everyday social epistemology. The extension is achieved in two, mutually reinforcing ways: either we can use our previous markers for counting processes of knowledge certification as reliable, or we can appeal to established reputation markers (the credit of the traditional media source).

The case of Wikipedia is similar, with an interesting twist. In the early days of the site, many seekers engaged in academic research were disinclined to trust Wikipedia entries. Their skepticism rested on Plato's old worry: the clamor of the uninformed would dominate. Yet, precisely because of that early skepticism, and because of the potential utility of the site, the situation changed. The Wikipedia staff explicitly encouraged contributions from people informed about technical topics. Academic seekers not only started to submit their corrections to articles within their domains of expertise, but this fact became widely known. Judith Simon (2010) discusses technological modifications of Wikipedia which were introduced to enhance the reliability of the online encyclopedia and, thus, its trustworthiness as an epistemic source. In particular she explains the *WikiScanner* (i.e. a

search tool that traces IP addresses and, thus, helps to disclose who is responsible for particular edits of content) and the *WikiDashboard* (i.e. the *talk pages* mentioned below). Both developments are meant to increase transparency. "Rendering the sources and editing patterns visible enables rational assessment of content provided on Wikipedia" (ibid., 350). Although Simon admits that these technological devices are useful tools to enhance the reliability of Wikipedia, she thinks that further modifications are necessary. Since not all users are aware of the relevance of talk pages and WikiScanner, she suggests that the results of these investigative tools should be connected to the respective content pages and thus draw the user's awareness to its degree of trustworthiness calculated on basis of this data (see ibid., 351).

Actually, technological developments have already altered the stage. WikiScanner was taken down in 2013 due to financial reasons (see http://virgil.gr/wikiscanner/, accessed June 22 2016).[21] There are, however, follow-up projects such as WikiWatchdog (see http://wikiwatchdog.com/, accessed June 23 2016), an open source software tool that allows to search for anonymous edits from certain organizations (e.g. political parties or companies).

Despite these changes in the technological landscape, Simon's suggestion has been realized in a way in the German version of Wikipedia.[22] Here, articles are recommended to readers and marked as *articles worth reading* (in German "lesenswerte Artikel", see https://de.wikipedia.org/wiki/Wikipedia:Lesenswerte_Artikel, accessed September 28 2015) or even as *excellent* ones (see https://de.wikipedia.org/wiki/Wikipedia:Exzellente_Artikel,

---

[21] We are grateful to an anonymous (sic!) reviewer for making us aware of these developments.
[22] It has to be added, however, that Simon suggested the implementation of marking processes based on algorithms, whereas the processes established are not automatically performed but still depend on users' reviews.

accessed September 28 2015). Moreover, Wikipedia also asks its users to mark articles of

bad quality and provides a list of potential shortcomings that could be indicated (see

https://de.wikipedia.org/wiki/Wikipedia:Bewertungsbausteine#M.C3.A4ngel, accessed

September 28 2015).

A consequence of these developments is that views of the process through which Wikipedia

entries came to be posted were modified, and, as in the example of WikiLeaks, familiar

markers could be applied to that process, validating Wikipedia as a source. The evolution of

Wikipedia's status culminated in the comparison of Wikipedia's accuracy with that of

traditional encyclopedias (see Fallis 2011). Previously skeptical seekers were able to deploy

their established markers to validate both the comparison and the sources now identified as

inferior.

However, a remaining problem concerning the role of expertise in Wikipedia entries is

brought to our attention by Lawrence M. Sanger (2009) – one of the founding members of

Wikipedia. He points out that the egalitarian ideal of Wikipedia drives off experts in the long

run (see ibid., 65). As there is no decision-making authority in the background, contributors

to Wikipedia are advised to discuss changes concerning particular articles on correlated *talk*

pages. On these pages authors and editors meet on an equal footing. Hence, cases of

continuous disagreement might occur, known as "edit wars". In such situations, more often

than not experts will back out as they do not have time for endless discussions.

Consequently, though experts might play a role in building up the stock of Wikipedia entries,

their contributions do not stay untouched, but will deteriorate in quality over time.[23]

---

[23] Sanger describes the qualitative development of many Wikipedia articles in the following way: "Over the long term, the quality of a given *Wikipedia* article will do a random walk around the highest level of quality permitted by the most persistent and aggressive people who follow an article" (Sanger 2009, 64).

