

# **A Global Charter for Truth: Social Epistemology for the Internet Age**

Joshua Habgood-Coote

---

IYTT WORKING PAPER No. 2  
JUNE 2021

in connection with IYTT's  
INTERNATIONAL YOUTH CONFERENCE 2020

# **A Global Charter for Truth: Social Epistemology for the Internet Age**

Joshua Habgood-Coote

---

IYTT WORKING PAPER No. 2

JUNE 2021

in connection with IYTT's  
INTERNATIONAL YOUTH CONFERENCE 2020

The **IYTT: INTERNATIONAL YOUTH THINK TANK** is a Gothenburg-based initiative mobilizing youth from diverse backgrounds across Europe with the aim of promoting a democracy movement based on open society values. Activities center around annual youth conferences in which participants develop and present policy proposals for strengthening an open and democratic society, while being brought together variously with executives from industry, academia, culture, politics, and civil society. Participants publish their proposals in a conference report and, engaged afterwards as Youth Fellows, develop them further into policy briefs through the "IYTT Bottom-Up Policy Advise Loop", a learning process involving open deliberations with decision-makers, scholars, peers in the IYTT European Youth Panel, and laypersons.

## **A Global Charter for Truth: Social Epistemology for the Internet Age**

Joshua Habgood-Coote

### **ABSTRACT**

This research overview focuses on policy proposals designed to address the crisis of truth which currently grips societies across the world. It builds on the proposals of the International Youth Think Tank's 2020 conference to reform social media, journalism, and data privacy, taking a step back to consider what problems for truthfulness are posed by these institutions, and what options we face to improve them. We start in section 1 by articulating 'post-truth' narratives about the crisis of truth, and arguing that they are at best an incomplete way to think about the crisis of truth. Section 2 offers an alternative conceptual framework taken from social epistemology. According to this proposal, the crisis of truth lies in the fact that we lack knowledge about topics which are important to us, and we ought to address this crisis by developing individual habits, social practices, and institutions which can help us to collectively acquire more knowledge about the things which are important to us. Sections 3, 4, and 5 apply this framework in turn to social media, journalism, and data privacy. Each section starts by framing the contribution of the area to the crisis of truth, before considering some different problems, and offering some possible policy proposals.

### **KEYWORDS**

social epistemology, post-truth, social media, journalism, data privacy, trust, free speech, democracy

### **CONTACT**

Joshua Habgood-Coote, [j.habgood-coote@leeds.ac.uk](mailto:j.habgood-coote@leeds.ac.uk)

The views expressed in this paper are those of the author and do not necessarily reflect the views of the IYTT.

## CONTENTS

Introduction	1
<b>1. What is the Crisis of Truth?</b>	2
<b>2. Social Epistemology and the Crisis of Truth</b>	4
<b>3. Social Media and Public Discourse</b>	16
<b>4. Journalism and the Pursuit of Truth</b>	28
<b>5. Data Privacy and Institutional Trustworthiness</b>	40
Conclusion	48
About the author	49
References	50

## Introduction

Over the last five years, the idea that we are living through a crisis of truth has become widely accepted and is transforming from a shocking warning into a mundane piece of common sense. Public discourse, popular books, and academic work are filled with concerns about the declining power of truth.<sup>1</sup> Truth is supposed to be under threat from fake news flowing through social media sites, from unashamed political mendacity, from algorithmically-produced echo chambers and polarisation, from the undermining of established media institutions, and from digital manipulation.

The aim of this research overview is to split up this tangle of complex issues into more manageable subproblems, and to give an overview of interventions which have been proposed to address these subproblems. We will focus on three parts of the supposed crisis of truth:

1. The role of social media in spreading falsehoods, enabling harassment, and commercialising the public sphere;
2. The function of the media as a reliable purveyor of truths, and how media funding, the consolidation of the media, the difficulty of distinguishing between reliable and unreliable sources, and attacks on journalists undermine this role;
3. The way in which unregulated data collection by technology companies undermines peoples' right to privacy, creates informational harms, and commodifies people.

Each of these subproblems are complex in their own right, and this overview cannot hope to claim to be comprehensive. The goal is not to offer the final word on any of these issues, but rather to get our arms around these problems, to understand the possible space of interventions, and to tease out some connections between these different issues.

The plan is as follows. Section 1 will interrogate the helpfulness of post-truth narratives for thinking about the role of truth, and suggest that we reframe the problem using ideas from social epistemology (the study of the importance of social factors to the pursuit of knowledge). Section 2 will offer a conceptual toolkit for thinking about the aims of different institutions, focusing on how to understand truth, knowledge, and trustworthiness. Section 3 will focus on the threats to truth in social media, section 4 will turn to journalism, and section 5 will consider data privacy.

---

1. See (Ball, 2017; D'Ancona, 2017; Davis, 2017; Fuller, 2018; McIntyre, 2018).

## 1. What is the Crisis of Truth?

One issue to address is understanding exactly what kind of problem (or problems) are involved in crisis of truth. Contemporary concerns about a crisis of truth have coalesced around the 2016 US election campaign, and the campaign that led up to the UK's vote to leave the European Union. Both campaigns were characterised by extremely prominent false claims, and by increasingly polarised factual disagreements. These events posed a threat to the establishments of the United States and the United Kingdom, and in response a number of authors and journalists concocted a post-truth narrative which promised to both explain these electoral events, and to offer a number of policy proposals to ameliorate the situation.

The term 'post-truth' originates with the Serbian-American journalist Steve Tesich. Commenting in 1992 on the Iran-Contra scandal—in which the US government secretly facilitated the sale of arms to Iran—Tesich writes:

We are rapidly becoming prototypes of a people that totalitarian monsters could only drool about in their dreams. All the dictators up to now have had to work hard at suppressing the truth. We, by our actions, are saying that this is no longer necessary, that we have acquired a spiritual mechanism that can denude truth of any significance. In a very fundamental way we, as a free people, have freely decided that we want to live in some post-truth world. (Tesich, 1992)

Tesich's rhetoric is undoubtedly compelling, but it is less clear than we might like. It is extremely unclear what the 'spiritual mechanism' that denudes truth of significance is, or how it is supposed to undermine the value of truth. Without an account of this mechanism, we are in no position to verify whether we have entered a post-truth era or not. More recent diagnoses of the advent of a post-truth era share Tesich's alarmism, and combine it with a wider political crisis narrative. In *Post-Truth: The New War on Truth and How to Fight Back*, the journalist Matthew D'Ancona writes:

Our own Post-Truth era is a taste of what happens when a society relaxes its defence of the values that underpin its cohesion, order, and progress: the values of veracity, honesty and accountability. These values are not self-preserving. Their maintenance is the product of human decision, agency and collaboration (D'Ancona, 2017, p. 112)

D'Ancona shares with Tesich the idea that truth has some central role to play in liberal democratic societies, and the concern that the citizens of western democracies have lost their grip on this value. The idea shared with other post-truth writers is that the political events of 2016 (especially in the UK and US) are indicative of a break, in which the established institutions which we historically relied upon for knowledge and reliable reporting have lost their power, allowing a combination of demagogic politicians, pseudo-scientists, technology

companies, and profit-motivated clickbait journalism to gain control over what people believe, with the consequences that harmful falsehoods are able to spread across society. According to this post-truth narrative, the remedy to the ‘post-truth’ era is to reestablish the power of established institutions, to replace demagogic politicians, rebuild trust in science and reliable journalism, and to regulate technology companies.

There are several problems with this post-truth narrative.

The first is that the post-truth narrative is historically inaccurate. The diagnosis of a post-truth era calls out for a return to the pre-post-truth (or just truth) era in which people trusted knowledge-producing institutions, and all was well with our intellectual lives. The problem is that there has never been a historical period in which knowledge-producing institutions worked without failures. These failures have just predominantly affected marginalised communities, and have been obscured by institutions with an interest in hiding their own failures. Newspapers and television news have a long history of spreading pernicious and persuasive falsehoods (consider media coverage of the AIDS crisis), of failing to cover important historical events (consider the lack of reporting lynchings in the American South), and of allowing themselves to be used as vehicles of government propaganda (consider the BBC during and after WW2). This is not to say that we ought to be sceptical of all media institutions, but to say that we ought to be critical friends of journalism, who are clear-eyed about historical problems with journalism that date back decades.

A second problem is that the post-truth narrative leads to a kind of exceptionalism about the present (Habgood-Coote, 2019, pp. 1056-7). Proponents of the post-truth narrative are fond of using neologisms like ‘fake news’, ‘echo chambers’ and ‘alternative facts’ to create the impression that the problems faced by democracies in the twenty-first century are unprecedented. While it is true that there are some novel features of the current information landscape and culture—such as the scale of connectedness enabled by the combination of broadband and Web 2.0 applications—the basic structure of the problems we face have a long history. The problems of distinguishing experts from imposters, of creating institutions to filter out false information and highlight important truths, and the problem of managing our dependence on other people for information have a history as long as the history of human sociality.

A related problem is that the neologisms on which the foundations of the post-truth narrative are built are poorly defined. Researchers have raised concerns that terms like ‘fake news’ are floating signifiers that shift between different meanings (Farkas & Schou, 2017), or nonsense terms that simply mean nothing (Habgood-Coote, 2019). Whether or not these diagnoses are true, terms like ‘fake news’ and ‘echo chambers’ are understood and used in pretty radically different ways, and using them can create confusion and verbal disagreement. By now these

linguistic problems are fairly widely accepted: both the UK Parliament's Digital, Culture, Media, and Sport committee and the European Commission's high-level commission on fake news have recommended dropping the term 'fake news', and replacing it with more clearly defined terms.

The third problem with the post-truth narrative is that it orients us towards conservative solutions. If the problem is that we have 'relaxed our defences' of the institutions which have historically filtered truth from falsity for us, then the solution is to rush back to the defence of those institutions, giving them greater funding, and urging other people to trust those institutions. While this defensive move may have a role in addressing our current situation, we should not exclusively focus our attention on rebuilding the media infrastructure of the 1960s and 1970s. We ought to at least consider more radical changes in the media landscape, including public ownership of media and technology companies, and efforts to build new kinds of institutions.

