

## Why We Can't Have Our Facts Back

NOORTJE MARRES<sup>1</sup>  
UNIVERSITY OF WARWICK

### Abstract

How do we make the case for “knowledge democracy” in the face of the growing influence of right-wing figures and movements that denounce experts and expertise? While the threats to knowledge posed by these movements are real, it would be a mistake to return to a classic intellectual strategy—the politics of demarcation—in the face of this danger. Examining practical proposals for combatting fake news and opinion manipulation on the Internet, namely so-called “fact-checking” tools and services, I argue that they threaten to enroll us in a problematic normative project, one that aims to re-establish a hierarchy between knowledge and its presumed opposite, non-knowledge, or anti-knowledge. I list a number of shortcomings of this strategy. Most importantly, it distracts us from the role of technology in the crisis of public evidence in today’s computationally-intensive societies. Social media are a truth-less public sphere by design. A politics of demarcation also puts us at risk of forgetting a key insight from the previous century that remains valid today: knowledge democracy is a re-constructive practice and an ideal. Instead of consolidating hierarchies of knowledge through facts that derive their authority from outside the public sphere, we need to recover the central role in public life of experimental facts: statements whose truth value is unstable. The experimental validation of public knowledge must happen in the public domain.

### Keywords

demarcation; knowledge technology; knowledge democracy; digital society; public facts

### Introduction

“We want our facts back,” I heard an academic say, quasi-jokingly, at a recent conference about Science and its Publics, during a debate about “fake news” and “post-truth.” The remark seems indicative of a wider mood: since right-wing nationalist movements like Britain’s Brexiteers, the Netherlands’ Geert Wilders, and the alt right in the US have gained influence and power, many people feel the need to adopt clear normative positions, as in this case, in support of facts. There is also an element of nostalgia afloat. To want your facts back is to evoke a time when experts still seemed to have unquestionable authority and felt secure in this authority, when government

---

<sup>1</sup> Noortje Marres, Email: N.Marres@warwick.ac.uk

ministers didn't say things like "people have had enough of experts,"<sup>2</sup> and the notion that evidence should play a central role in public debate and politics had widespread, even taken-for-granted, institutional support. While I recognize its logic and appeal, I am also struck by the convenience of this position, and the creative interpretation of the recent past on which it relies. As I remember it, before Brexit, Wilders, and Trump—before disdain for facts was so openly expressed in public discourse as it is by these right-wing figures—those who argued that knowledge is central to public life were actually in the minority. Indeed, in the 1990s and 2000s, it was widely recognized by many that evidence-based debate was *not in itself a solution* to the problem of how to ensure respect for factual truth in democracies: the ideal of evidence-based debate itself was considered problematic insofar as it puts public life and politics at risk of overreliance on the authority of experts (Turner 2003).

There is no doubt that the status of facts in public life is under threat today. From newspapers proudly reporting un-truths ("Jeremy Corbyn was a Czech intelligence informant"),<sup>3</sup> to a UK government minister issuing inflammatory public statements based on unpublished research ("immigration put house prices up by something like 20%"),<sup>4</sup> public speech today appears more indifferent to facts than I can remember. The willingness of powerful figures to lie in public, and their indifference to being exposed for doing so, is only the tip of the iceberg. Today's changing architectures of the public sphere seem uniquely suited to the manipulation of people's opinions, as every week brings a new scandal about how social media, big data analytics, and digital "dark arts" were used to influence elections and referenda. This state of affairs may or may not be similar to other historical periods (the Cold War, the aftermath of the 1930s financial crisis). But there is also a normative question we should ask: in making the case that public discourse should respect factual truth, what ideal of public knowledge should we invoke? What role for facts in the public sphere should we strive towards? Numerous initiatives currently underway seek to reinstate respect for empirical evidence in public life, including awareness campaigns about disinformation online, knowledge literacy programs, and the development of so-called digital fact-checking services and tools. In this article, I want to argue that these efforts are at risk of re-instating an outdated strategy for securing the role of facts in public debate, one in which public respect for knowledge is based on authority. The strategy not only fails to address problems with knowledge in today's changing public spheres; it also reinstates a highly problematic normative hierarchy of knowing and un-knowing subjects, which ultimately stalls the quest for a thriving knowledge democracy. To make this argument concrete, I will focus my discussion on social media.

---

<sup>2</sup> As in the memorable statement by Michael Gove during an interview in the run up to the Brexit referendum. See Mance, H. (2016) "Britain has had enough of experts, says Gove." *Financial Times*, June 3, <https://www.ft.com/content/3be49734-29cb-11e6-83e4-abc22d5d108c>

<sup>3</sup> Ferguson, K., "Ex Czech spy Jan Sarkocy said Corbyn was a 'very good source' in TV interview." *Daily Mail*, 25 February 2018, <http://www.dailymail.co.uk/news/article-5432377/Czech-spy-says-Jeremy-Corbyn-good-source.html#ixzz5CdjUuM00>

<sup>4</sup> Shipman, T., "Tory housing minister Dominic Raab warns that immigration has pushed up house prices." *The Times*, April 8, 2018. Subsequently the "data" on which this claim was purportedly based was made public: a model of house prices produced by the National Housing and Planning Advice Unit in 2007.

Online platforms like Twitter and Facebook have emerged as key battlegrounds in conflicts about the role of facts in the public sphere. The companies behind them have responded to growing evidence of their role in the manipulation of public opinion by announcing strategic investments in technological fixes for disinformation online. A perfect example was Facebook's Mark Zuckerberg recent claim to the US Congress that software and AI are "going to be the scalable way to identify and root out most of this harmful content."<sup>3</sup> I would like to show here that the type of projects realized under this rubric threaten to set us back significantly, not only in relation to societal objectives of inclusion in public debate, but, equally, in relation to possibilities for a knowledge democracy. Realizing the latter ideal will require a different vision of the role of facts in public life, and it will require institutional and media-technological arrangements that *are significantly different* than the forms of "evidence-based debate" that had such seemingly widespread support before today's backlash against expertise. Trying to "get our facts back" is therefore the wrong strategy if the aim is to secure a central role for knowledge and inquiry in contemporary public life. What is the right one? I argue that we should not want our facts back, but instead develop new strategies to secure a central role for knowledge in public life, ones that address transformations of the very category of "public fact" in today's digital societies.

### **Knowledge vs. Anti-knowledge: The Return of "Demarcationism"**

A renewed commitment to "facts"—as what is needed to put public debate and politics back on sound footing—is today expressed in all corners—not just by academics, who, after all, have a professional investment in the virtue of knowledge, but equally by politicians, civil servants, activists, and tech companies. In response to recent and on-going scandals about the use of social media to manipulate public opinion during election and referenda campaigns in the US, Europe and Asian countries including the Philippines and Myanmar, companies like Facebook and Google have taken steps to implement solutions. Prominently, these include so-called fact-checking tools and services. Using a variety of techniques, from "human moderation" of online content to automated fact verification, these tools and services apply mostly journalistic procedures for the verification of statements to published content in online environments (Graves 2018). Over the last years, online platforms have invested in moderation teams, which sift through large amounts of online content to detect and remove harmful and duplicitous material. They support user literacy campaigns, as in the case of the UK Charity Full Fact funded by the Google News Initiative among others (more about which below). These efforts also include the development of automated tools for content verification, where software is used to identify and

---

<sup>3</sup> Harwell, D., "AI will solve Facebook's most vexing problems, Mark Zuckerberg says. Just don't ask when or how." Washington Post, April 11, 2018. Since this hearing, harmful online content seems increasingly defined in public debate in moral terms, with "hate speech" as a prominent placeholder term. This moral coding of "fake news" is interesting as the role played by social science research in online opinion manipulation received much public attention around the same time. It was a research app designed by Cambridge University psychologists that enabled the collection of the Facebook data used to influence the US Presidential election of 2016. Hern, A. and O. Solon, " Facebook closes 538 million fake accounts in the first three months of 2018." The Guardian, May 15, 2018.

flag dubious statements online. A case in point is the Facebook "disputed content" banner (see Figure 1).<sup>6</sup> These interventions, furthermore, have been well publicized in such prominent media as the New York Times and the Financial Times, and they appear to do double duty as badges of corporate responsibility for "big tech."