Not all successful extensions of seekers' powers to recognize sources are like these, however. WikiLeaks and Wikipedia operate in many domains quite remote from those in which people have substantive prior knowledge – that is why they are so useful. Seekers turn to them for information of types they can't independently check. When the topic concerns matters a seeker can subsequently evaluate for herself, the extension of markers goes differently. The seeker consults Tripadvisor before booking a hotel room, or reads the Amazon reviews before ordering a product. In advance she is quite uncertain about the process through which the "scores" on the sites are generated. Careful reading of the reviews might help to decide whether the features she is most interested in are found in the hotel or in the product, but her perusal will often be insufficient to yield a confident verdict. Perhaps all those who have submitted their assessments are so different from her in their judgments that the recommendation she gleans from the site is quite unreliable in anticipating her own reaction. Yet often there will be little harm in experimenting. She can try the recommendation and see whether her own subsequent judgment accords with it. Proceeding in this fashion, she calibrates various sites, learning inductively which ones are good guides for her own choices. (Perhaps she even comes to "follow" particular reviewers – unconcerned by the fact that they write under pseudonyms, since what matters is the similarity between their judgments and hers.) Here, there is no endorsement of a process. The seeker understands that the site simply records the opinions of the people who submit reports to it, and the collection of reviews might express judgments highly discordant or perfectly harmonious with hers, or anywhere in between. Because she can establish a track record, she can find her way to sites, or to contributors, that serve as sources for her.

To summarize, we've seen two ways in which the worry about epistemic losses can be overcome. The anonymity of the internet need not imply epistemic opacity, and seekers can deploy their antecedently available markers to validate the processes through which statements are posted, thus identifying new sources.[24] Alternatively, epistemic opacity might only be a transient problem, as seekers' independent abilities to check conclusions enable them eventually to identify some sites (or contributors) as sources and others to be avoided.

The sites we have so far considered prove relatively conservative in refining the traditional division of epistemic labor. From the consumer's perspective, the problem of deciding whether to trust a new source is resolved relatively easily: seekers can separate *knowledge* from mere *opinion*. Conservative extensions of the seeker's stock of social markers emerge because, in a new learning environment, the existing division of epistemic labor is straightforwardly extended. So there is democratization of type (3), but far less (if any) of type (2). The newly validated sources are, in effect, the old experts, now available to deliver their well-grounded views over a wider set of channels.

Going Further?

How far should optimism extend? A more general and more thoroughly democratic process is more widely available, and characteristic of a far larger range of internet sites. Any voice can be heard. There are to be no checks or encouragement of contributions by people antecedently designated as experts. Instead, the public discussion, in its full form, is available to seekers, who can form their own views after reviewing it. Internet democracy,

---

[24] Basing our reliance on processes instead of people's trustworthiness with respect to web content is also suggested by Simon (2010).

like democracy in general, thrives on the clash of opinions in a public arena. Mill already pointed the way.

Mill's famous – and inspiring – brief for free and open discussion rests on a number of assumptions, some but not all of them explicit in his writings. He imagines a debate about some large (morally or politically important) issue, in which various perspectives are publicly presented. People concerned to make up their minds on this issue are able to attend to all that is said in this debate. At the conclusion of the presentations, these people are able to weigh the evidence offered, to assess arguments and counter-arguments, and they make their decisions on this basis. If those decisions are instrumental in yielding a policy for the society to which they belong, this procedure is the best that society could devise.

We have no quarrel with Mill's specification of an *ideal*, the ideal of the *Millian arena*. Nevertheless, it's important to recognize that the arena can only generate the supposed benefits, only qualify as the best way of social decision-making, if certain conditions are present. To put the point negatively, there are plenty of ways in which things can go badly wrong.

1. Those who decide may not understand the evidence presented, or even the question at issue.

2. The representation of alternative perspectives may be incomplete.

3. Some perspectives may receive more "air time" than others.

4. Presentations may include false statements whose falsehood the audience is in no position to detect.

5. Some perspectives may be presented with more rhetorical skill than others.

6. Those who decide may lack the skills required for proper weighing of the evidence.

The items on our list are matters of degree, and actual public debates may involve all of these shortcomings without compromising Mill's case for the value of free and open discussion: even though social decision-making doesn't go perfectly, it may be a satisfactory approximation to the ideal. Yet there are examples in which the lapses are truly egregious. Discussion of climate change (particularly in the USA) is a case in point.