The final problem is that trying to wrap up all of the elements of the current crisis of truth into one 'post-truth' narrative oversimplifies them. In reality, we are facing multiple crises of truth spanning communication technologies, media funding, political culture, democratic culture, and data privacy. A singular narrative will tend to focus on one problem, whereas we need to address a number of distinct problems all at the same time.

Do the problems with the post-truth narrative mean that there is no crisis of truth? Can we relax? As nice as this would be, I don't think that the problem with this narrative is that it is creating problems from thin air, but rather that it is framing real problems in an unhelpful way. There is a crisis of truth (Hicks, 2020), but it is not a new problem, and the solution is not to try to recreate an ideal past media environment. We ought to look elsewhere for an orienting framework to frame the problems that might be associated with a crisis of truth. In the next section, I want to suggest that we can look to philosophy for the conceptual tools to more clearly see the problems which make up our current crisis of truth.

## **2. Social Epistemology and the Crisis of Truth**

If we want to address the crisis of truth, then we will need the right tools for the job. Some of these tools will be practical—new forms of technology, education, and policy proposals—but to deploy the practical tools appropriately, we need to have the right *conceptual* tools to think about what we want, and how to get it. We need conceptual tools that can help us to understand what is troublesome about our current situation, and what a better outcome would look like. In this section, I will make the case that philosophy can provide these tools, and lay out some groundwork for how we ought to think about knowledge, truth, and trust.



Why is philosophy the right discipline to look to? Philosophy is often associated with a kind of monkish detachment from the world and a concern with questions of little or no practical implication. Cultural representations are not filled with examples of philosophers contributing to practical problems in the same way that engineers, natural scientists, and economists do. In her essay “Philosophical Plumbing” Mary Midgley suggests a metaphor that illustrates why—despite its lack of visibility in everyday life—philosophy does have an important role in everyday life:

Plumbing and philosophy are both activities that arise because elaborate cultures like ours have, beneath their surface, a fairly complex system which is usually unnoticed, but which sometimes goes wrong. In both cases, this can have serious consequences. Each system supplies vital needs to those who live above it. Each is hard to repair when it does go wrong, because neither of them was ever consciously planned as a whole. (Midgley, 1992, p. 139)

Midgley’s point is that the value of both plumbers and philosophers is often overlooked, because their activity is focused on the maintenance and repair of systems that we rely on in everyday life without thinking about their workings. If a plumber has done her work fitting a new boiler, she won’t be needed until something goes wrong with the system. Much as the plumber is only called out when you have water flooding your bathroom, the philosopher’s work is only really seen when a conceptual system goes wrong. Conceptual systems can go wrong by giving rise to paradoxes, by leading to inconsistent sets of beliefs, or by leading to hermeneutic gaps in which we lack concepts to think about issues of concern. I suggest that we think about the problems with the ‘post-truth’ narrative as a philosophical plumbing problem. We are facing a bundle of problems about our collective pursuit of truth and knowledge, and the conceptual framework given to us by the ‘post-truth’ narrative is not adequate to represent either why our current situation is bad, or what a better situation would look like.

## 2.1 Epistemology and the Analysis of Knowledge

What conceptual plumbing should we put in the place of the ‘post-truth’ narrative? The natural suggestion would be to look to *epistemology*—the branch of philosophy that is concerned with knowledge, and the kinds of inquiry that aim at truth (as opposed to kinds of inquiry that focuses on *what to do*) (Nagel, 2014). Epistemology focuses on two central questions:

**The Nature Question:** What is knowledge?

**The Acquisition Question:** How can we get knowledge?

Setting to one side the impersonal sense of knowledge (as in ‘that book is full of knowledge’), the nature question focuses on what conditions are required for *someone* to know something. The central contrast to be explained is between *knowing* something is the case, and merely *thinking* it.<sup>2</sup> Consider two sentences:

Ann *knows* that her keys are in the bowl.

Bernard *thinks* that his keys are in the lock.

These sentences have two important differences.

The first difference concerns the status of the claims Ann and Bernard believe. Ann only knows that her keys are in the bowl if it is true that her keys are in the bowl. If her keys are behind the sofa, then she can’t know that that they are in the bowl (she merely thinks it!). By contrast, Bernard can think that his keys are in the lock no matter where they actually are. Philosophers and linguists call a verb *factive* when saying that someone stands in that relation to a claim presupposes that the claim is true.<sup>3</sup> ‘Knows’ is factive, and ‘thinks’ is not, meaning that knowledge can only relate us to true claims, in other words facts.

The second difference concerns Ann and Bernard’s support for their beliefs. Even if Bernard is right that his keys are in the lock, he might fail to know that proposition because he didn’t believe it for the right reasons. If he believes this claim because he made a random guess, or because he simply hoped it was there, then does not know it. By contrast, to know where her keys are Ann must believe a true claim, and have appropriate support for that claim, either in the form of an ability to explain why it is true, or in the form of a reliable source from which she gained this belief (perception, or a competent witness might be good candidates for a reliable source).<sup>4</sup>

- 
2. In many countries the educational curriculum replaces the distinction between merely thinking and knowing with the distinction between *facts* and *opinions*. This is an inadequate piece of philosophical plumbing. It is treated as an exclusive distinction, but it is quite possible to have an opinion about a fact; my belief that snow is white is an opinion concerning a true matter of fact. It is also not a distinction between the same kinds of things: facts are *true claims*, but opinions are mental *attitudes* which we have about claims which might be true or false. This distinction is often associated with the idea that in a disagreement between people with opposing opinions, there can be no fact of the matter about who is right (it’s just a matter of opinion!). This is false: opinions about questions of fact can be true or false, or be believed for better or worse reasons.
  3. For example, ‘regrets’ is a factive verb, but ‘remembers’ is not factive.
  4. The first kind of support (accessible reasons) is often associated with *internalist* theories of justification, and the second kind of support (a reliable source) is often associated with *externalist* theories of justification. Internalist theories say that justification must be consciously accessible, and externalists deny this. This distinction won’t be important for our purposes, and we will assume that both accessible reasons, and reliable sources are routes to possessing knowledge.

These two observations point us towards a view which has become central in epistemology in the last seventy years: *the justified true belief account of knowledge*.<sup>5</sup> This view offers the following account of what knowledge is:

A person *S* knows a claim *p* if and only if i) *p* is true, ii) *S* believes that *p* is true, iii)  
*S* is justified in believing that *p* is true.<sup>6</sup>

In ordinary language: to know a claim, you need to believe it, have justification for believing it, and the claim needs to be true. This view proposes that knowledge involves an ingredient external to the knower—a true claim—and an ingredient internal to the knower—a justified belief.

A couple of clarificatory points about the justified true belief analysis of knowledge.

This account does not entail that everything which we ordinarily call knowledge involves a justified true belief. Knowledge can be attributed incorrectly: we can say that someone knows something which they don't, and fail to recognise that they know what they do.<sup>7</sup> Some of the things that we think we know are false, others are unjustified, and some of them aren't believed. Even claims that we are certain about can turn out to not involve knowledge: early modern scientists might have been sure that combustion was explained by phlogiston (and they had some good evidence for that claim!), but they did not know it, because it is a false claim. Some philosophers have denied the distinction between calling a belief knowledge and it being knowledge, but we will assume that knowledge is like gold, pedigree dogs, and love, in the sense that we can make mistakes about whether something is the real deal.

The justified true belief analysis is an account of knowledge, and not an account of truth, or justification. But it does presuppose that there are such things as true claims and better or worse reasons for believing something. For our purposes, we can assume that a true claim is simply a claim that corresponds to the way that reality is.<sup>8</sup> We will set to one side anti-realist philosophers who deny that there is such a thing as truth, argue that truth is relative to a subject's perspective, or argue that there are no such things as good reasons to believe something. Responding to these views is beyond the scope of this research overview.

This is far from the last word about the Nature Question, but the justified true belief account of knowledge provides us with a helpful tool to think about the crisis of truth, offering us a

- 
5. This view is often traced back to Plato's *Meno*, and presented as the accepted view, but recent history of philosophy suggests that it is a much more recent development. (Dutant, 2015)
  6. The phrase 'if and only if' is used by philosophers to mark equivalence claims. The idea of this analysis is that having a justified true belief in a proposition is both *necessary*, and *sufficient* for knowing that proposition.
  7. Most properties work like this, although there are exceptions. Arguably, if we all think that something is money, then it is.
  8. For details of theories of truth, see <https://plato.stanford.edu/entries/truth/>

way to think about the problems with our current situation, and what a better outcome might look like.<sup>9</sup>

We are collectively in a bad spot right now, because lots of us don't have knowledge about topics that matter to us (the safety of vaccines, how to tackle climate change, whether the 2020 US election was legitimate). The justified true belief account reminds us that there are a number of (possibly overlapping) ways to lack knowledge about some claim: 1) by having a false belief, 2) by not having justification, 3) by not having a belief at all. This reminder is important, because it makes clear to us that the crisis of truth is not merely about the number of false claims in circulation, but also about peoples' beliefs, and what justification is available in a social context. One way to destroy knowledge is to circulate lots of false beliefs, but another way is to simultaneously circulate true and false beliefs which are difficult to distinguish, with the consequence that people withhold judgement.<sup>10</sup>

What would a better situation look like? Well, we want to be in a situation where we can have more knowledge about subjects that matter to us. We don't want knowledge about just anything—how many blades of grass there are in the park is for most purposes an unimportant question—rather we want knowledge about subjects which are either theoretically or practically important. We want to have knowledge about the explanation for the motions of the planets, what kinds of political systems will allow us to manage conflict, and what kinds of diets will allow us to stay healthy. We don't necessarily want to maximise how much each individual person knows — in some cases we will be happy with a division of labour in which specialists have knowledge about topics (plumbing, theoretical physics, epidemiology) which are not commonly known—but we certainly are invested in increasing the sum of collective knowledge.

If we want to get to a situation in which we know more about important topics, we should turn to the acquisition question: how we can gain knowledge, both as individuals, and as groups.