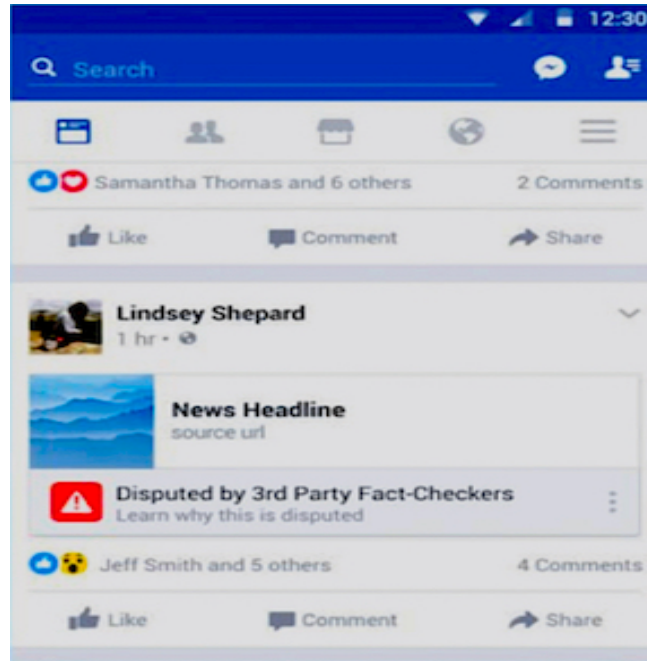


Figure 1: Facebook's disputed content banner (New York Times, 23 October 2017)

This press coverage has by no means all been positive, however. Gillian Tett has written about the bad working conditions in the click farms where the epistemic labor of fact checking is performed in today's digital societies.<sup>7</sup> Yet receiving less attention to date are the knowledge models and methodologies that inform these fact-checking services. Arguably, however, ideals of public knowledge, scientific citizenship, and knowledge democracy, are partly what lends these high-tech pro-fact campaigns their credibility. What is remarkable about these projects is their endorsement of an idea that only a few years ago was mostly considered outdated, namely, that quality is an attribute of information itself. In the early 2010s, it was claimed that in our digital

<sup>6</sup> Many of these projects build on earlier initiatives. Intel, for example, launched a prototype "Disputefinder" tool in 2010 (for a discussion see Marres 2015).

<sup>7</sup> Tett, G. (2018) "The low-paid workers cleaning up the worst horrors of the internet." Financial Times Magazine, 16 March 2018 <https://www.ft.com/content/b1042958-2712-11e8-b27e-cc62a39d57a0>. Scholars like Lilly Irani (Irani et al. 2013) have done extensive research on the new, often exploitative forms of epistemic labor enabled by digital platforms.

age, information had become “social,” in that what mattered most in this age was not content in itself, but rather how it circulated, how messages spread and gained influence by engaging people and communities (Gerlitz and Helmond 2013; Nascimento and Polvora 2013). In contrast, fact-checking efforts shift the focus back to the properties of content itself (Bounegru et al. 2017). *What is said to be in need of attention and intervention is the “veracity” of online statements and the potential duplicity of online sources.*

To be sure, this change of tune can partly be accounted for by the constant swings in the hype-prone field of digital innovation. Nevertheless, the implementation of this core idea has significant implications for the problem definitions, proposed solutions, and allocation of responsibilities now underway. The example of the start-up Factmata can help to make this clear. This company is developing new software tools that will “help people to become their own fact-checkers—making it easy to attain factual data about news stories you come across.” In February 2018, the London-based company received £1 million in seed funding from a Twitter co-founder and others to develop a prototype that uses AI to provide Internet users with “factual context in real time.”

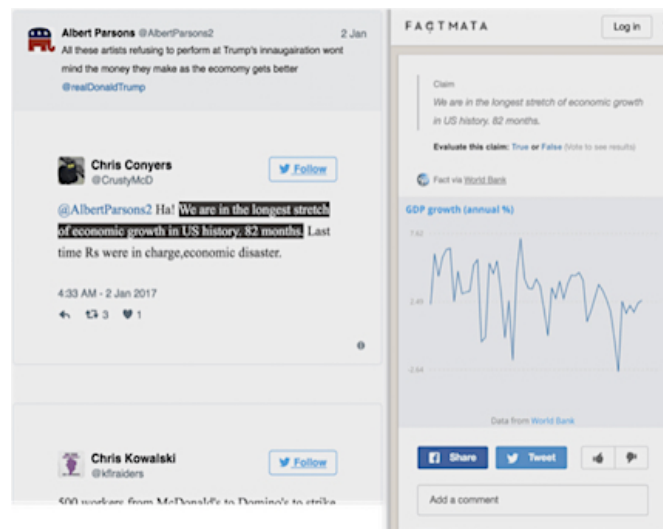


Figure 2: An early mock-up of the Factmata system (The Next Web, 27 June 2017)

Factmata's objective is to cross-reference any claim circulating online onto a database of so-called verified statements, in order to validate or invalidate it. As Dhruv Ghulati, the start-up's CEO, put it: “An example of how this would work is that when you'd open an article or a

• Lunden, I. (2018) “[Factmata closes \\$1M seed round as it seeks to build an ‘anti fake news’ media platform.](https://techcrunch.com/2018/02/01/factmata-closes-1m-seed-round-as-it-seeks-to-build-an-anti-fake-news-media-platform/)” TechCrunch, 1 February 2018  
<https://techcrunch.com/2018/02/01/factmata-closes-1m-seed-round-as-it-seeks-to-build-an-anti-fake-news-media-platform/>

statement about Muslim immigration, you'd suddenly see the actual data about the issue. This would come in the form of charts, tables and so on. That layer is a layer of data on top of information" (Figure 2).<sup>5</sup> In proposing to implement this methodology of verifying online statements by matching them to a database of validated facts, Factmata reproduces a well-known philosophical approach: the correspondence model of truth.

This approach is most famously associated with early 20th Century Viennese logicians and philosophers of science, such as Rudolf Carnap, who transformed a classic philosophical question—"How can we distinguish between legitimate and illegitimate knowledge claims?"—into an implementable procedure. The logical empiricists designed and advocated generally applicable procedures of statement verification that make it possible to decompose expressions into more basic observational statements, which could then be matched against experience or against other verifiable statements (Carnap 1995). For example, "the feather hits the floor at  $t=42$ " could serve to ground a more complex statement about objects falling with uniform acceleration. In some respects, it should not surprise us that this scientific ideal is today "channeled" by solutions proposed by tech companies like Factmata. The logical positivists built on the work of logicians like Gottlob Frege, which subsequently informed the development of computer science and artificial intelligence. Factmata is promoted as an implementation of "artificial intelligence" technology, which will be used to automate the detection of duplicitous statements. On a more general level, its promise to implement a complex epistemological procedure—which statements qualify as fact?—accounts for Factmata's attractiveness, just as it did for logical empiricism, except that in Factmata's case it also explains its valuation as a start-up.