The evidence for climate change is intricate and technical, and serious study is required to grasp it. Prominent climate scientists have worked hard to write for the general public (Hansen, Schneider, Mann), but their books and articles are probably still too difficult for at least half of the American electorate. Even the question of climate change is badly understood: many people confuse the claim that the global mean temperature is increasing with the thesis that *every* place on the planet is getting warmer. Public discussions of climate change almost never convey all the potential differences of the future environment; temperature is the focus, and such effects as ocean acidification are usually ignored. Thanks to the influx of funds from companies that profit from fossil fuels, perspectives denying climate change or minimizing its significance receive a disproportionately large share of the time or space available (see Oreskes and Conway 2008). The spokesmen for these perspectives often make claims climate scientists have rebutted again and again – and continue to reiterate the claims without citing any new evidence in their favor. Much effort is devoted to "marketing" the idea that climate change is "bad science" (the "Hide the decline" video is a powerful example[25]). Finally, the task of identifying and weighing the evidence is extremely onerous – and probably beyond the skill of anyone outside the expert community.

---

[25] See https://www.youtube.com/watch?v=WMqc7PCJ-nc, accessed December 16 2014.

The result of this caricature of the Millian arena is an American public that does not realize the importance of the issue, and cannot arrive at any reasonable judgment about it. Pessimism about internet democratization of knowledge stems from thinking that large swaths of online discussion will be pervaded by similar failures, possibly in even more extreme forms. When seekers move away from organized filtering found in WikiLeaks and Wikipedia to which they can expand their usual social markers, and when they leave behind the relatively mundane topics on which they can use their own judgments to calibrate sites, they will have no satisfactory markers for assessing the credibility of the statements made. Just as public presentations on controversial issues often reflect the attitudes wealthy donors wish to inculcate, so too internet sites will serve the wishes of those whose contributions, or whose advertising, support those who labor to make them attractive. Just as citizens turn for information to media whose messages resonate with their antecedent predilections and values, so too the internet will be partitioned into niches, serving to "inform" people who want to protect the beliefs they already hold.

The deepest form of pessimism worries about what we'll call the *cascade*. At the heart of Clifford's celebrated analysis of the ethics of belief is the hypothesis that, as erroneous views become prevalent, the judgment of new items of potential information becomes ever more compromised (see Clifford 1999). Deep pessimism views the early stages of internet democratization as involving an accumulation of false beliefs that eventually undermines seekers' abilities to deploy markers to identify genuine sources. People become so confused about the truth that their body of beliefs can no longer work in tandem with previously useful markers. Error begets more error and erodes the practice of relying on others. There is a disastrous cascade.

This problem becomes even more pressing when taking into account that internet technology can screen-off divergent opinions. This worry is expressed by Cass R. Sunstein who points to the fundamental role that the public forum plays in democracy (see Sunstein 2008, 96ff). Such a forum allows "chance encounters" with kinds of information and alternative opinions not pre-selected. Thus, the trend to fragment societies into groups that draw from a restricted subset of internet sources will reduce any broad public forum. In isolated groups, easily generated via internet personalization strategies which we will discuss next, opinions are constantly reinforced and become more and more extreme (see ibid., 99ff.). Within such communities, Clifford's cascade finds an amplifying environment in which confirmation biases flourish.

The Dangers of "Personalized" Searches

One possible way in which people can become locked in to misguided strategies for modifying their corpus of beliefs stems from the use of personalized searches.[26] Following Neil Thurman, we distinguish two types of personalization: "Explicit personalization uses direct inputs; implicit personalization infers preferences from data collected ..." (Thurman 2011, 397). In either case, the hit list resulting from a search is shaped by factors beyond the search terms entered – either from other information consciously entered by the seeker or by the "profile" the search engine (e.g. Google) has previously constructed. Google employs a policy of mandatory implicit personalization, using data gained by previous logins to different Google services (for example, the social media platform Google+, the email program Gmail, earlier Google searches), together with web history cookies. The result of

---

[26] This problem is also discussed by Boaz Miller and Isaac Record (2013), though they do not take into account the distinction between explicit and implicit personalization.

the profiling is to order the sites presented to the seeker, whereby the ordering is claimed to be in accordance with the relevance of the sites listed to the seeker's search request. The more data Google is able to collect, the closer (presumably) the profile fits the seeker's attitudes, interests, and preferences, so that the sites appearing at the top of the hit list are likely to be those the seeker will find most relevant.