Philosophers have tended to think about the Acquisition Question in an individualistic way. They have focused on routes to knowledge for individuals—typically perception, memory, and reasoning—neglecting sources for knowledge that involve other people. This is surprising, given the extent to which we are dependent on other people and institutions for our everyday beliefs. In a day we rely on others to an extraordinary extent. We rely on newspapers to report current events, on our friends and neighbours to inform us about goings-on where we live, and on scientists and health professions to give us advice about how to avoid illness. As an

---

9. In fact, this analysis is only really the start of inquiry into the Nature question. For more, see <https://plato.stanford.edu/entries/knowledge-analysis/>

10. <https://www.nytimes.com/2016/11/28/opinion/fake-news-and-the-internet-shell-game.html>

exercise it is worth thinking about your reasons for some beliefs you take yourself to know to see how many of them essentially rely on trust in other people.

Over the last thirty years, philosophers have tried to remedy this situation, considering the importance of *social* routes to knowledge. This field has come to be known as *social epistemology*, and it investigates a number of topics, including how to think about the acquisition of knowledge from testimony (Adler, 2006), how to identify experts on specialist topics (Goldman, 2001), and how institutions can be designed to promote knowledge (Goldman, 2010). If we want to think about ways in which the crisis of truth might be ameliorated, social epistemology has a range of useful conceptual tools. Discussions about when we should rely on testimony can help us to think about how individuals ought to manage their dependence on others for information. Discussions of expertise can help us to think about how people should identify trustworthy sources on topics that lie outside of their ken. And discussions of institutional design can help us to think about how knowledge-producing institutions, including social media sites, and newspapers can be better designed to deliver their users knowledge.

Social epistemology offers us an alternative to ‘post-truth’ narratives as a way to think about policy questions about how to address the crisis of truth. Instead of focusing on the somewhat abstract task of *defending truth*, we can think together about how we can improve our institutions, social practices, and individual habits in order to increase the amount of knowledge we have about important topics. There will be an important role for abstract theorising in thinking about what the intellectual goals of our institutions and social practices are (do we want knowledge, truth, or understanding), but the project of designing better institutions will also need to rely on empirical research, and models of intellectual communities to understand how we can achieve our collective intellectual goals.

## 2.2 The Epistemology of Democracy

Researchers working in social epistemology have paid particular attention to the role of democratic political institutions in the acquisition of knowledge. In her paper “The Epistemology of Democracy” (Anderson, 2006), the philosopher Elizabeth Anderson discusses three competing different models for thinking about the knowledge-generating powers of democratic societies.

The Condorcet Jury theorem—named after the 18<sup>th</sup> century mathematician the Marquis de Condorcet—focuses on the properties of voting systems. The theorem states that in a situation in which voters are choosing on a binary either/or choice in which the majority position wins, the voters vote independently, and they are on average more likely to get things right than wrong, the probability that the majority result will be correct will be *higher* than the

average probability that a voter will get things right. Not only that, but the more voters you add, the more likely that the majority position will be correct, with larger groups tending towards perfect reliability in the limit case. This theorem can be proven using fairly straightforward maths, and can be generalised to apply to more cases (for example, allowing for more than two options, and for different voting rules [List & Goodin, 2001]). The Jury theorem is extremely significant because it means that in certain situations, the reliability of a group's vote can outperform the reliability of individual voters, meaning that the collective performs better than its members, and giving an intellectual rationale to majority voting systems with a large franchise.

The Diversity Trumps Ability theorem—proven by Hong and Page (2004)—focuses on smaller groups, and states that when problems are hard (meaning that no individual performs perfectly), but have a finite set of solutions, groups composed of randomly chosen problem-solvers will outperform groups composed of more reliable individuals. This theorem gains support from computational models, and from a mathematical proof (which relies on modelling assumptions). This result has been used to support efforts to diversify workplaces and educational systems, and to replace expert-run systems with diverse committees, but does come with a couple of caveats (Grim et al., 2019). Firstly, the theorem concerns diversity in problem-solving, which may not straightforwardly correspond with diversity in social identity. Second, it is not clear that the high-performing individuals in Hong and Page's model are really experts. What the Diversity Trumps Ability theorem does establish is that when we are constructing a problem-solving group, we ought to pay attention not just to individual reliability, but to the way individuals with different approaches mesh together, and the fact that team performance often depends on the heterogeneity of team members.

Anderson argues that the Condorcet Jury theorem and the Diversity Trumps Ability theorem are at best partial models for the knowledge-generating functions of democracy. The Jury theorem captures the idea that we want to include as many people in voting systems as possible, but fails to model the importance of socially distributed knowledge for democratic society, and focuses on elections to the exclusion of the everyday dynamics and discussion that make up democratic life. The Diversity Trumps ability theorem also fails to deal with the street-level processes of democracy, and fails to account for the importance of universal inclusion to democracy (for it, diversity matters more than size). These results might remain important for understanding parts of our democratic lives: for example the Jury theorem is important for understanding the value of voting systems, and under which conditions voting might be useful as a problem-solving tool (as opposed to being a tool for generating political legitimacy), and the Diversity Trumps Ability theorem might be a useful result for thinking about the make-up and value of small decision-making bodies, like citizens' juries (see the IYTT research overview on citizens' juries [Geib, 2021]).

Anderson's alternative model is what she calls *Deweyian Experimentalism*, named after the American philosopher John Dewey. Deweyian Experimentalism conceives of democratic society as the application of collective intelligence to practical problems, and thinks of our everyday political discussions as part of a wider collective enterprise of solving political issues, much as a scientist's everyday lab work is part of the collective endeavour of chemistry. This model requires *diversity* because socially important knowledge (for example knowledge about the effects of policies, or about current social problems) is widely distributed amongst society, *discussion* as a mechanism for bringing together people from different walks of life to discuss the public interest, and *dynamism* because although deliberation concerns matters of fact, there is no guarantee that discussion will yield the right result first time around.

### 2.3 Free Speech and the Epistemology of Democracy

The previous two sections have argued that we should be framing the crisis of truth by asking how we can design social institutions that will contribute to the acquisition of knowledge about important topics, and that democratic institutions like voting, citizens' juries, and public discourse have a crucial role in contributing to the common stock of knowledge. Before we start considering how we might change our currently existing institutions, we need to consider how this institution-building project relates to the value of freedom of expression. There are two issues to consider:

1. Do individual rights to freedom of expression conflict with the project of designing institutions to produce socially important knowledge?
2. Does the idea that democratic societies produce knowledge through a system of freedom of expression provide us with a fairly simple answer to the question of how to design our knowledge-producing institutions?

Let's take these issues in turn.

Many of the interventions by social media companies which are designed to address false, misleading, and uncooperative contributions have been criticised on the grounds that these interventions undermine the value of freedom of speech.<sup>11</sup> These responses often come from conservative politicians, who are committed to a narrative that claims that social media companies are systematically biased against them. (If anything, the opposite appears true

---

11. <https://eu.usatoday.com/story/tech/2020/11/17/facebook-twitter-dorsey-zuckerberg-donald-trump-conservative-bias-antitrust/6317585002/>  
<https://eu.usatoday.com/story/tech/news/2016/11/18/conservatives-accuse-twitter-of-liberal-bias/94037802/>

[Barrett & Sims, 2021]). Setting to one side the evidence for these claims of bias, there is a general worry that any of the interventions to make democratic institutions better for generating knowledge will undermine individuals' right to freedom of speech. This kind of concern might be associated with absolutist views of freedom of speech which think that free speech is both a fundamental right, and a good to be maximised by whatever means necessary.

There are a couple of things to say about this worry. First, lots of interventions which aim to improve the intellectual quality of public discourse will have no effect on freedom of expression. For example, a fact-checking service that issues corrections to claims that are demonstrably false without stopping those claims from being made will have no effect on the freedom to say false things. Secondly, the view that freedom of speech is an exceptionless right that overrides other rights is simply an implausible view. In John Stuart Mill's classic defence of the freedom of speech in *On Liberty* (Mill, 1859), he argues that the right to freedom of speech (and the right to freedom of action more generally) applies only in cases in which other people are not harmed. If a piece of speech will cause significant harm to other people, then on Mill's view it can legitimately be restricted (Mill's famous example is shouting 'fire!' in a crowded theatre, causing a rush for the exits which leads to several people being trampled). Harm-based exceptions mean that it is possible to restrict hate speech and other kinds of harmful speech without undermining individuals' right to freedom of speech. Thirdly, as Cass Sunstein argues in *#Republic* (Sunstein, 2017), many of the defences of the legal right to freedom of expression are based in the knowledge-generating power of public discourse. For example, Oliver Wendell Holmes writes in his famous dissenting opinion from 1919 that:

When men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by the free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes can safely be carried out. (Abrams v. United States 250 U.S. 616 [1919])

Rather than contrasting the goals of protecting freedom of speech and pursuing the truth, in Holmes' view the value of the 'free trade in ideas' comes from its ability to help us collectively reach the truth. This view is shared with many other defenders of the right to freedom of expression, including Mill, and means that tweaks to public discourse can be motivated from the same grounds that we justify the value of freedom of speech. These considerations suggest that the value of freedom of speech does not preclude attempts to improve the intellectual quality of public discourse, although there are some policies to improve the intellectual quality of public discourse that would undermine freedom (for example, a policy



which punished anyone who said something false by permanently removing their ability to contribute to public discourse).

The idea that our best bet for pursuing the truth is Holmes's marketplace of ideas brings us to the second issue. If free exchange on the model of a market for trading goods is our best bet for collectively reaching the truth, then one might think that we ought to take up a policy of non-interference in public discourse, letting the market do its work. The metaphor with commercial markets for other goods highlights the problems with this idea. Although markets can be effective ways of producing and distributing goods that people need, they can only work well when they are regulated for example by antitrust legislation to prevent the formation of monopolies, and by consumer protection legislation to outlaw people being sold goods that harm them. Similarly, the marketplace in ideas will only be a useful way to collectively pursue the truth if it is regulated to prevent the characteristic harms of marketplaces. These regulations might include restrictions to harmful speech (on the grounds of Mill's harm principle), special protections for the speech of minority groups who might lack the social power to get uptake for their views, measures to prevent monopolies in the provision of information, and protections against certain kinds of informational harms that impede the collective pursuit of knowledge.<sup>12</sup>

In this section we've seen that the right to freedom of speech does not preclude interventions which aim to improve the intellectual quality of public discourse, and that interventions motivated by our interest in collective knowledge are in fact consonant with an important tradition for justifying the right to freedom of speech.