It is not strange, then, that the valuation of artificial intelligence as the current next big thing in "tech" is accompanied by the return to and re-valuation of a logical-empiricist conception of knowledge. But something else also follows: while fact-checking tools are designed to implement a specific, empirical methodology and offer a particular service—namely to automatically verify the truth value of online expressions—these devices simultaneously reactivate a fundamental conceptual distinction, that between legitimate and illegitimate claims to knowledge. This should set our alarm bells ringing. Notwithstanding its framing as a mere "technical procedure"—a methodical operation upon statements—demarcation draws a normative boundary, and I want to argue here, a deeply problematic one.

The technical distinction between valid and invalid statements invokes the broader normative opposition between legitimate and illegitimate claims, between knowledge and non-knowledge. Philosophers such as Richard Rorty (2009 (1979)) refer to this conceptual apparatus as "demarcationism," as it defines validity and invalidity as binary attributes of individual statements: can claims be traced back to observational sentences, or can they not? It also imposes as a normative objective the establishment, and policing, of the boundary between these two types of statements. Furthermore, demarcationism enables the allocation of responsibility to actors for the epistemic quality of debate. Demarcation secures respect for factual knowledge

---

<sup>5</sup> Maack, M. (2017) "Fake news is your fault, this start up is trying to help." The Next Web, June 27 <https://thenextweb.com/artificial-intelligence/2017/06/27/users-will-have-to-take-responsibility-for-fake-news/>

through the validation and de-validation of statements, and these selective operations upon content produce a normative distinction between "good" sources capable of producing valid statements and those that don't, between "good" users capable of discerning valid statements and those who aren't. I refer to this normative operation as the politics of demarcation.

In line with this approach, fact-checking arrangements promoted by tech industries undertake this same allocation of responsibility:<sup>20</sup> it is *the type of information sources* that the Internet makes available, on one hand, and *the users' lack of skills capable of discerning the difference between valid and invalid statements, one the other*, that are said to be responsible for the prevalence of dodgy content in this media environment. As an online article about Factmata put it only quasi-jokingly: "Fake news is your fault, this start-up is trying to help."<sup>21</sup> The aforementioned Full Fact on its Web site provides practical guidelines for social media users to enable them to detect "fake news." In telling users what to look out for, Full Fact provides a list of tell-tale signs of disinformation, which notably includes "look closely at the URL" and "investigate the source." Fact checking here comes down to training users to ascertain the veracity of individual sources. A report on Full Fact's activities in the Guardian newspaper helps to make clear what kind of normative judgments this approach may lead to: information sources that don't comply are at risk of being singled out as deviant. Characterizing Full Fact as a service that eliminates fake news by tracing stories back to their source, the Guardian cites as an example the false claim "Labour leader Jeremy Corbyn hints he would pay £85bn Brexit 'divorce bill' if PM."<sup>22</sup> According to the newspaper, this claim originates with the following source: "...an Essex teacher who until now has tried to guard his anonymity [...] has confirmed to the Guardian that he owns the domain YourBrexit, but said Walter White was the pseudonym of a 19-year-old student in Southend 'who is fed up with the liberal elite.'" But, should we really believe that actors like "a 19-year old pushing fake news from his basement in Essex"<sup>23</sup> are responsible for fake news? Or do suggestions like these—inadvertently or not—apportion blame in a problematic way, shifting responsibility to the familiar suspect of undisciplined youth from out of town? Fact checking, and

---

<sup>20</sup> Demarcationist approaches to the Internet are not new: the idea that digital media lack mechanisms for separating the "wheat from the chaff" was prominent in the 1990s also. Similar arguments about the deluge of un-validated content that circulates unchecked, and users who lack the ability to discern knowledge from non-knowledge, held sway in the late 1990s. Demarcationist technologies have been proposed and tried over many years, not only Intel's Disputefinder (2010) but also earlier "fact versus fiction" pages which became a prominent feature of the Web in the late 1990s. However, whereas in previous decades it was the "lone crackpot" or "fringe sub-cultures" that were identified as the stereotypical producers of chaff (anti-vaccination, flat earthers), with the mainstreaming of the Internet this category seems increasingly marked in social terms invoking deviance, notably, with education levels as a prominent marker.

<sup>21</sup> Maack, M. (2017) "Fake news is your fault, this start up is trying to help." The Next Web, June 27 <https://thenextweb.com/artificial-intelligence/2017/06/27/users-will-have-to-take-responsibility-for-fake-news/>

<sup>22</sup> Booth, R. (2017) "Truth seekers: inside the UK election's fake news war room." The Guardian, May 19 <https://www.theguardian.com/politics/2017/may/19/truth-seekers-inside-the-uk-elections-fake-news-war-room>

<sup>23</sup> Booth, R. (2017) Truth seekers: inside the UK election's fake news war room, The Guardian, May 19 <https://www.theguardian.com/politics/2017/may/19/truth-seekers-inside-the-uk-elections-fake-news-war-room>

the wider principle of demarcation, risks reinforcing stereotypical oppositions between those who are trained to be discerning and those who are not, between those capable of knowledge and the others. Worryingly, the distinction between knowledge and non-knowledge roughly maps onto a polarity that defines today's political force fields, namely the opposition between educated progressives and, on balance, less educated supporters of populist and nationalist causes, a point to which I will return in the conclusion.

### **Crises Behind the Crisis: Commercial Media, Polarization, Software that Cheats**

The Science & Technology Studies scholar Sergio Sismondo (2017) stated that "Twitter may be part of the dissolution of the modern fact," citing as evidence that celebrities and trolls dominate online discourse.<sup>14</sup> However, the opposite may equally be the case: Facebook and other tech giants are today bent on re-instating the modern fact in digital media environments by implementing fact checking and animating the demarcationist agenda behind it. Although the threats these corporations are combatting are real, this is an inadequate response. There are a number of reasons for this, of which I will discuss four.

The first is that, as a solution, demarcationism does not directly address the role that digital media technology itself plays in undermining respect for knowledge in public discourse. There is widespread evidence that the algorithms that regulate access to online media sources promote duplicitous content. Digital platforms rely on automated methods of source selection—such as search engine return lists—to determine what content reaches their users. All of the major platforms have been shown to favor sensational, often outrageous and extremist, material—of the sort that states, "Barack Obama may have planned a coup d'etat"—over well-balanced, factual information (Tufekci 2015).<sup>15</sup> As a recent study of the spread of messages on Twitter put it provocatively: it took the "truth" about six times as long as "falsehood" to reach 1,500 people on Twitter (Vosoughi et al. 2018).

Despite the secrecy with which tech companies protect their valuable algorithms, we have a pretty good idea of why duplicitous content "outperforms" verified statements online, to use BuzzFeed's telling verb.<sup>16</sup> Algorithms for online content selection are designed to maximize

---

<sup>14</sup> In support of his claim, Sismondo (2017) quotes Guardian commentator Lindy West, who upon announcing that "she is abandoning Twitter" wrote that "it may simply be impossible to make this platform usable for anyone but trolls, robots and dictators (West 2017)."