The basis of Google's personalization strategy is its PageRank algorithm. It evaluates the number of links of a particular website with other sites on the Web. The more links are connected with a particular page, the higher it is ranked in the resulting hit list. This initial idea of the Google search engine has been continuously developed further over the last years so that today the company announces that it makes use of "more than 200 signals and a variety of techniques" to rate the indexed websites (https://www.google.com/intl/en/about/company/philosophy/, accessed November 07 2014). Hence, when personalization proceeds by constructing profiles that track the seeker's earlier search behavior, the list presented will be tailored to give high priority to sites that connect to others the seeker has previously visited. Assuming that linkage indicates like-mindedness, it is not hard to understand how implicit personalization inclines seekers to remain in the epistemic circles in which they have previously traveled – thus generating the possibility of constantly relying on a family of unreliable sources, the worry we derived from Clifford.

Yet it is easy to understand how personalization, even the implicit personalization that appears to cause trouble, is attractive. As several authors have argued, seekers need help in dealing with the information overload on the Web (see Beam and Kosicki 2014). Especially for those unskilled in choosing search terms, seeking answers to specific questions is often

frustrating, the would-be search yielding, again and again, a host of irrelevant discussions. When search engines accurately construct a profile of the seeker, the chances of steering her to pertinent sites go up. From an epistemological point of view, however, the dangers of cascading misinformation seem to outweigh the benefit – at least with regard to contemporary technology.

One difficulty stems from the fact that many people aren't aware of implicit personalization. In consequence, they may be led to conclude that the hit list on their screen presents the most authoritative website at the top of the list, or that it offers a comprehensive view of the topic. Some, perhaps many, seekers welcome the internet as an opportunity for enlarging their sense of the range of views with respect to a particular question, and constructing profiles is diametrically opposed to that background interest. In principle, this class of problems could be readily solved. Search engines might be required to provide both (implicit) personalization and non-personalization options, available to a user with a single mouse click.[27] In that way, seekers would be helped to achieve the *objective perspective*, celebrated by Thomas W. Simpson. "'Personalization' is the development in Internet-delivered services ... and consists in tailoring online content to what will interest the individual user. But objectivity may require telling enquirers what they do not want to hear, or are not immediately interested in. So personalization threatens objectivity. Objectivity matters little when you know what you are looking for, but its lack is problematic when you do not know" (Simpson 2012, 427). Clearly, objectivity, taken as a neutral and independent perspective on what is going on in the world, is threaten by personalization strategies that are likely to introduce confirmation biases.

---

[27] A similar suggestion is put forward by Simon (2010). The focus of her proposal, however, are so-called recommendation systems such as Tripadvisor and not search engines such as Google (see ibid., 353f.).

The hypochondriac who worries constantly that he will develop a gastric ulcer returns again and again to sites about ulcers. When severe stomach pains actually occur, his search takes him to sites that magnify his anxieties – whereas any competent physician would have diagnosed a case of indigestion. An important difference between the doctor and the search engine, of course, is that the former is concerned with the patient's welfare.

Experts who recognized a seeker's desire for a comprehensive view of a controversial topic, and who were concerned to help her, might recommend a variety of sources, the best representatives of the major rival positions. Teachers often aid their students in this fashion. Could companies offering web services such as Google's search engine be required to allow the option of presenting diversity, or even to give priority to sites at variance with the profile of the seeker? Perhaps. Yet it's important to note that concern for the welfare of the seeker is secondary – satisfaction is to be fostered only insofar as it advances the goals of the company behind the software. Those goals are, of course, economic.[28] They are achieved by attracting customers who place advertisements. To neglect the seeker's interests completely would be self-defeating – for a search engine widely viewed as inferior would quickly lose its commercial clients. By the same token, however, advertisers typically want to find themselves aligned with particular perspectives and opinions, aiming to reach people with particular profiles, which is also why Google's strategy of personalization is so attractive to them. So there is an economic pull against any systematic attempt to achieve the diversity seekers sometimes want[29]. Locking people within a particular family of opinions and ideas is

---

[28] At least as long as we are not concerned with open source products where normally economic profit is not the driving force behind providing and developing software.

[29] Or, at least, citizens *should want* such a plurality of opinions when being engaged in democratic decision making, for example.

well suited to the paymasters' bottom lines, and thus to the commercial search engines they support.

Thus, although the internet – as a technology – allows all perspectives to speak, the economic arrangements that stand behind the relevant infrastructure, such as search engines, databases or social networks, permit some voices to muffle others or even to drown them out entirely. One major consequence is a version of the Cliffordian cascade and of Sunstein's isolated groups sharing radicalized opinions (reviewed in the previous section). Is there any chance of doing better? Our last discussion will look at one possibility.