## 2.4 Trust and Trustworthiness

Another piece of conceptual plumbing which we need to make sense of the crisis of truth is the distinction between *trust*, and *trustworthiness* (O'Neill, 2002; 2020). Trust has an important role to play in the acquisition of knowledge. Trust has a central role in enabling human sociality, and that role is also played out in our collective pursuit of knowledge. By trusting other people we can gain knowledge about the world beyond our immediate experience, including knowledge about specialist topics about which we have limited understanding. For example, I couldn't start to explain how the Large Hadron Collider works, or how theoretical physicists use statistics to understand the data it produces, but if I trust a theoretical physicist who tells me that the Higgs boson exists, I can come to know that the Higgs boson exists, and that its existence provides some evidence for the standard model of physics. The social

---

12. For a discussion of the ways even passing on *true* information can impede public discourse, see (O'Connor & Weatherall, 2019).

importance of trust means that surveys indicating decreasing levels of public trust in experts are worrying, and can motivate calls for increasing public trust in authorities.

Commentators often claim that low levels of public trust are a problem and call for measures to increase the public's trust in experts and knowledge-generating enterprises. Although these calls are well-meaning, they fail to appreciate the importance of the distinction between trust and trustworthiness, which has been stressed by the philosopher Onora O'Neill. Trust is important for us, but by itself more trust is not necessarily a good thing. It would be bad if the public increased levels of trust by placing their confidence in scientists who were interested in pursuing financial gain, or in newspapers which were politically biased. Speaking generally, we should place our trust in others when they both intend to live up to their commitments, and are competent in doing so. As O'Neill points out, what we want is not simply *more trust*, but for trust to be placed in *trustworthy* individuals and institutions. Presumably some institutions are already plenty trustworthy, and the problem is just to get people to recognise this fact. But in many other cases, public distrust stems from breaches of trust. For example, in the United Kingdom, rates of uptake of COVID vaccines amongst BAME people has been rather lower than in other ethnic groups.<sup>13</sup> At first pass, we might think that this distrust marks a failure of these communities to recognise the authority of scientific judgement. However, if we look back at the history of the relationship between these communities and the medial establishment, there are a number of striking cases in which medical professionals have abused the trust of BAME communities (notably, the Tuskegee experiments, and rates of maternal death amongst Black women). In this case, the remedy is not for people to simply trust experts, but rather for the experts to become more trustworthy, and to display their trustworthiness in an open and accountable way. In *Mistrust*, Ethan Zuckerman (Zuckerman, 2020) makes a related point, arguing that public mistrust is not an intrinsic problem, but rather a signal that institutions need to transform in order to meet legitimate expectations by making themselves more *trustworthy*.

## 2.5 What is to be done?

So far, this section has given us a toolbox of evaluative tools for thinking about knowledge, the ways in which democratic societies function as a knowledge-producing enterprise, the value of free speech, and the importance of trustworthiness. In closing I want to move to a more practical frame, to consider the kinds of policy interventions which might be within the scope of our discussion.

---

13. <https://www.medrxiv.org/content/10.1101/2021.01.25.21250356v3>

We will be discussing several social practices and institutions, with an eye to how well they are doing with respect to providing us with knowledge. Some of these practices and institutions are actively causing harms, and in other cases they simply are not performing as well as they might. Many of these practices and institutions are partially or completely based on market mechanisms, which in some cases will help us explain why they are working badly, either because market-based systems are not good at producing knowledge in the relevant area, or because the market has become a monopoly.

We will discuss a few different kinds of policy interventions:

1. Legal regulations: when practices and institutions are actively harming their users or society at large, it will be appropriate to think about whether their operations might be legally restricted in order to avoid these harms. For example, we might want to consider making certain kinds of data collection illegal (as the EU's GDPR legislation has).
2. Breaking up monopolistic markets: much of the concern about social media and technology companies has been framed in terms of their size and control over markets in advertising and technology products. We might want to consider how the concerns about the monopolies held by technologies impacts on the role of their products in knowledge-acquisition, and whether antitrust legislation might ameliorate some of these problems.
3. Changing the way institutions are run: in some cases, the problems with journalism and social media will not involve the kinds of harms which are suitable for legislation, so we might want to consider whether voluntary changes to these institutions will help to address their problems. For example, we might not need legislation to persuade social media companies to put contextual watermarks on posts from news organisations.
4. Establishing new institutions: institutions are path-dependent, in the sense that how they work depends on how they have developed through time. It might be that some of the problems with our knowledge-providing institutions are determined by their development, meaning that we need to propose new institutions to avoid the problems of the ones we currently live with. For example, we might need to think about constructing new kinds of social media companies to avoid the privacy invading effects of currently existing companies whose financial model is intimately tied to surveillance capitalism (Zuboff, 2019).

### 3. Social Media and Public Discourse

In the *Phaedrus*, Plato's Socrates recounts an Egyptian myth about the invention of writing. Theuth is an industrious god, who makes many advances in mathematics and geometry, including inventing a system for writing. He takes these inventions to the god Thamus, the King of Egypt, hoping to provide these inventions to all of the inhabitants of Egypt. Socrates recounts Theuth's sales pitch for writing, and Thamus's response:

'This,' said Theuth, 'will make the Egyptians wiser and give them better memories; it is a medicine both for the memory and for wisdom.' Thamus replied: 'O most ingenious Theuth, the parent or inventor of an art is not always the best judge of the utility or inutility of his own inventions to the users of them. And in this instance, you who are the father of letters, from a paternal love of your own children have been led to attribute to them a quality which they cannot have; for this discovery of yours will create forgetfulness in the learners' souls, because they will not use their memories; they will trust to the external written characters and not remember of themselves. [275a-b Jowett trans].

Socrates concurs with Thamus' dim view of writing:

I cannot help feeling, Phaedrus, that writing is unfortunately like painting; for the creations of the painter hold this or that pose as if alive, and yet if you ask them a question they preserve a solemn silence. And the same may be said of speeches. You would imagine that they had intelligence, but if you want to know anything and put a question to one of them, the speaker always gives one unvarying answer. And when they have been once written down, they are tumbled about anywhere among those who may or may not understand them, and know not to whom they should reply, to whom not: and, if they are maltreated or abused, they have no parent to protect them; and they cannot protect or defend themselves. [275e-f Jowett trans].

Plato's fable about the invention of writing offers a useful model for thinking about debates about new technologies. New technologies are marketed with hyperbolic claims about the way they will extend and enhance human abilities. Theuth markets writing as a solution for the frailty of human memory; the founders of Facebook, Twitter, Weibo, Reddit, and MySpace sold social networking sites as solutions to enable human connections and democratise discourse.<sup>14</sup> Just as Theuth overlooked the flaws in writing technology due to his parental affection, the founders and developers of social media sites failed to see the problems which these sites were creating (which were widespread long before 2016).<sup>15</sup> And just as Socrates

---

14. For a classic discussion of the ideology of 1990s Silicon Valley, see <https://www.metamute.org/editorial/articles/californian-ideology>

15. <https://slate.com/technology/2019/04/black-feminists-alt-right-twitter-gamergate.html>

raises concerns about the effects that writing will have on human intelligence, making dire claims about writing technology disrupting the value of conversation, critics of social media have made disastrous predictions about the effects of social media on knowledge production and individual intellectual character.

It is difficult to take Plato's warnings about the dangers of writing too literally (they do come to us in the very medium he decries). The lesson we should take away from Plato's story is that neither the parents of new technologies, nor their immediate critics are the best judges of their merits and demerits. While the parents of social media were caught up in their own marketing, making grand claims about connecting people and creating a new public sphere, some of the more hyperbolic critics of social media are caught up in warnings about the imminent demise of democratic society.<sup>16</sup> We need to account both for the negative and positive features of social media. This is especially important given the incredible heterogeneity of communities on social media sites.

Social media sites have some good features, including:

1. Social media allow resistance movements to form quickly, mobilising large numbers of people to co-ordinate around issues of public concern (Tufekci, 2018; Jackson, Bailey, & Foucault Welles, 2020);
2. Social media allows marginalised minority groups to form communities in ways that might be difficult or impossible in physical space (Brock, 2020);
3. Social media sites (especially in their earlier iterations) enable rich textual and visual communication between people, allowing them to cultivate the social good of friendship (Briggle, 2008).

It is also uncontroversial that social media sites have serious problems. We will focus on three problems:

1. Social media sites allow—and sometimes promote—kinds of communication that undermine the collective pursuit of knowledge;
2. Social media sites allow for large scale harassment and hate speech against minority groups, with the consequence that they are marginalised from the public sphere;

---

16. This hyperbole can lead to some fairly out-there policy proposals, such as limiting social media use to 30 minutes a day <https://www.vox.com/recode/2019/7/31/20748732/josh-hawley-smart-act-social-media-addiction>

3. Social media sites are designed to facilitate the pursuit of profit, meaning that they are designed to promote features which undermine the pursuit of knowledge (including gamification, algorithmic filtering of information, and promoting content which maximises engagement).

After sketching out these three problems, we will consider some policy proposals which might help to address these problems, focusing on moderating attempts to derail inquiry, dealing with online harassment, and alternative financial models for social media.