<sup>15</sup> The Facebook News Feed that serves topical articles directly to social media users; the Youtube video recommendation systems that suggests similar titles ("people who watched X also watched Y") and the Google Snippet tool that presents fragments from sources in its return lists: they have all been demonstrated to favor sensational and extremist content. Fischer, M. and A. Taub (2018) "How Everyday Social Media Users Become Real-World Extremists." *New York Times*, April 25. Condliff, J. (2017) "Google has been totally obtuse about these issues." *MIT Technology Review*, March 6 <https://www.technologyreview.com/s/603796/googles-algorithms-may-feed-you-fake-news-and-opinion-as-fact/> Morris, D. (2018) "How YouTube Pushes Viewers Towards Extremism." *Fortune*, March 11, <http://fortune.com/2018/03/11/youtube-extreme-content/>.

<sup>16</sup> Silverman, C. (2016) "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook." *Buzzfeed*, November 16



circulation—or what social media commentators stubbornly persist in calling "engagement"—as it makes sense from an advertising point of view to privilege messages that reach many users in a short time, thereby enabling them to reach yet more users. Online platforms, then, reward messages that spread instantly and widely with even more visibility, and, as tabloid newspapers invested in maximizing advertising revenue also found out in previous decades, sensational rather than factual content turns out to satisfy this criterion of maximal "share-ability" best.<sup>27</sup> A commercial logic here gives rise to a circular content economy, one without referent: content that gets shared a lot is rewarded with more visibility, thereby increasing its share-ability. The methods and tools of fact checking that are implemented in services like Facebook's Disputed Source or Factmata leave untouched these algorithmic selection processes. Since those fact-checking techniques operate according to a correspondence model, by checking individual statements against a database, the *circulation* of claims as such is left out of account. To put this differently, fact-checking services are designed to militate against "source bias." They therefore leave unaddressed the role of algorithmic selection in enabling the circulation of particular types of content and the associated production of "network bias." This methodological blind spot enables the more general, moral-political blind spot to which I alluded above: *fact checking, and the wider principle of demarcation, encourages us to allocate responsibility for the quality of online information to individual sources ("content providers") and addressees ("users"), and it thereby displaces critical attention away from the technologies of source selection that regulate content circulation online.* However, given the role of algorithms in the circulation and promotion of dubious content, surely we do need to lay some blame on the technology for the quality of expression online.

While academics and journalists have done important work to expose the role of algorithms on the production of online bias (Ziewitz 2016; Kroll et al. 2016), it is important that we also examine the formation of knowledge communities online. As Bounegru, Gray, Venturini and colleagues (2017) have argued, it is not just the principles implemented in social media algorithms, but the particular ways in which messages circulate between machines and people that are behind the disregard for truth in online media spaces today. Analyzing the social networks through which specific fake news articles—"Trump offering free one-way tickets to Africa and Mexico for those who wanna leave America"—spread on Facebook, Bounegru et al. show that the communities that share this type of content barely overlap with communities invested in debunking these false claims (Figures 3a, 3b). For these particular claims, Facebook users invested in fact verification— quite possibly including those who use tools like the Dispute

---

<https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>

<sup>27</sup> Criticisms of the social media logic of circulation that maximizes share-ability have been made on various grounds. The study by Vosoughi et al. (2018) that compares the share-ability of true versus false information does this according to a prescriptive, epistemological model. Elsewhere I have criticized the socio-logical principle of "trend formation" promoted by social media platforms, on the moral-political ground that it risks producing a "tyranny of the majority," and I contrast this with the distributed, socio-ontological logics of issue formation (Marres 2017, p. 71-72). Here, below, however, I focus on the ways in which the design of social media locates knowledge and truth outside the space of public debate (see on this point also Kennedy and Moss 2015).

finder—and those invested in the propagation of fake news form different communities in social media. This finding is important because it suggests that users of fact-checking tools may present only a sub-set of all the users implicated in the circulation of fake news.<sup>18</sup> It also highlights a basic but valuable point on which the sociology of knowledge has long insisted: that different forms of knowledge find their basis in different communities (see also Angemuller 2018). Equally important is the *type* of knowledge communities that have formed in the wake of fake news on Facebook: here, fake news-propagating communities are pitched against fact-based communities. This finding, that some epistemic—and anti-epistemic?—communities "barely overlap" on Facebook, highlights the salience of the opposition between "knowledge" and "anti-knowledge" as a principle of social organization, or rather, of social division and polarization. And this brings us to a second shortcoming of demarcationist approaches to safeguarding knowledge online: *they risk reinforcing the societal dichotomies mobilized and propagated by anti-truthers, namely, the opposition between evidence-loving progressives and lie-condoning conservatives.*

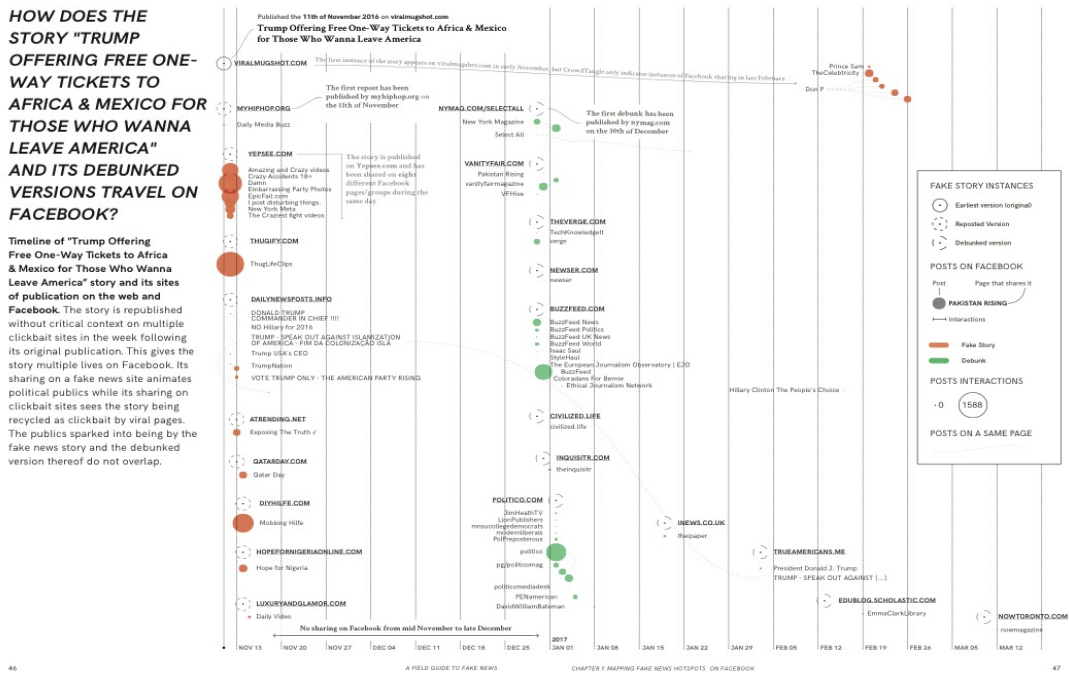


Figure 3a: Are debunking Web pages acknowledged by the publics of fake news? (Bounegru et al. 2017)

<sup>18</sup> <https://fakenews.publicdatalab.org/download/SAMPLE-field-guide-to-fake-news.pdf>

ARE DEBUNKING WEB PAGES ACKNOWLEDGED BY THE PUBLICS OF FAKE NEWS?