Open Source Communities

We have argued that attempting a radical revision of the division of epistemic labor has obvious dangers. It may entrench unwarranted opinion, and amplify the economic forces that thwart epistemic endeavors. Although we have found a positive role for some internet sites and platforms in their provision of greater opportunities for information consumption (our thesis (3)), that role depends on a conservative extension of the prior division of epistemic labor, and, correspondingly, no significant inclusion of new voices (contrary to thesis (2)). Democracy in the epistemic domain thrives on a more active participation. One promising development seems to be the formation of *open source communities*. Paul de Laat defines them as groups of "peers producing content together on a voluntary basis[30], without direction from markets or managerial hierarchy, and posting their created content in

---

[30] Richard Stallman explains in "The GNU Project" what the initial motivation to engage in a non-commercial collaborative project was (see https://www.gnu.org/gnu/thegnuproject.en.html, accessed December 14 2014). Most of all, he points out that such projects shall enable competent people to develop software and tools that fit their needs and interests which is only possible if they have free access to the source code.

a virtual commons accessible to all" (de Laat 2010, 328).[31] Apparently, open source

communities expand the class of producers and the opportunities for consumption alike.

How have open source communities achieved their apparent success as complex

collaborative projects? De Laat analyzes trust-based cooperation in open source

communities (see de Laat 2010, 2012) by reconstructing the developmental history of such

communities.[32] A central problem, as we might expect from earlier discussions, is to handle

the anonymity of the participants (see de Laat 2010, 330). How were these communities able

to sustain their high performance rate, even when the number of participants started

growing rapidly?

De Laat assigns a prominent role to a shared attitude, a "hacker ethic".[33] He claims that this

attitude, a commitment to causing no harm while being engaged in hacker activities, was in

place even before the invention of the internet, held by some computer scientists who had

engaged in offline collaborative projects. The attitude was transferred to cyberspace,

becoming the basis of mutual trust for internet projects (see de Laat 2010, 331).

However, the new groups soon faced serious difficulties. More and more people wanted to

participate, and many of them were epistemically opaque. Thus, the shared background

attitude could no longer be presumed. Consequently, alternative ways to ensure reliability

were developed. Groups introduced rules for managing collaborative projects. Four main

features emerged: *modularization, formalization, division of roles*, and *decision making* (see

ibid., 333ff.). Groups decided to divide large projects into small subunits, to establish

---

[31] The most famous example in this context might be the open source software LINUX. The beginning of this huge project by Linus Torvalds is documented on https://groups.google.com/forum/#!topic/comp.os.minix/dlNtH7RRrGA[1-25], accessed December 14 2014.
[32] A first-hand description of the developmental process is offered by Stallman (see https://www.gnu.org/gnu/thegnuproject.en.html, accessed December 14 2014).
[33] On the notion of hacking see also Stallman (https://stallman.org/articles/on-hacking.html, accessed December 14 2014).

standardized tools and procedures for collaborative enterprises, to assign particular roles to

participants (such as 'read only' or 'edit'), and to introduce more hierarchical structures of

decision making by subdividing responsibilities. The new hierarchies often clearly moved

away from the initial democracy of the group. If, for example, a maintainer, i.e. the project's

leader, does not agree with a particular amendment, this new part of the code will not be

passed on. Furthermore, new participants have to prove their competence and reliability

with regard to the common project before being allowed to make substantial contributions

(see ibid., 335ff.).

Plainly, these mechanisms are aimed at addressing epistemic opacity. Traditional strategies

for identifying reliable informants – such as testing their qualifications and establishing a past

track-record of more or less successful contributions – have been used to screen those

allowed into the group discussion. Moreover, such endeavors are not subject to the critique

put forward by Frost-Arnold (2014), namely that "[...] abandoning anonymity can undermine

the sharing of true beliefs and the detection of false beliefs by decreasing the epistemic

activity of threatened epistemic agents. While those arguments were based on the premise

that such groups will participate less in non-anonymous epistemic communities, this fourth

argument maintains that their participation will, without anonymity, be given less epistemic

authority than it deserves" (ibid., 72). As it is *epistemic opacity* and not *strict* anonymity that

is the target of disclosure in the above suggestion, even otherwise oppressed users are still

able to raise their voice without being threatened. In the end, the development of open

source communities resembles the conservative extensions of the prior division of epistemic

labor we considered earlier. To achieve a genuinely beneficial expansion of contributors to

knowledge, one can only build on the social markers already in place. Although the store of

markers is expanded and the division of epistemic labor refashioned, our examples reveal no radical shifts toward inclusion of previously unheard voices.