### 3.1 Knowledge-Undermining Communication

In November 2016, BuzzFeed News published a story written by Craig Silverman and Lawrence Alexander about the imminent US presidential election being influenced by people living in Veles, North Macedonia (then known simply as Macedonia).<sup>17</sup> For reasons that remain obscure, young people in Veles had realised that if they set up sites with highly partisan content about the US election, they could attract enormous amounts of traffic which would allow them to make decent money from adverts. To maximise visitors to sites like TrumpVision365.com, USConservativeToday.com, DonaldTrumpNews.co, and USADailyPolitics.com, their owners populated them with sensational and often false headlines taken from fringe US sites, featuring headlines declaring the Pope's support for (then-candidate) Trump, and imminent criminal proceedings against Hillary Clinton. According to Silverman and Alexander, the owners of these sites were disinterested in the outcome of the election, or in the truth or falsity of the stories they posted, and were instead motivated by the opportunity to earn money (in 2016, North Macedonia had youth unemployment rates of 48.2%).<sup>18</sup>

Headlines of the form 'Macedonian Teens Run Fake News Sites!' became emblematic of concerns about the role of social media sites in our intellectual culture, tying together concerns about political propaganda, worries about the financial incentives of the online advertising market, and an avaricious adversary motivated by financial advantage. In the aftermath of the 2016 election, social media sites responded to these worries by presenting themselves as 'fighting fake news'.<sup>19</sup> Through the 2020 election they trialled several

---

17. <https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo#.bkVqlqB57Y>

18. <https://data.worldbank.org/indicator/SL.UEM.1524.ZS?locations=MK>

19. <https://www.facebook.com/formedia/blog/working-to-stop-misinformation-and-false-news>  
[https://blog.twitter.com/en\\_us/topics/product/2020/updating-our-approach-to-misleading-information.html](https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html)

<https://www.reuters.com/article/us-eu-tech-tiktok-idUSKBN23G2XM>

interventions, including labels on contested content,<sup>20</sup> nudges to comment rather than simply retweet,<sup>21</sup> and algorithmic detection of misinformation.<sup>22</sup> If anything, throughout 2020 and 2021, concerns about widespread false information became even more serious, as claims about COVID-19 which posed immediate harms to people were widely circulated (Cinelli et al., 2020).<sup>23</sup>

Research on social media by journalists and academics has established that false or misleading stories are widespread on social media. In another study from 2016, Silverman found that false election stories generated much more engagement than true stories.<sup>24</sup> Research on Twitter suggests that false stories spread further, faster, and more broadly than true stories.<sup>25</sup> Much remains unknown, and there are important questions about how effective political misinformation is,<sup>26</sup> but it is clear that there is a significant problem posed by persuasive false information on social media.

Since 2016, terms like ‘fake news’ and ‘misinformation’ have been caught in the cross-fire of what can only be described as *lexical warfare*, as commentators from the left, centre and right with different factual and ideological views try to establish what kinds of claims are acceptable, and which are not. I propose that we declare neutrality and avoid these debates, for a couple of reasons. Firstly, it is not at all clear that there is an informative way to draw a line around a category of problematic views of claims, besides taking Mill’s liberal view that claims which are likely to lead to harm ought to be restricted. Secondly, framings of the crisis of truth that focus on ‘fake news’ or ‘misinformation’ often fail to reckon with the heterogeneity of threats to knowledge, which as we saw in section 2 can be undermined not only by false claims, but also by claims that undermine justification, and claims that lead people to withhold belief. Thirdly, operating with a category of ‘fake news’ which is somehow problematic or appropriate to censor opens the door for oppressive regimes to apply this label to what ever news stories and news sources they don’t like. This is a pattern which we have seen world-wide, including in the Philippines, East Africa, China, and Hungary.<sup>27</sup>

---

20. <https://www.theverge.com/2020/8/12/21365305/facebook-covid-19-warning-notification-post-misinformation>

21. <https://www.theverge.com/21524092/twitter-temporarily-changing-retweet-quote-tweet-election>

22. <https://ai.facebook.com/blog/using-ai-to-detect-covid-19-misinformation-and-exploitative-content/>

23. [https://www.who.int/health-topics/infodemic#tab=tab\\_1](https://www.who.int/health-topics/infodemic#tab=tab_1)

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

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

26. <https://web.stanford.edu/~gentzkow/research/fakenews.pdf>

27. <https://www.theguardian.com/media/2018/apr/24/global-crackdown-on-fake-news-raises-censorship-concerns>

I want to take a different tack and suggest that we shouldn't think about good and bad kinds of content, but about different ways in which contents relates to collective knowledge-production.<sup>28</sup> If social media sites function to enable democratic conversation, then—following Anderson's Experimentalist view of public discourse—we can see social media as a forum for the collective pursuit of knowledge about matters of common political concern. This kind of forum aims to collect and produce knowledge about what is going on in society, and what can be done about it. Lots of people are committed to this effort, but various people—for whatever reason—aim to systematically derail collective inquiry. They might do this by making persuasive false claims, by spreading a mix of true and false claims which leads people to suspend judgement, by spreading doubts about justified claims, by trying to discredit other parties, or by spreading old-fashioned political propaganda (Stanley, 2015).<sup>29</sup> The point is that all of these strategies undermine the collective effort to find out what's going on and what to do about it. By shifting our efforts from trying to strictly define forms of problematic content to an approach which focuses on intentions and intended effects, we can better recognise the diversity of derailing efforts, and remain open to new techniques.

One important point which is often overlooked is that efforts to derail collective inquiry often interact with racism, and other systems of political oppression (Mills, 2008; Noble, 2018). Dylan Roof, the American white supremacist who murdered nine African Americans in a church in South Carolina, claims that he was radicalised by typing 'black on White crime' into Google, which directed him to several websites promoting the idea that the United States is in the grip of a crisis of Black violence targeted against white people (Noble, 2018, Ch. 4). White supremacists have always been early adopters of internet technologies and use websites and social media as a tool to promote their views and recruit others.<sup>30</sup> Technology companies are not neutral bystanders in the spread of racist misinformation: Google should have known that it was promoting white supremacist search results, and could easily have adjusted its results to filter out explicitly racist websites. Online racism and misogyny is often presented as problem specifically about *hate speech*, which is to say speech that expresses hate, or an intention to harm a particular social group, but many of the most worrisome kinds of misinformation are about racialised minorities, women, and gender diverse people.

---

<https://www.economist.com/international/2021/02/11/censorious-governments-are-abusing-fake-news-laws>

28. See also <https://points.datasociety.net/agnostology-and-epistemological-fragmentation-56aa3c509c6b?gi=24206fdd781b>

29. 'Propaganda' is also a contested term, but it is vital is we want to think about the way political power undermines collective inquiry. See <https://www.nytimes.com/2016/12/26/books/how-propaganda-works-is-a-timely-reminder-for-a-post-truth-age.html>

30. <https://journals.sagepub.com/doi/full/10.1177/1536504218766547>



### 3.2 Harassment and Public Discourse

Different people use social media sites to do a dizzying variety of different things, and have drastically different experiences of what it is like to spend time online. One issue which was largely overlooked by the mainstream press—at least until Gamergate<sup>31</sup> in 2014—is the daily reality of online harassment for women, people of colour, and LGBT people. A PEW research study in 2020 found that 41% of Americans report have personally experienced online harassment, with 25% experiencing severe harassment. Although women do not report more harassment, they appear to be targeted by more severe forms of harassment. 70% of LGBT people reported being harassed, with 50% reporting severe harassment, and around half of Black and Hispanic people reported being harassed because of their race.<sup>32</sup> Harassment is also targeted at journalists and politicians, often in ways that are compounded by race and gender.<sup>33</sup> A study by Amnesty International found that during the 2017 UK election, the Labour MP Diane Abbott received half of all harassment on Twitter.<sup>34</sup> And a study by the International Centre for Journalists<sup>35</sup> found that three-quarters of female journalists report online harassment, with 20% of respondents reporting that online harassment had led to offline attacks, and qualitative research backs up the seriousness of the threats faced by female journalists.<sup>36</sup> Online harassment is not simply a problem about malicious citizens, in some cases the harassment is perpetuated by the government.<sup>37</sup>

Online harassment is a problem for at least two reasons. First, it leads to psychological and material harms for its targets. Secondly, it has an exclusionary effect on the public sphere, leading marginalised groups to self-police their speech in order to avoid harassment. Sarah Sobieraj argues that the targeted online harassment of women leads to a serious democratic deficit, whereby women are unable to contribute to public discourse on equal footing. She draws a helpful analogy with offline harassment, saying: “just as inhibited use of physical public spaces is a spatial expression of gender-based expression, inhibited use of chat rooms, social media platforms, blogs, vlogs, and online gaming must be understood as a digital expression of these power dynamics.” (Sobieraj, 2018, p. 1701).<sup>38</sup> If the intellectual ideal for a

---

31. <https://www.vox.com/culture/2020/1/20/20808875/gamergate-lessons-cultural-impact-changes-harassment-laws>

32. <https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/>

33. <https://www.ndi.org/tweets-that-chill>

34. <https://global.oup.com/academic/product/the-culture-of-connectivity-97801999970780?cc=gb&lang=en&>

35. <https://medium.com/@AmnestyInsights/unsocial-media-tracking-twitter-abuse-against-women-mps-fc28aeca498a>

36. <https://www.dw.com/en/new-research-online-attacks-on-women-journalists-lead-to-real-world-violence/a-55712872>

37. <https://firstmonday.org/ojs/index.php/fm/article/view/11071/9995>

38. <https://demos.co.uk/project/engendering-hate-the-contours-of-state-aligned-gendered-disinformation-online/>

39. <https://www.tandfonline.com/doi/abs/10.1080/1369118X.2017.1348535>

democratic society is the open pooling of situated knowledge from different groups which is able to benefit from the intellectual diversity of the groups that make up a state, then online harassment is a serious threat to that ideal. Online harassment is not simply a threat to the individuals it targets, it undermines the collective workings of democratic society.

### 3.2 Commercialisation of the public sphere

Why have social media companies allowed ignorance and harassment to spread so freely across their platforms? At least part of the reason is that social media companies have designed their sites around profitability, rather than the public good.

In order to gather as much valuable behavioural data as possible, Facebook has encouraged people to connect with large numbers of other users (van Dijck, 2013), creating dense informational networks which are vulnerable to misinformation and propaganda. When WhatsApp took measures to combat misinformation, one of the first changes was to limit the number of other users which one could forward a message to, creating a curb to the spread of viral information.<sup>39</sup> Facebook's newsfeed is also designed to maximise engagement, which many commentators have suggested creates a system which is designed to amplify sensational false stories which will generate high engagement.<sup>40</sup> Many writers have connected concerns about polarisation and ideological sorting of users on social media to algorithmic filtering of the newsfeed, suggesting that a newsfeed algorithm focused on maximising engagement will end up only showing users what it thinks they want to see (Sunstein, 2017).

Relatedly, the content moderation systems which social media companies present as a solution to hate speech and other kinds of problematic information have operated under what Sarah Roberts calls a *logic of opacity*<sup>41</sup> which hides the existence of a large and extremely exploited<sup>42</sup> workforce of humans behind (false) claims about algorithmic moderation, and obscures the actual rules which govern what is moderated (leading users to construct 'folk' theories about what the rules are)(Roberts, 2019).<sup>43</sup> This logic of opacity extends to the function of content moderation, which Roberts argues is focused on brand management and advertising revenue, rather than ameliorating harms.