Fake news pages and debunking web pages have different publics on Facebook. Only six of the public pages that share fake news stories have acknowledged web pages which aim to debunk them in our CrowdTangle dataset. While Google looks to prioritise debunking web pages, on Facebook it is fake news stories that circulate better. While both progressive and conservative pages share fake news stories it is primarily progressive Facebook pages and those pertaining to journalists and fact-checking initiatives that share web pages which aim to debunk fake news stories.

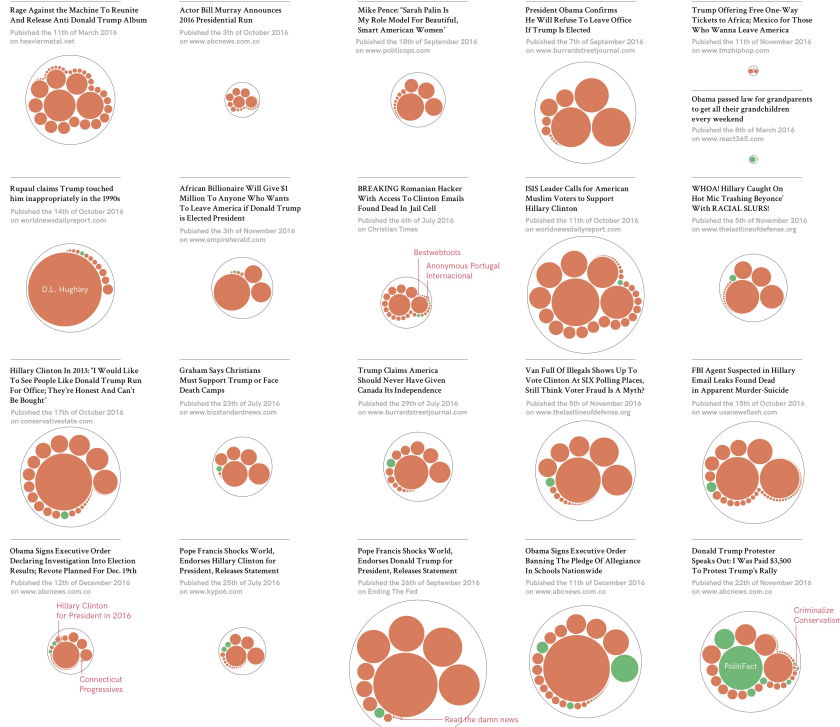
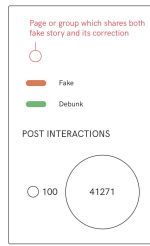


Figure 3b: Are debunking Web pages acknowledged by the publics of fake news? (Bounegru et al. 2017)

The Internet's problem with knowledge, then, is not just a problem of toxic content, illiterate users, or badly-behaved algorithms pushing said toxic content into the mainstream, but has to do with the type of knowledge communities enabled by digital media. The key question concerns how we characterize these communities: do we subscribe to a normative opposition between "literate" and "illiterate," (epistemically capable and epistemically incapable) sources and users? Or do we understand fake news communities reflexively, recognizing that part of their point may well be to challenge this quite possibly self-indulgent normative hierarchy of knowing and "un-knowing" subjects? The point is, an approach focused on the attributes of "content itself" does not enable us to address let alone answer these questions. However, insight into this societal dimension of online discourse should certainly not lead us to shift our attention away from knowledge. This points to a third shortcoming with fact checking and demarcation: *it does not enable us to understand the problem with knowledge on the Internet, or indeed, beyond it.* To be critical of fact checking, as a solution, is not to deny that there is a crisis of public evidence in today's digital societies, to the contrary. But this crisis goes well beyond a disregard for facts in digital media environments. Not just in the domain of news media and public politics, but in a

variety of societal spheres, computational tools, services and architectures are today enabling interventions in society that undermine the very status of "facts" as the empirical basis of public accountability. A case in point is Dieselgate, the scandal surrounding the Volkswagen cars that have been computationally enabled to "cheat" during emissions tests. The software running on these cars include a so-called defeat device which made it possible for these cars to "perform to the test": when the software detects that cars are being driven in accordance with "test conditions," it adjusts the cars' behavior, keeping the exhaust of NOX and other gases low. When they are driven on the street, the software lets harmful emissions go up to levels that are dramatically higher than when in the lab (for a discussion see Marres 2018; Lippert 2016). In this way, smart software enables the car industry to hold up a deceptive appearance of empirical compliance with regulations.

This is not the place to discuss the computational manipulation of test-based regimes of governance in more detail, but it highlights a crucial point: not just in the domain of social media, but in other sectors and environments too, *the insertion of computational technologies into public infrastructures have resulted in deception and manipulation of the empirical record*. Digitization is challenging not just the status of facts in public discourse, it is undermining institutional architectures that ground accountability in empirical evidence across societal domains.\* (This is another reason why Sergio Sismondo's statement that we are observing "the dissolution of the modern fact" in digital environments is apt.) To be sure, the grand questions of which type of evidentiary regimes are appropriate for a digital age, and how the empirical basis for accountability can be re-instated in computationally-intensive societies, remains to be answered. But it seems highly unlikely that the correspondence model of factual truth—the one on which fact checking is based—will be adequate to addressing the new conditions. According to the principle of correspondence, empirical statements are validated by checking them against pre-given, stable referents and such an approach is ill-adjusted to the dynamic model of information, communication and feed-back that is today implemented across society by computational means. The correspondence model of truth is by definition retrospective, or "backward looking," and in many cases this is not adequate in today's dynamic information environments, in which readings and behaviors are constantly adjusted as conditions change. As I will discuss in the section on experimental facts below, we need to replace a retrospective with a dynamic conception of the role of empirical truth in public debate. While campaigns to implement fact checking in public discourse today concentrate on social media, they mis-define not just the problems with knowledge on the Internet, but also those of wider digital societies.

---

\* As Doctorow puts it: "Software can say, 'If there's a chance I'm undergoing inspection, then I will be totally honest'—but cheat the rest of the time." Except it is the other way around: software cheats when tested. But the broader point holds: machinic agency now extends into the epistemic, and this challenges empirical forms of governance and accountability (Doctorow 2017).

### **Behavioral Media: A Truthless Public Sphere by Design**

Lack of respect for facts in online discourse, then, has deep and complex causes, which go well beyond the qualities of information sources and users on the Internet, and instead have to do with the very designs of digital infrastructures and how these operate on conditions for public knowledge. Here I want to discuss one further aspect of digital media platforms that is relevant to this (and which will bring us to the fourth and last shortcoming of the politics of demarcation). This is the behavioral vision that has informed the design of social media architectures and encourages a conception of users as influencable subjects, not knowledge agents. The designers of social media platforms like Facebook have drawn inspiration from a variety of sources, including the sociological analysis of networks (Healy 2015), but they have equally adopted models of communication from biology and physics. What the data scientist Alex Pentlandt calls “social physics” proposes that the mobilization of people and the formation of opinion can be understood using the same principles as viral contagion or the flocking of birds (for a discussion see Marres 2018). Crucially, computational social scientists prize online platforms for enabling a science of society that does *without interpretation*: on Facebook and Twitter, the argument goes, social and political processes can be measured by tracing action—by what is shared, linked, clicked, purchased. As one group of data scientist put it: “we do not have to consider people’s opinions at all” (Newman et al. 2007), considering it progress that in digital media public opinion formation does not need to be defined as an interpretative process. In this sense, one could say that social media enable a truth-less public sphere by design.