Conclusion

We have attempted to review a variety of claims about the democratization generated by the internet, exploring them across a range of web platforms and contexts. It turns out that although the internet as a technology may allow more people to raise their voices and, in this sense, to gain the *opportunity* to have their opinions taken up by others, their chances of success in disseminating their views are limited. For the mechanisms giving access to these new claims to knowledge – the search engines, the social networks, the internet platforms – continue to sustain an authoritative framework. Filtering still goes on. Sometimes there are small positive steps toward democracy. Web services such as WikiLeaks might encourage people who otherwise would not dare to share their information to participate in knowledge distributing processes. By calling attention to this information such service providers might foster a more pluralistic information supply. But unless the new authorities remain close to the traditional assignments of expertise they risk creating a cascade of misinformation, locking people within previously held opinions, as the example of Google shows.

So far as we can see, if democratization is to be welcome, the inclusion of new contributors to knowledge must be relatively conservative. Perhaps the expansion of opportunities for knowledge consumption can be carried through more dramatically. As we have argued, however, the economic forces behind the internet might actually subvert democratic values. Without careful regulation, the internet cannot bring unmixed democratic blessings in the epistemic domain.

Acknowledgements

APPENDIX

WENDY DONIGER AND WIKIPEDIA: A CASE STUDY

Wendy Doniger, Mircea Eliade Distinguished Service Professor of the History of Religions at the University of Chicago, is internationally known for her studies of Hinduism and of mythology. Today, in 2016, she has a perfectly normal Wikipedia page, one that begins by listing some of her many achievements.

It was not ever thus. In 2009, Doniger's Wikipedia page looked very different. A "Background" section offered a very brief outline of her career. The bulk of the page consisted in a "Criticism" section, consisting of quotations from various commentaries. According to the passages cited, Doniger was "known for being crude, rude and very lewd in the hallowed portals of Sanskrit Academics". Her translations from Sanskrit were characterized as "unreliable and idiosyncratic." The entry placed Doniger in a virtual pillory.

People who read it got the message. One of them had a friend – someone we'll call M – who had long been interested in Hinduism and who had profited from reading some of Doniger's books. M recommended Doniger to others with similar interests, and was told that Doniger was "very controversial." The Wikipedia page was cited as the authority.

So, in 2009, M looked at that page. M also looked at the comparable pages for other plainly controversial figures – political leaders, for example. None of them showed the same lack of balance. Appalled by Wikipedia's complete silence about the many accomplishments that had earned Doniger an enviable international reputation, M decided to write to her about the page. Having recently – temporarily – retired after a successful business career, M volunteered to work on fixing the page. Doniger responded by explaining that others had already tried to do so, and that the venture was hopeless. Corrections were always quickly taken down, so that the page reverted to the hatchet job M had read.

In 2009, Doniger's work was the focus of a huge controversy. That year, Viking/Penguin had published *The Hindus: An Alternative History*, in which Doniger looked at Hinduism from the perspectives of outsiders, rather than presenting the orthodox vision of the Brahmin authors who had previously defined it. The book received highly positive reviews in many places, and was a finalist for a National Book Critics Circle award. In India, however, it was subjected to vitriolic denunciations, especially in the blogosphere, and Penguin/India eventually agreed to withdraw the book from circulation. (In fact, many copies continued to be sold, usually under plain wrappers, and the book was eventually republished in India by another company.)

Although M was aware of all this, the extreme mischaracterization on the Wikipedia page provoked a resolution to set things right. M spent a month researching Doniger's career, and

constructed a careful biography, deliberately neutral in its tone. The biography was sent to Doniger, who approved it – although she continued to believe in the impossibility of setting the record straight.

On August 4 2009, M entered the lists, writing under the pseudonym meetoohelp. Correction began with the insertion of a more complete list of Doniger's publications. That was followed with the biography, and some other amendments to the Wikipedia page, in particular a more balanced collection of comments from reviewers of her work.

M's last alteration was submitted at 2.54 p.m., 26 minutes after the first correction had been entered. Less than a minute later, another interested party, Priyanath, deleted everything M had submitted. Patiently, M reinserted the material, responding to the reason Priyanath had given for the bulk deletion. From 5.30 p.m. until early the next morning, the exchange is comical – Priyanath takes away, and M puts back. M is consistently courteous, showing a restraint rarely found on the internet.