A well-functioning market would allow for the emergence of competitors that might be able to challenge the currently existing extractive and anti-democratic social media platforms.

---

39. <https://www.theverge.com/2019/1/21/18191455/whatsapp-forwarding-limit-five-messages-misinformation-battle>

40. <https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation/>

41. <https://firstmonday.org/ojs/index.php/fm/article/view/8283/6649>

42. <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>

43. <https://journals.sagepub.com/doi/10.1177/1461444818773059>

However, as Nick Srnicek argues in *Platform Capitalism* (Srnicek, 2016), a combination of network effects (the benefits that social media companies get from having large numbers of users), well-financed companies being able to buy out competitors (this is particular tendency in the late 2010s, due to historically low interest rates), and the sheer size of the technology companies that own digital platforms means that it is difficult for competitors to emerge. The monopolistic nature of parts of the digital economy has been recognised for a while, and at the time of writing, both the EU and US are in the process of putting through antitrust legislation to break up technology companies.<sup>44</sup>

It is important to remember that social media companies are—despite their claims about providing social goods—after all companies, and one big take away message which one might want to draw from their troubles is that the commercial model for social media is a source of a great many of its problems. A well-functioning public sphere for discourse might have to be designed and run on non-commercial principles.

### 3.4 Solutions

The range of problems around the derailing of collective inquiry, harassment, and the commercialisation of the public sphere are complicated, and interact in complex ways. It is important not to shut down imagination and experimentation in thinking about ways in which social media and other internet sites might be redesigned to avoid these problems.<sup>45</sup> It is also important to experiment with different ways of running social media, and consider the already existing diversity of different social media sites.<sup>46</sup> Nonetheless, there are a number of policy proposals which are plausible candidates for addressing these problems.

There are a number of interventions which could plausibly limit the power of knowledge-undermining communication. The majority of these interventions are changes to the design or rules that govern social media sites, and they might be implemented voluntarily, or by legislation.

1. **Media education:** Starting with the users of social media platforms, many commentators have proposed greater media education. This kind of program might include critical thinking training, some basic statistics, the principles of media reporting, and some signs that an article is likely to be false or misleading. Although this kind of intervention might be helpful, it will

---

44. <https://www.reuters.com/technology/eu-hit-apple-with-antitrust-charge-this-week-source-2021-04-27/>  
<https://www.ft.com/content/4be47818-e889-4442-a009-1d1adda25b0d>

45. <https://www.nytimes.com/interactive/2019/11/30/opinion/social-media-future.html>

46. For a series of essays exploring this theme, see <https://knightcolumbia.org/authors/ethan-zuckerman>

probably be of limited effectiveness unless the media landscape also improves.

2. **Watermarks on media:** One of the strategies that derailers have used to great effect is to run sites that are difficult to distinguish from legitimate news outlets, sometimes going as far as mimicking the layout of specific news sites.<sup>47</sup> One way in which this specific strategy might be addressed would be to establish a system of badges or watermarks for news sources which comply with minimal journalistic standards. This system could also be extended to distinguish between independent reporting and reporting coming from news agencies, and to distinguish between reporting and opinion pieces. This is a fairly limited strategy, as badge-less media may gain a countercultural cachet, and established media often gets things wrong, despite purporting to live up to journalistic standards. It is also important to highlight the fact that citizen journalism and alternative journalism is an extremely important part of the media landscape which would be left out by these proposals.
3. **Slowing down content:** One of the structural reasons why sensational falsehoods have spread so quickly and far on social media is that most sites are designed to amplify content which gets high engagement, so that more and more people see it, creating a feedback loop. One way to change this situation would be to put brakes on virality, for example by setting a limit on how many people a post can be shown to in a given period, or by decreasing the weight which is assigned to engagement in newsfeed algorithms. This intervention would have a mixed effect, limiting the amplification of animal videos and important and surprising truths, whilst at the same time decreasing the audience for misinformation.
4. **Consequences for derailing inquiry:** One of the consequences of framing the function of social media in terms of open and diverse collective inquiry is that we have a clearer and more systematic basis for excluding people. As things stand, social media community guidelines, and their procedures for locking and banning accounts are hidden behind same the logic of opacity that obfuscates content moderation. Regulation of social media could require that sites 1) make the rule for exclusion clear, and 2) include grounds for exclusion on the basis of systematically derailing collective inquiry, for example by

---

47. <https://www.theguardian.com/technology/2017/aug/18/experts-sound-alarm-over-news-websites-fake-news-twins>

spreading misleading information, by discrediting minority groups, or by spreading racist misinformation.

5. **Dealing with political propaganda:** Social media sites have tried to maintain a veneer of political neutrality in their community guidelines, meaning that they have either refused to regulate political speech (as with their refusal to ban Donald Trump's account until the 2021 January insurrection),<sup>48</sup> or have restricted *all* political speech (as with Facebook and Twitter's bans on all political adverts in the 2020 campaign).<sup>49</sup> The refusal to regulate political speech has meant that Facebook in particular has been complicit in significant political violence, internationally, especially in Myanmar, where Facebook allowed government posts attacking Rohingya people.<sup>50</sup> Once we recognise that political propaganda is an enormous threat to the intellectual life of democracy, this neutral position is not tenable. One way to address this would be for social media sites to include rules against political propaganda in their community guidelines, which would be implemented by an independent oversight body.

There are also a number of actions which can be taken to address widespread online harassment:

1. **Informed and transparent content moderation:** Social media sites have for the most part tried to devalue content moderation work, hiding it behind claims about algorithmic systems, and opaque community guidelines. The extent of online harassment suggests that social media companies need to rethink the value and functions of content moderation. Rather than treating content moderation as a PR exercise, it is important to recognise that protecting people online is a complex task, which requires skilled workers. Only by creating transparent guidelines about online harms, and employing content moderators on good contracts to implement these guidelines do social media companies have a chance of addressing the extent of online harassment.

---

48. <https://www.nytimes.com/2021/01/07/technology/facebook-trump-ban.html>

49. <https://www.theguardian.com/technology/2020/oct/07/facebook-stop-political-ads-policy-3-november>  
<https://www.bbc.co.uk/news/world-us-canada-50243306>

50. <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>

<https://www.nytimes.com/2018/11/06/technology/myanmar-facebook.html>

<https://www.theguardian.com/technology/2021/apr/12/facebook-fake-engagement-whistleblower-sophie-zhang>

2. **Implementation of hate speech legislation in online spaces:** Most countries have legislation against hate speech, and there have been successful prosecutions of people for online hate speech. However, prosecution remains somewhat patchy, and it might have a salutary effect on online spaces if countries systematically tried to implant existing legislation on their citizens, or perhaps set up international legislation governing hate speech.
3. **Different private and semi-private spaces for marginalised groups:** one underappreciated solution to online harassment of members of marginalised groups is to build private and semi-private spaces for marginalised groups to avoid the harassment which they might attract elsewhere.<sup>51</sup> These kinds of spaces can provide respite from a hostile society, and can be the basis for effective online counterpublics (Fraser, 1990), which allow marginalised people to develop the intellectual resources to represent minority needs in the main public sphere. These counterpublic spaces can be an effective counterweight to public spheres which are inhospitable to minority groups. There are several models which are worth pursuing, including professionally moderated spaces, community moderation, anonymous and pseudonymous communities.

Finally, if we want to trace the problems of currently existing social media sites back to their financial model,<sup>52</sup> there are several alternatives which we might want to consider:

1. **Public ownership:** if social media sites ought to be pursuing social goods—collective knowledge, democracy, and friendship—we might think that their financial model should reflect this fact. One proposal that would get us in this direction would be public ownership for social media sites.<sup>53</sup> Although this proposal can sound dystopian if we think about a government controlled version of Facebook, the reasonable<sup>54</sup> versions<sup>55</sup> of this proposal<sup>56</sup> put forward a slimmed-down social media site providing limited services, which is publicly accountable and has clear social goals. It is worth remembering that early

---

51. <https://demos.co.uk/project/a-room-of-ones-own-a-guide-to-private-spaces-online/>

52. <https://www.common-wealth.co.uk/reports/common-platform-tech-utility-antitrust#chapter-1>

53. <https://knightcolumbia.org/content/the-case-for-digital-public-infrastructure>

54. <https://www.theguardian.com/politics/2018/aug/23/corbyn-proposes-public-facebook-as-part-of-media-overhaul>

55. <https://newpublic.org/signals>

56. <https://logicmag.io/scale/the-data-is-ours/>

competitors to the US-based internet in France (minitel)<sup>57</sup> and Chile (Cybersyn)<sup>58</sup>, were state-run.

2. **Platform co-operatives:** an alternative financial model, which has been proposed as an alternative to gig working companies like Uber and Deliveroo is platform co-operatives, which would be owned by their workers on a co-operative model.<sup>59</sup> If we think that it is important to secure a distinction between the state and social media, or if social media companies are to maintain significant financial interests (such as advertising, or shopping functions), this model might be worth considering. Perhaps an alternative to Instagram could be owned by models and influencers, and an alternative to YouTube could be owned by musicians and people making instructional videos.
3. **Non-profit:** at the time of writing, the only non-profit website in the top 50 most-visited sites is Wikipedia,<sup>60</sup> which is owned by the Wikimedia foundation. Although Wikipedia had a reputation for unreliability in the early 2000s, it has established itself as a reliable (and mind-blowingly broad) source of factual information which is comparably reliable to the Encyclopaedia Britannica (Giles, 2005). Just as Wikipedia ownership by a foundation enables it to pursue a public good (providing accurate information), a social media site might be better able to pursue social goods (productive democratic discourse) if it was owned by a foundation.
4. **Extended bottom line:** One proposal which has been made to make companies responsible for non-financial goods and ills is the device of the extended bottom line, which requires companies to submit accounts for social goods along with their financial returns. For example, a company might have to submit a report about its positive and negative social impact, and about its carbon emissions, alongside its financial returns. Although this is a more conservative approach, one might want to consider whether social media companies should be obligated to submit a report about the prevalence of harmful false information and harassment on their site.