However, we *cannot* infer from this that the social media imaginary is a “knowledge-free” environment. It’s not just that plenty of knowledge claims are put forward here. Instead, social media present a research-centric apparatus, in that their design directly reflects the epistemic needs of the data scientists whose analytic operations are key to their commercial model: to target information to groups of friends, to track shares and likes in the aggregate. This analytic architecture is shot through by behavioral assumptions: the activities that platforms enable—to influence, to make trend—have the “manipulability” of users as their primary feature. Take Facebook’s “lookalike” audience tool that makes it possible for advertisers to match particular topics—say a brand or politician—with a groups of users that “like it” and then to target messages at identified user segments. As one Facebook executive put it, this type of “demographic targeting” is able to find just “the right admixture of age, geography, time of day, and music or film tastes that demarcate a demographic winner of an audience.”<sup>20</sup> Such a tool makes the *influence-ability of users* the defining feature and the condition of success for social media communication. It is also to say, the use of social media for the purpose of “opinion manipulation” was not simply an invention of rogue campaigners, or of the “populist right,” but also, to an extent, the realization of a *scientific* vision. Until the election scandals of 2016 and 2017, the ability to influence voters and voting behavior was highlighted by data by scientists as a key

---

<sup>20</sup> A tool like Lookalike Audience makes the influence-ability of its users the defining feature and the condition of success for social media communication. See Garcia-Martinez (2017) “I’m an ex-Facebook exec: don’t believe what they tell you about ads.” The Guardian, 2 May  
<https://www.theguardian.com/technology/2017/may/02/facebook-executive-advertising-data-comment>.

strength of social media platforms. For example, the oft-cited research by Bond et al. (2012), which took the form of a “61-million person experiment in social influence and political mobilization,” demonstrated that “social messaging” via Facebook—showing users which of their friends had clicked the “I voted button”—directly influences voter turnout. The article concludes with an “upbeat” vision, which proposes that such techniques can be used to make “real-world behaviors amendable to online intervention.”

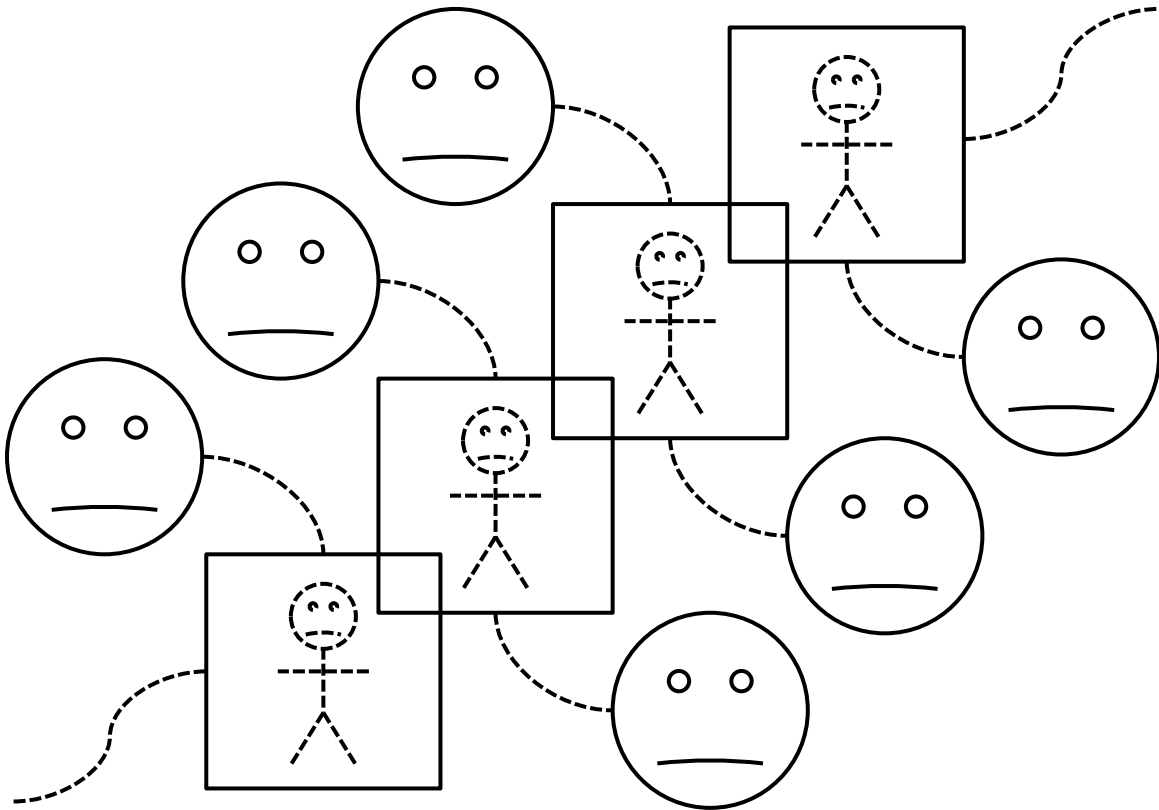


Figure 4: The behavioral Shibboleth (Muniesa 2018)

My point here is certainly not to challenge the legitimacy of behavioral science as such, but, rather, to argue that the disrespect for knowledge in social media has something to do with the dominant conception and design of social media platforms as a behavioral domain, where users are often framed as influence-able subjects.<sup>23</sup> This behavioral methodology has been aptly

<sup>23</sup> Several aspects of the social process of opinion formation are ignored by such an approach, such as the fact that convictions aren't simply attributes of actors but accomplishments of communicative processes. “Online content,” or “social media networks” are not just reflective of people's “opinions” or “attitudes” or the

schematized in a diagram by Fabian Muniesa (see Figure 4, reproduced from Muniesa 2018) entitled the “dotted agent.” It visualizes the relation of influence-ability, and the steep knowledge differential, between active, knowing analysts (data analysts) and the passive, known subjects of behavioral analytics (users). This, then, is a further and last reason why demarcationism does not enable us to adequately grasp the Internet's problems with knowledge: the normative opposition between knowledge (good) and non-knowledge (bad) that it imposes makes it difficult to see that epistemic ideals—like behaviorism—themselves have played a role in encouraging a disregard for knowledge on the Internet. Social media are not simply marked by a growing public disrespect for knowledge, they signal the prevalence of a distinctive type of research regime. Much like Muniesa's dotted agents diagram, *social media present an environment in two halves, where, on the one side, we find users with “influence-able” and “target-able” opinions, tastes, and preferences, while, on the other side, we have authoritative data analysts who “know” the population's fine-grained and ever-changeable preferences and tastes.* Scientists—the proponents of knowledge—haven't been by-standers but active participants in the crafting of a media architecture designed to enable the influencing of users' actions. <sup>2</sup>

If we understand demarcationism on a general level as a normative principle predicated on the separation of the sphere of knowledge (legitimate truth claims) and the sphere of opinion (expression without a claim to knowledge) then today's digital media architectures are already the embodiment of this principle: on the one hand, we have a domain of analysts who are supposedly equipped to know online populations, and, on the other hand, we have a space inhabited by influence-able subjects whose attributes—at least by design—include preference and opinion but not knowledge. Seen from this angle, demarcationism is not the solution, it is part of the problem. This is why Sismondo's declaration—social media present the dissolution of the modern fact—requires a corollary: while it may be true that the modern fact is being dismantled in online discourse, facts of a kind continue to do very well in the wider digital economy, culture and politics. The online domain of swayable opinion is only half of the story of social media platforms. Users are exposed to influence and manipulation, and as such they may seem unlikely to be capable of more than opinion. But social media have equally given rise to a veritable army of social data scientists who monitor, measure, and seek to intervene in this behavioral theatre. Facts continue to reign but they do so outside the public realm: we are observing not the wholesale dismantling of the authority of facts—experts continue to exert considerable authority in institutional and commercial settings—but the far more specific dismantling of public facts, of the Internet as a space where knowledge can be practiced publicly, among experts AND citizens.