By August 6, the page seemed to have stabilized. M's interventions had decisively altered the original tone, offering a more accurate picture of Doniger's work, the praise it had received and the criticisms leveled against it. Priyanath had apparently settled on making smaller changes. The page continued to be revised by Goethean, another Doniger champion, who attended to many details. Then in October 2009, Spdiffy launched a series of modifications, aimed at reverting to a negative picture of Doniger's work.

At this point M and Goethean joined forces to turn back the latest challenge. Plainly, however, the passions aroused by Doniger's work would be expressed in further attempts to undermine her credibility. M had decided to return to work, and would no longer have the time to monitor the page at the required frequency. Learning from Doniger that the

University of Chicago had a small fund available to enable its faculty to hire assistants for scholarly purposes, M proposed that Goethean be enlisted to keep an accurate page up to date; if the funds were inadequate to cover all the hours Goethean would have to devote to the task, M volunteered to make up the difference.

The arrangement has worked. Doniger's current Wikipedia page provides a reasonably accurate account of who she is and what she has done. But maintaining it has been an enormous struggle.

How does this episode bear on the issues about epistemic democracy that have been our central concern? First, it shows how democratic exchanges can lead to the construction of a reliable channel of information. Even without the invocation of official experts – hordes of Sanskritists and specialists in Hinduism testifying to some account of Doniger's achievements – the page eventually stabilized on something informative and accurate. But it also shows how chancy the process was. Without M's persistence and – unusual – politeness, Priyanath, Spdiffy, and others who followed them might easily have carried the day. Without Goethean's constant hard work, and, behind it, the support provided by the University of Chicago and (potentially) by M, Doniger's opponents could surely have chipped away around the edges (at very least). The success depends on highly unusual background conditions.

Even before the controversy about Doniger's page, Wikipedia had rules about writing "Biographies of Living Persons". They were an attempt to provide something like a constitutional protection against character assassination, by setting limits to what the "citizens" (the contributors) could do. As far as we can tell, the version of the page Priyanath, Spdiffy and their successors wanted to retain was in clear violation of the rules. It still took a

determined effort to produce a more balanced account of Doniger's intellectual contributions.

We see this case as supporting the main conclusions of our essay. The epistemic democratization offered by the internet is a conservative extension of the existing division of epistemic labor. But we see a thorough survey of examples like this one as a potential route to going further than we have done. A study of a variety of instances in which what is found at a particular site is shaped by protracted struggles among parties with incompatible views, might help improve the epistemic democracy the internet has so far provided.

References

APUZ Aus Politik und Zeitgeschichte "Digitale Demokratie" 62(7), Feb 2012.

Beam, Michael A. and Gerald M. Kosicki. 2014. Personalized news portals: Filtering systems and increased news exposure. *Journalism & Mass Communication Quarterly* 91(1): 59–77.

Clifford, William K. 1999. *The ethics of belief and other essays.* reprinted ed. Amherst NY: Prometheus Books.

Coady, David. 2012. *What to believe now: Applying epistemology to contemporary issues*. Malden and Oxford: Wiley-Blackwell.

Craig, Edward. 1990. *Knowledge and the state of nature – an essay in conceptual synthesis*. Oxford: Clarendon.

Daston, Lorraine and Peter Galison. 2007. *Objectivity*. New York: Zone Books.

de Laat, Paul B. 2010. How can contributors to open-source communities be trusted? On the

    assumption, inference, and substitution of trust. *Ethics and Information Technology*

    12(4): 327–341.

———. 2012. Open source production of encyclopedias: Editorial policies at the intersection

    of organizational and epistemological trust. *Social Epistemolog*y 26(1): 71–103.

Engler, Robert L., James W. Covell, Paul J. Friedman, Philip S. Kitcher, and Richard M. Peters.

    1987. Misrepresentation and responsibility in medical research. *New England Journal*

    of Medicine 317: 1383–1389.

Fallis, Don. 2011. Wikipistemology. In *Social epistemology: Essential readings*, eds. Alvin I.

    Goldman and Dennis Whitcomb, 297–313, Oxford: Oxford University Press.

Ferdinand, Peter. 2000. The internet, democracy and democratization. *Democratization* 7(1):

    1–17.

Feyerabend, Paul. 1978. *Science in a free society*. London: Verso.

Feyerabend, Paul. 1987. *Farewell to reason*. London: Verso.

Fleck, Ludwik. 1979 [1935]. *Genesis and development of a scientific fact*. Eds. Thaddeus J.