---

57. <https://spectrum.ieee.org/tech-history/cyberspace/minitel-the-online-world-france-built-before-the-web>

58. <https://www.jacobinmag.com/2015/04/allende-chile-beer-medina-cybersyn/>

59. <https://www.nesta.org.uk/report/platform-co-operatives/>

60. [https://en.wikipedia.org/wiki/List\\_of\\_most\\_popular\\_websites](https://en.wikipedia.org/wiki/List_of_most_popular_websites)

## 4. Journalism and the Pursuit of Truth

Facebook's internal motto in its early years was *move fast and break things*. In the 2010s this slogan read as countercultural and radical. From the perspective of the early 2020s it has a rather different meaning. Among the many things which Facebook—and the wider technology sector—have broken is the news media. Declining trust in established news sources,<sup>61</sup> the move of advertising from print and broadcast to digital removing a key source of revenue for media companies,<sup>62</sup> and political polarisation around the news agenda<sup>63</sup> have created an extremely difficult situation for media companies. The closure of local newspapers,<sup>64</sup> layoffs and increasing precarity amongst professional journalists,<sup>65</sup> and increasing consolidation of media ownership<sup>66</sup> are symptoms of this inhospitable media environment.

Why should we care about journalism given our interest in truth?

From an individual perspective, news media is a purveyor of informational goods, providing individuals with knowledge about the subjects that they want to know about. From this perspective, the closure of a local newspaper is bad because it will frustrate individuals' desires to know what's going on in their area. In cases where there is disagreement about factual issues, journalism is also an important source of expert opinion. Citizens in a democracy are much of the time in the position of novices, facing competing experts, with few tools to discriminate (Goldman, 2001). In this situation, there are a number of tools that novices can rely on: they can try to understand the argument presented by different putative experts, consider the views of other experts in the domain, look for evidence of competing interests, they can look at the track-records of predictions, and they can look to meta-experts who are able to determine expertise in a range of domains. Journalists are sometimes experts in the domains they report in, but even when they aren't, they are able to gather expert opinion and report on the basis of others' expertise. This kind of reporting will not always give simple answers - often scientific and medical reporting involves reporting on a disagreement in which both sides have legitimate reasons. When they have domain expertise, journalists function as experts and their employment by newspapers functions as a strong (although not infallible) indicator of their expertise. When journalists are reporting on the basis of others'

---

61. <https://reutersinstitute.politics.ox.ac.uk/what-we-think-we-know-and-what-we-want-know-perspectives-trust-news-changing-world>

62. <https://www.pressgazette.co.uk/report-predicts-five-years-of-steep-global-decline-for-newspaper-industry-revenue-print-and-online/>

[https://www.iab.com/wp-content/uploads/2020/05/FY19-IAB-Internet-Ad-Revenue-Report\\_Final.pdf](https://www.iab.com/wp-content/uploads/2020/05/FY19-IAB-Internet-Ad-Revenue-Report_Final.pdf)

63. <https://www.journalism.org/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/>

64. <https://www.usnewsdeserts.com/reports/expanding-news-desert/loss-of-local-news/>

65. [https://www.cjr.org/business\\_of\\_news/five-findings.php](https://www.cjr.org/business_of_news/five-findings.php)

66. <https://rm.coe.int/media-ownership-market-realities-and-regulatory-responses/168078996c>



expertise, they are functioning as meta-experts (i.e. experts in identifying experts) and it is their distinctive skill as journalists which they are relying on.

Journalism is also crucial to the collective workings of democracy. In modern democracies, it is simply not possible to understand the range of public perspectives or keep up with relevant political issues by means of face to face communication—there’s just too much going on. So, we rely on news media to provide us with a filter on current events and the opinions of our fellow citizens which can help pick our relevant and reliable information, and give a platform to a diverse range of opinions about issues of common concern (Habermas, 2006). Seen in this way, a free, open, and diverse press is a crucial component of a well-functioning democracy that is capable of generating knowledge about what is going on and what to do (Anderson, 2006). This is especially important when governments and other powerful parties are undertaking socially damaging actions in private. Journalism is an important part of the systems which democratic societies use to hold politicians, and the powerful to account, and journalists may require special legal protections in order to effectively perform this function.

For proponents of the ‘post-truth’ narrative, the problems with traditional journalism is an important part of their narrative. Reading books on post-truth by Davis, D’Ancona, or Ball one gets the impression that until the 2000s, print media was a bulwark of disinterested and noble reliability, until social media came along and redirected public trust towards an algorithmically generated public sphere, which was rife with misinformation and sites pretending to be ‘real’ news. This is supposed to be a great disaster, since digital media (in the form of blogs, websites, and social media) lacks the traditional filtering processes of editors, and ends up being parasitic on traditional media (newspapers, national broadcasters, and state institutions). The preferred solution of these thinkers is to shore up traditional news media, either by establishing the value of existing newspapers and broadcast news, setting up news media outlets with traditional journalistic values,<sup>67</sup> or by trying to re-establish public trust in the news.

Although there is a lot to be lauded in traditional print and broadcast media, we should be careful with any crisis narrative that tries to recreate the media landscape of the 1990s. There were and are significant problems with news media that prevent it from playing the filtering role which it ought to be playing in democracy (Habermas, 2006). Failures to provide diverse media representing different groups, a lack of high-quality local news,<sup>68</sup> consolidation of media ownership,<sup>69</sup> sensationalism, and failures of accurate reporting were all problems long

---

67. For example, D’Ancona is the editor of *tortoise*, which positions itself as something like a membership version of a traditional newspaper. <https://www.tortoisemedia.com/>

68. <https://www.nesta.org.uk/report/mapping-news-media-and-journalism-landscape-research-report/>

69. [https://www.mediareform.org.uk/wp-content/uploads/2021/03/Who-Owns-the-UK-Media\\_final2.pdf](https://www.mediareform.org.uk/wp-content/uploads/2021/03/Who-Owns-the-UK-Media_final2.pdf)  
<https://rm.coe.int/media-ownership-market-realities-and-regulatory-responses/168078996c>

before the advent of social media. Although levels of trust in media are low across Europe (in 2018 a survey by the European Broadcast Union found that 47% of respondents tend to trust print media),<sup>70</sup> we might see this as more indicative of a failure of media to be *trustworthy* rather than a failure of citizens to trust reliable institutions.<sup>71</sup> The conflict between digital and established media is not a fight between diverse and reliable established media and a consolidated technology sector providing unreliable information (Fenton & Freedman, 2017). Both sides are consolidated and have their distinctive intellectual failings.

I want to group issues with news media provision into three areas:

1. Problems with news provision and funding;
2. Consumer-side issues about distinguishing between genuine journalism and counterfeit journalism;
3. Concerns about threats to journalism and the safety of journalists.

My primary focus here will be on the European context—the EU27 and the UK—but many of the same issues are important across the globe. There are a number of ways that one can divide up media providers (for example, by medium, or purpose), but I will focus on ownership structures, distinguishing between publicly-owned media (France Radio, Sveriges Television, the BBC), privately owned media (News Corp, EGS), and media owned by foundations (the Guardian, co-operative owned newspapers), and distinguishing between traditional media (newspapers, public radio, broadcast news), and digital media companies (GalDem, BuzzFeed, Huffington Post). The issues discussed will not be identical in every European country (for example, funding is less of an issue in Sweden due to state subsidies), but they give a shape for thinking about the challenges for 21<sup>st</sup> century journalism. Rather than diving into comparisons between countries, I will rely on examples to illustrate the problems discussed.

#### **4.1 Problems with Media Companies**

Commercial media faces a number of challenges around funding, diversity in news provision, and media consolidation which threaten their ability to play their role in providing trustworthy information, and filtering issues for public discussion.

---

70. [https://cor.europa.eu/en/events/Documents/Europcom/David\\_Fern%C3%A1ndez\\_Quijada\\_EuroPCom.pdf](https://cor.europa.eu/en/events/Documents/Europcom/David_Fern%C3%A1ndez_Quijada_EuroPCom.pdf)  
[https://www.slideshare.net/Edelman\\_UK/edelman-trust-barometer-2018-uk-results/1](https://www.slideshare.net/Edelman_UK/edelman-trust-barometer-2018-uk-results/1)

71. This impression is backed up by higher levels of trust for public service media in the same study.

The move in advertising revenues from traditional media companies across to technology companies has contributed to loss in revenues, along with falling print circulation,<sup>72</sup> and the difficulties in developing alternative financial models. The financial crisis created by this situation has led to layoffs at many companies (compounded by the COVID-19 pandemic),<sup>73</sup> and an increased pressure to blur the boundaries between reporting and entertainment. The core function of the media is to provide public informational goods, and there is a deep tension between this aspiration and the profit-focused model which is the basis of the commercial media. Although there may have been a period in the late 20<sup>th</sup> century in which the availability of advertising revenue covered this tension, we appear to be entering another period in which the commercial model of media companies undermines their function for democracy.

Traditional media companies are based on a model of mass communication in which they seek out a mass audience. This means that they tend to cater for majority social groups, leaving minority groups underserved. There are specialist news outlets serving the intellectual and political interests of minority groups (for example, LGBT+ newspapers, websites aimed at women and people of colour, and local news sites and newspapers), but the smaller audiences for these outlets means that they face significant financial challenges. Although it is possible for large companies to provide some minority-interest journalism, as a profession journalism remains unrepresentative. A report on journalism in the UK from the National Council for the Training of Journalists based on information from in 2018-2020 found that in 2020, 92% of journalists are white (compared to 88% of the population), and that 75% of journalists are from the three highest parental occupations (a proxy for social class), compared to 45% of the population).<sup>74</sup> Lack of diversity in media provision is important for a couple of reasons: it undermines the epistemic diversity which is important for the success of democratic knowledge-aggregation systems, and it compounds the marginalisation of minority groups in public discourse.