---

make-up of social groups. Instead, public opinion *emerges out of interactions between* people, technologies, and ideas. In my view then, these methods and media are not quite social.

<sup>2</sup> As ESTS editor Daniel Lee Kleinman observed: it may also be taken to suggest that the rise of some experts (e.g. data scientists) is accompanied by the demise of others (e.g. public climate change scientists).

### **In Praise of Experimental Facts**

The opposition between epistemic and anti-epistemic communities, between those who respect the facts and those that spread disinformation, hides important ambiguities. Just as the problem of harmful content and illiterate users isn't the whole story, so we need to understand that there is much more to today's disregard for facts in public discourse than a dislike of expertise and experts. An indication of this can be found in an online debate map that the fact-checking team at Full Fact produced in the period following the Brexit referendum. The map was intended to help human fact-checkers to identify the thematic areas where fake news is particularly prone to arise, but the resulting "issue scape" suggests *there is logic behind the apparent confusion of circulating untruths*: fake news tends to attach itself to serious issues such as living standards, immigration, public services, and the NHS. Reading the map, I was struck by just how faithfully fake news apparently follows the public debates and controversies that define Austerity Britain, where the UK government's objective to cut spending has over the last 8 years dramatically and contentiously reduced funding across social policy areas from schooling to mental health, disabilities, youth, and prisons. Indeed, precisely these issues were deployed, albeit instrumentally, in the run up to the Brexit referendum, to pitch "Brexiters" against "Remoaners." But this also raises a question: if disinformation connects to matters of public concern, how well do current efforts to combat disinformation address this connection? It seems to me the answer is that they do it badly.

Fact-checking tools and services divide public information into a legitimate and an illegitimate corpus of claims, and, in so doing, they are all too likely to consolidate or even amplify public divisions between virtuous knowledge and sinful non-knowledge. That does not seem to me helpful at all if the aim is to advance mutual understanding among concerned publics: while fact checking prides itself on its empiricism, the ways in which it divides the public confirm or even aggravate prejudice. This arguably was the case in one-to-many broadcast media too, but the effect is amplified in participatory media where content providers in principle include the public, rendering the politics of demarcation much more problematic. We need alternative, less judgmental ways of carving up public issue spaces.<sup>23</sup> One way to do this would be not simply to check the factual statements put out there by individual sources, but to test public media for "experimental facts": to examine the robustness of a different type of empirical claim put forward in our changing media environments today, namely propositions whose truth value may fluctuate over time, and of which the political color may also change.

To be sure, the stable facts on which fact checking is mostly focused have their place in public discourse. Statements that are solidly supported by "ground truths" are important ingredients of public debate. They present the background conditions that constrain or should constrain opinion formation. But different and equally crucial elements in public debates are experimental facts, or those statements whose veracity is unstable and whose epistemic status changes over time. Take as an example the infamous claim that "Brexit will make available 350

---

<sup>23</sup> What we do not need is an anti-social epistemology to combat an anti-epistemic conception of sociality and repeat this back-and-forth endlessly with no methodological innovation on either end.



million pounds per week for the NHS." When this statement was first plastered onto a red bus during the referendum campaign, it qualified as a robust example of "post-truth politics." According to the UK Office of Statistics, the statement presented a "clear misuse of statistics."<sup>24</sup> However, in later speeches, debates, and analysis, the claim came to signal something else as well: it highlighted the role of cuts and austerity in swinging the Brexit vote. The statement, as time passed, remained highly problematic but subsequent operations upon it arguably made this claim less un-true. As Simon Stevens, Head of NHS England stated: The "health service 'wasn't on the ballot paper, but it was on the battle bus."<sup>25</sup> The popularity of the Brexit bus, that is, came to serve as evidence for the widespread public concern with the state of the NHS. The rise to prominence of the 350 million NHS claim, then, is not just indicative of a public "literacy" problem. As the NHS England Head's statement indicates, public interest in this claim that materialized as a bus on the country's roads may even indicate a valid attempt at issue articulation: something must change in this country so that public investment can return to vital services across regions. That this conviction was put to opportunistic purpose by public figures busy inventing themselves as "persons of influence" during the Brexit campaign does not rob it of its validity. The public legitimacy—or even the truth value—of this 350 million NHS claim may change as it is reformulated by different actors over time.

While digital platforms are today widely held responsible for the decline of public knowledge culture, it was precisely digital media that back in the 1990s promised to reinvigorate knowledge democracy, as they enabled a more dynamic approach to the epistemological evaluation of public sources and claims. Early platforms like the search engine Google were designed to disclose and valorize dynamic content, and search engines were celebrated as enabling exciting new applications of analytic techniques to public discourse, such as the detection of "emergent topics" and "networked relevance" (See for a discussion Rieder 2012; Marres and Weltevrede 2013).<sup>26</sup> With social media, the dynamic disclosure of digital content was re-defined in non-epistemic terms: social trends and viral memes emerged as the defining dynamics of these platforms. I agree with campaigners against disinformation that it is high time that we recover and revalue the epistemic dimension of online discourse. But the fact-checking methodology that guides many efforts towards this end only enables us to do half of the job. Yes, checking whether public statements correspond with known facts can help to secure the background conditions of public debate—to mark the stuff "everybody knows," or does well to know. But *in these not-quite modern times, the statements that we can or should "all" be able to agree about, and about which prescriptive normativity can be securely exerted, tend to be the less crucial,*

---

<sup>24</sup> Kentish, B. "Boris Johnson 'misused' figures with £350m Brexit claim, UK statistics chief says." The Independent, 17 September 2017 <https://www.independent.co.uk/news/uk/politics/boris-johnson-350m-brexit-nhs-misled-uk-statistics-authority-the-telegraph-sir-david-norgrove-amber-a7951711.html>

<sup>25</sup> Agerholm, H. "Brexit: NHS England boss demands £350m-a-week promised to Leave voters." The Independent, 8 November 2017 <https://www.independent.co.uk/news/health/brexit-latest-nhs-england-350-million-per-week-pledge-eu-simon-stevens-brussels-a8043191.html>

<sup>26</sup> Insofar as correspondence-based approaches to fact checking bracket this dynamic quality of digital knowledge, we may refer to these methods as "non-digital," notwithstanding their implementation by computational means.