    Trenn, Robert K. Merton and Fred Bradley, Chicago: University of Chicago Press.

———. 1986 [1947]. To look, to see, to know. In *Cognition and fact*, eds. Robert S. Cohen.

    Boston Studies in the Philosophy of Science, 129–151. Dordrecht: Reidel.

Fricker, Miranda. 2010. *Epistemic injustice – power and the ethics of knowing*. reprinted ed.,

    Oxford: Oxford University Press.

Frost-Arnold, Karen. 2014. Trustworthiness and truth: The epistemic pitfalls of internet

    accountability. *Episteme* 11(1): 63–81.

Goodman, Nelson. 1953. *Fact, fiction, and forecast*. Indianapolis: Bobbs-Merrill.

Hanson, Norwood Russell. 1958. *Patterns of discovery.* Cambridge UK: Cambridge University Press.

Harris, Paul L. 2012. *Trusting what you're told: How children learn from others*. Cambridge and London: Belknap Press.

Jacob, Daniel and Manuel Thomas. 2014. Das Internet Als Heilsbringer Der Demokratie?. *APUZ Aus Politik und Zeitgeschichte* 64(22-23): 35–39.

Kitcher, Philip. 1993. *The advancement of science.* New York: Oxford University Press.

———. 2011. *Science in a democratic society.* Amherst NY: Prometheus Books.

Kneuer, Marianne (ed.). 2013. *Das Internet: Bereicherung oder Stressfaktor für die Demokratie?* Baden-Baden: Nomos.

Kuhn, Thomas S.   1962. *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Laudan, Larry. 1981. A confutation of convergent realism. *Philosophy of Science* 48: 19–49.

Longino, Helen E. 2001. *The fate of knowledge.* Princeton: Princeton University Press.

Magnus, P. D. 2009. On trusting Wikipedia. *Episteme* 6(1): 74–91.

Miller, Boaz and Isaac Record. 2013. Justified belief in a digital age: On the epistemic implications of secret internet technologies. *Episteme* 10(2): 117–134.

Mößner, Nicola. 2010. *Wissen aus dem Zeugnis anderer – der Sonderfall medialer Berichterstattung*. Paderborn: Mentis.

———. 2012. Die Realität wissenschaftlicher Bilder. In *Visualisierung und Erkenntnis. Bildverstehen und Bildverwenden in Natur- und Geisteswissenschaften*, eds. Dimitri Liebsch and Nicola Mößner, 96–112. Cologne: von Halem.

———. 2013. Photographic evidence and the problem of theory-ladenness. *Journal for General Philosophy of Science* 44: 111–125.

Munn, Nicholas John. 2012. The new political blogosphere. *Social Epistemology* 26(1): 55–70.

Neurath, Otto. 1991. *Gesammelte Bildpädagogische Schriften*. Ed. Rudolf Haller, Vienna: Hölder-Pichler-Tempsky.

Oreskes, Naomi and Erik Conway. 2008. *Merchants of doubt*. New York: Bloomsbury.

Pfister, Damien Smith. 2011. Networked expertise in the era of many-to-many communication: On Wikipedia and intervention. *Social Epistemology* 25(3): 217–231.

Sanger, Lawrence M. 2009. The fate of expertise after Wikipedia. *Episteme* 6(1): 52–73.

Sellars, Wilfrid. 1953. Empiricism and the Philosophy of Mind, In *Minnesota Studies in the Philosophy of Science*, Volume 1. Minneapolis: University of Minnesota Press.

Shapin, Steven and Simon Schaffer. 1985. *Leviathan and the air-pump*. Princeton: Princeton University Press.

Shirky, Clay. 2009. A speculative post on the idea of algorithmic authority. http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority/. Accessed 05 October 2016.

Simon, Judith. 2010. The entanglement of trust and knowledge on the web. *Ethics and Information Technology* 12(4): 343–355.

Simpson, Thomas W. 2012. Evaluating Google as an epistemic tool. *Metaphilosophy* 43(4): 426–445.

Sunstein, Cass R. 2007. *Republic.Com 2.0*. Princeton: Princeton University Press.

———. 2008. Democracy and the internet. In *Information technology and moral philosophy*, eds. Jeroen Hoven and John Weckert, 93–110. Cambridge: Cambridge University Press.

Thurman, Neil. 2011. Making 'the daily me': Technology, economics and habit in the mainstream assimilation of personalized news. *Journalism* 12(4): 395–415.