*conditional statements that indicate the margins of public debate, while the claims that are at the center of public debate, and help organize it, are often marked by epistemic dynamism. The statement 350 million+NHS+bus may be reformulated, its truth value and the way it carves up the political landscape (for or against the cuts; for or against Europe) may change as time passes. We need methods to recover this processual character of public knowledge claims “in formation.”*

Besides the implementation of fact checks—which match claims to pre-authorized statements—in public discourse, then, we need to develop new ways of evaluating and new types of tests of public statements that will enable us to determine the robustness of experimental facts as they are formulated and reformulated over time. One contribution to this project is David Moats’ (2017) proposal to analyze public controversies as the curation of media ensembles. Moats highlights that online public debate and statement formation involve the activation of diverse media sources, which are brought into relation—sembled—through the circulation of engaging claims. To use the example of the Brexit bus again, the 350 million claim started off by mobilizing publics to vote for leaving the EU, but it subsequently activated a range of different sources, from NHS advocates to Greenpeace campaigners who bought and covered the bus with “thousands of questions... about what Brexit means for the environment.”<sup>27</sup> Public statements, then, are not and should not only be “in accordance with pre-given facts”; there is also a class of statements that may or may not agree with certified knowledge but that actively contribute to the formulation and re-formulation of new empirical truths by different actors. And while the public formulation of experimental facts is susceptible to dynamics of polarization along known dividing lines, just like other facts, it equally has the potential to carve up the public along less predictable and more productive lines than those that pitch epistemic against anti-epistemic communities: they may, for instance, assemble advocates of substantive causes (public health care, environment) into new alliances.

## Conclusion

To be sure, the methods and techniques that would allow us to establish the validity of experimental facts are in need of development. But at least this project opens up alternatives to the politics of demarcation. To highlight the limits of prevalent fact-checking methods, as I have done in this article, is **not** the same as arguing that we must give up on the delineation between more and less valid contributions to public knowledge. The point is *not*, as it seemed to be for Richard Rorty and some other critics of demarcationism in the 1980s, to declare an end to the normative project of establishing hierarchies of epistemic value. Rather, to test public media for experimental facts is to recognize that the category and genre of “fact” is undergoing transformation in contemporary societies. While the concern with “public knowledge” in times of crisis can seem elitist (Mirowski 2017), the coding of the relation between knowledge and publics as tenuous and brittle also indicates something else: traditional models for the validation of

---

<sup>27</sup> Greenpeace (2016) “Boris Johnson’s ‘Vote Leave’ battle bus re-branded outside Parliament.” 18 July. <https://www.greenpeace.org.uk/press-releases/boris-johnsons-vote-leave-battle-bus-re-branded-outside-parliament-20160718/>

public facts may be losing their efficacy.<sup>»</sup> In particular, it is no longer self-evident that public respect for facts can be secured through an appeal to the authority of experts or grounded in the authority of statements that are validated outside the public domain. Indeed, as long as factual knowledge derives its authority from an inaccessible realm beyond the public, it is all too likely to remain the target of anti-elite movements. In today's dynamic and diverse public spheres, epistemic authority will also have to be earned the hard way, through an exchange between epistemically diverse viewpoints.

This is why those who simply want their facts back can't have them (yet). As long as the work of re-constructing "the public fact" as a dynamic category remains incomplete, to want your facts back will amount to empty nostalgia: to say it will sound like a refusal to contribute to the unfinished project that is the re-construction of knowledge democracy. Indeed, if the latter ideal is to gain wider credibility and support, the validation of public facts will partly need to take place in the public sphere. Knowledge democracy as a reconstructive project requires transformation of epistemic ideals, including that of facticity. Facts are too important to be reduced to vehicles of the restoration of authority: their validity is always experimentally acquired, and the experimental validation of public facts must today happen in the public domain. It will take time, but we will need to re-envision what a public fact is for a world that is not only marked by contingency but also by epistemic diversity and dynamism, and where, consequently, experts aren't and shouldn't be fully secure in their public authority.

### **Acknowledgements**

I am indebted to Liliane Bounegru, Gerard de Vries and Tanja Bogusz for insightful comments. I also want to thank the Public Data Lab for their generous invitation to join the collective.

### **Author Biography**

Noortje Marres's work contributes to the interdisciplinary field of Science, Technology and Society (STS) and investigates issues at the intersection of innovation, everyday environments and public life: participation in technological societies; the role of mundane objects and devices in engagement; living experiments; the changing relations between social life and social science in a digital age. Noortje also works on research methodology, in particular issue mapping, and is interested in developing creative forms of inquiry between the social sciences, technology and the arts. Dr. Marres is currently Associate Professor at the Center for Interdisciplinary Methodologies at the University of Warwick (UK), a Visiting Professor in the Centre for Science and Technology

---

<sup>»</sup> Instead of giving up on demarcation, I propose that we move from a politics of demarcation to a politics of selection: while demarcation concentrates on the retrospective establishment of correspondence of public statements with presumably stable, pre-given atomistic statements, a politics of selection progressively establishes a referent for claims through an iterative process of locating and evaluating statement-networks in formation.

Studies at the University of Leiden (the Netherlands), and a Mercator Fellow in the Media of Cooperation Research Programme at the University of Siegen (Germany).

## References

- Angermuller, J. (2018). Truth after post-truth: for a Strong Programme in Discourse Studies. *Palgrave Communications*, 4(1), 30.
- Bounegru, L., Gray, J., Venturini, T. & Mauri, M. (2018) A Field Guide to Fake News. Amsterdam: Public Data Lab.
- Carnap, R. (1995 (1974)) *Introduction to the Philosophy of Science*, Dover: Dover Publications.
- Doctorow, C. (2017) "Demon-Haunted World." September, <http://locusmag.com/2017/09/cory-doctorow-demon-haunted-world>
- Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, 15(8), 1348-1365.
- Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19(5), 613-631.
- Healy, K. (2015). The performativity of networks. *European Journal of Sociology*, 56(02): 175–205
- Irani, L. C., & Silberman, M. (2013, April). Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 611-620). ACM.
- Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2016). Accountable algorithms. *University of Pennsylvania Law Review*, 165, 633.
- Lippert, I. (2016) "Umwelt – „Version 2b“: Das Programmieren ökologischer Fehlentscheidungen und Grundlagen für eine neue Umweltpolitik." *Leviathan* 44(3)
- Marres, N. (2018) What if nothing happens? On street trials of driverless cars as experiments in participation. *Sociology of Knowledge Yearbook*. S. Maessen, S Schneider, and S Dickel (eds). Dordrecht: Kluwer.
- Marres, N. (2017) *Digital Sociology: the Reinvention of Social Research*, Cambridge: Polity
- Marres, N. (2015) Why Map Issues? On controversy analysis as a digital method, *Science, Technology and Human Values*, September 2015 vol. 40 no. 5 655-686
- Mirowski, P. (2017). "Against citizen science." Aeon, 20 November, <https://aeon.co/essays/is-grassroots-citizen-science-a-front-for-big-business>
- Moats, D. (2017). From media technologies to mediated events: a different settlement between media studies and science and technology studies. *Information, Communication & Society*, 1-16.
- Muniesa, F. (2018) The Behavioural Shibboleth, In: *Inventing the Social*, N. Marres, M. Guggenheim and A. Wilkie (eds), Mattering Press.
- Nascimento, S., & Pólvara, A. (2013). "Opening up technologies to the social: Between interdisciplinarity and citizen participation." *Design Issues*, 29(4), 31-40.
- Newman, M., Barabási, A. and Watts, D., (2007). *The Structure and Dynamics of Networks*, Princeton, NJ: Princeton University Press.

- Rieder, B. (2012). What is in PageRank? A historical and conceptual investigation of a recursive status index. *Computational Culture*, 2.
- Rorty, R. (2009 (1979)). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
- Sismondo, S. (2017) "Post-truth?" *Social Studies of Science*: 3-6.
- Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *J. on Telecomm. & High Tech. L.*, 13, 203.
- Turner, S. (2003). *Liberal democracy 3.0: Civil society in an age of experts*. London and New Dehli: Sage.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.
- Ziewitz, M. (2016). Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*, 41(1), 3-16.