Consolidation of media across Europe is also a serious challenge. Media consolidation refers to the ownership of commercial media by a small number of companies. This phenomenon underlines the way in which the financial structure of commercial media can affect the way it performs its democratic functions when the market becomes consolidated. A study from 2015 finding that most European countries are at high risk from media consolidation.<sup>75</sup> While the UK is particularly stark example (with 90% of the print circulation of newspapers being

---

72. <https://www.pressgazette.co.uk/uk-national-newspaper-sales-slump-by-two-thirds-in-20-years-amid-digital-disruption/>

73. [https://www.cjr.org/business\\_of\\_news/five-findings.php](https://www.cjr.org/business_of_news/five-findings.php)

74. <https://www.nctj.com/downloadlibrary/Diversity%20in%20journalism%202021.pdf>

75. <https://rm.coe.int/media-ownership-market-realities-and-regulatory-responses/168078996c>

controlled by four companies),<sup>76</sup> media consolidation is an issue in many European countries. For example in Italy, Silvio Berlusconi (the former prime minister) is the major shareholder of the biggest commercial television company (Mediaset), the biggest publisher (Mondadori), and the biggest advertising company (Pubitaliana). Two of Italy's biggest daily newspapers are owned by his brother (Il Giornale), and (until 2015) by his ex-wife (Il Foglio) (Richeri & Prario, 2016). Media consolidation is a major issue insofar as it undermines the ability of journalism to hold power to account, increases the power of the owners of media companies over public discourse, and reduces the diversity of public discussion. Returning to our discussion of the role of markets in the pursuit of knowledge, it should be pretty clear that media consolidation is an example of a case in which a market-based system needs to be regulated in order to provide the goods which the system functions to provide.

#### 4.2 Issues for News Consumers

News consumers face an increasingly heterogeneous media landscape which is presented in a strikingly homogenous digital format. This flattening out of presentation is part of the reason why so-called 'fake news' sites were able to fool so many users—the indicators of an established news provider are fairly easy to counterfeit on a website, whereas it would take considerable effort to fake a physical newspaper.

There has been a vigorous debate about what the term 'fake news' means (and another debate about what it *ought* to mean). One of the key lessons from this debate is that there is a sense of 'fake news' connecting to counterfeit news stories which picks out an important problem (Pepp, Michaelson, & Sterken, 2019; Fallis & Matheson, 2019). A *counterfeit* news story is a story which has not been through standard journalistic processes being presented and circulated as if it had been through these processes, giving it an undeserved boost to credibility. Much of the concern about the 2016 US election which we discussed in 3.1 involved counterfeit stories which were false or misleading, but achieved wide circulation in part by mimicking the format and names of established sites.

In a context in which news is provided by a mix of large traditional media companies, smaller digital-only companies, and increasingly single journalists operating newsletters,<sup>77</sup> it will be important to understand the different journalistic processes which lie behind news stories. This is not to say that there is something intrinsically better about a large broadcast organisation compared to a single journalist: these operations have different virtues. A single

---

76. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/720400/180621\\_Mediatique\\_-\\_Overview\\_of\\_recent\\_dynamics\\_in\\_the\\_UK\\_press\\_market\\_-\\_Report\\_for\\_DCMS.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/720400/180621_Mediatique_-_Overview_of_recent_dynamics_in_the_UK_press_market_-_Report_for_DCMS.pdf)

77. <https://www.newyorker.com/magazine/2021/01/04/is-substack-the-media-future-we-want>

journalist may be able to cover a beat which is too specialist for a large broadcaster, but the large broadcaster will have more editorial support, a broader base of expertise, and better legal support which allows them to take on complex stories, and undertake investigative journalism. It is important to remember that not all large media companies follow proper journalistic processes—as well as tabloids that sensationalise news stories, there is a history of newspapers persistently printing false stories (in the UK, the Sun’s reporting of the Hillborough disaster is a notable example), and failing to print important truths (newspaper reporting of the climate disaster providing a very high-stakes example).

I suggest that the rise of counterfeit news, the existence of unreliable major news outlets, and increasing numbers of small digital sites and journalists working independently means that it is increasingly difficult for citizens to know how much credence to give to news stories. The changes in publication and communications technologies has meant that the media environment has become so complex that traditional indicators (like writing for a major newspaper—again, not an infallible indicator) do not always help us to distinguish expert from non-expert journalists.

What kinds of solutions might we want to think about to address this problem?

Consider an analogy with medicine. Medics provide individual and social goods in the form of individual and public health. In some cases these goods can be effectively produced by market mechanisms (people will be willing to pay to have their broken bones fixed), but in other cases market incentives undermine the provision of health (it might be more profitable to recommend a costly drug for a chronic illness rather than carrying out a one off operation which has better outcomes). Although the majority of medics will be competent and well-intentioned, an incompetent or malicious medic may be able to do a considerable amount of harm to their patients. Crucially, the success of medics in producing the good of health relies on them being trusted by their patients. Setting to one side the placebo effect, effective medical treatments (a course of antibiotics, rest after an operation, a change of diet) will only improve a patient’s health if they trust their physician enough to follow their advice. Because of the importance of trust to medical treatment, and the dangers of incompetent and malicious medical practitioners, in almost every country in the world medical professionals are registered or licensed by a professional body in order to practice medicine. In the UK, this body is the general medical council (GMC), which tests applicants (including medics who are registered in other countries), deals with some cases of medical misconduct, and sets the standards for medical education.

Although there are important differences between medicine and journalism, the reasons for professional accreditation for medics have pretty clear analogies for journalism. Journalism produces individual and social goods, in the form of individual knowledge, and contribution

to a democratic public discourse, and in some cases market incentives may undermine these goods (a newspaper might sell more papers if it prints a sensationalised false story than a boring true story). Although the majority of journalists are competent and well-intentioned, mistakes and malice can cause significant harm. If a newspaper prints details of criminal allegations which turns out to be false, this could easily ruin the accused's life. And just as the medical profession requires trust to perform its role, so does journalism require trust in stories to perform its role in democratic discourse.

If we take this analogy seriously, it suggests that there may be an important role for professional accreditation bodies for journalists. A professional registration body would need to be independent of the state to allow journalists to hold power to account, would require some kind of test to be accredited, and would have the power to remove the accreditation of journalists who carry out journalistic misconduct or do consistently bad reporting. Although we might not want to legally require accreditation to practice as a journalist (it might not be practical for citizen journalists to be accredited), a practice of accrediting and 'striking off' journalists would both give readers a strong indicator of which journalists are reliable, and establish a social structure which would hold journalists accountable for their reporting. There are some existing accreditation schemes: in the UK, the National Union of Journalists<sup>78</sup> and the chartered institute of journalists<sup>79</sup> have schemes to accredit journalists with press cards, and the international federation of journalists does the same thing internationally.<sup>80</sup> In Portugal it is a legal condition of working as a journalist that one has a licence from the CCPJ,<sup>81</sup> and in France professional journalists are required to get a press card issued by the CCJIP which comes along with various conditions, including not taking certain other forms of work.<sup>82</sup> However, none of these schemes are as central to journalistic practice as medical licensing is to medicine: accreditation is not a condition for working as a journalist (except in Portugal), and there does not seem to be a practice in which the body removes registration from journalists who have flouted professional norms. Lani Watson discusses a related idea in her book *The Right to Know*, suggesting that some of the problems with contemporary media might be addressed by something like a Hippocratic oath for journalists (Watson, 2021, pp. 102-103).

This scheme for accrediting journalists could be combined with the watermarks system discussed in section 3.4, with accredited journalists receiving something like Twitter's blue tick next to their byline, and with accredited newspapers having some similar indicator of

---

78. <https://www.nuj.org.uk/>

79. <https://cioj.org/press-cards/the-national-press-card/>

80. <https://www.ifj.org/press-card.html>

81. <https://www.ccpj.pt/en/professional-licenses/professional-journalist-license/>

82. <http://www.ccijp.net/rubrique-2-la-carte-de-presse.html>

reliability. This would allow the accreditation scheme to play a public role in assuring public trust besides its behind the scenes role in motivating journalists to behave in a trustworthy manner.

### 4.3 Journalistic Independence and Safety

In their round-up of 2020, Reporters Without Borders (RSF) document 387 journalists who had been detained, 54 who have been held hostage, and 50 who have been killed in connection with their work.<sup>83</sup> These figures demonstrate the risks which are involved in on-the-ground reporting. Of the 50 journalists who were killed, 84% were deliberately targeted, and 68% were killed in countries which were (at least nominally) at peace. In Europe, RSF highlight Hungary's undermining of a free press, and the failure of prosecutors in Slovakia and Malta to successfully prosecute anyone for the murders of Ján Kuciak and Daphne Caruana Galizia.<sup>84</sup> The issue of journalistic safety has been a particular issue during the COVID-19 pandemic, and there are several cases in which European governments have arrested reporters seemingly because of their coverage of the pandemic.<sup>85</sup> Safety is an important issue for all journalists, but it is an especially salient issue for women working in journalism,<sup>86</sup> and for other minority groups who are at heightened risk of violence and abuse.

There are two ways in which we can think about protections for journalists. The first is to ground a right to investigation in the right to freedom of expression. If people have a right to hold and to express views without fear of threats or violence which gives them rights to protection, then we might think that there is a similar right to inquire and gather information about topics with a view to disseminating that information without fear of threats or violence. The second way to think about the morality of protection for journalists is to argue that independent investigative journalism is an important part of a well-functioning democracy. Anderson's Experimentalist model of democracy presents democracy as the application of collective intelligence to matters of common concern. Although this kind of collective project ought to be inclusive, it will also require a division of labour in which specialised institutions and experts take on certain issues. For example, the scientific establishment in a democratic society takes on the role of investigating scientific issues of public concern, besides investigating topics of intrinsic intellectual interest. Following this analogy, we might think that part of the job of investigative journalists is to take on the specialised role of finding out

---

83. [https://rsf.org/sites/default/files/rsfs\\_2020\\_round-up\\_0.pdf](https://rsf.org/sites/default/files/rsfs_2020_round-up_0.pdf)  
[https://rsf.org/sites/default/files/bilan\\_2020\\_en-tues\\_.pdf](https://rsf.org/sites/default/files/bilan_2020_en-tues_.pdf)

84. <https://rsf.org/en/2020-rsf-index-europes-journalists-face-growing-dangers>

85. <https://rsf.org/en/news/repressive-laws-prosecutions-attacks-europe-fails-shield-its-journalists-against-abuse-covid-19>  
<https://t.co/IW57wBM3tQ?amp=1>

86. <https://www.mediasupport.org/publication/the-safety-of-women-journalists/>

about issues that the government and powerful institutions are trying to keep out of the public eye. If journalists are to perform this role successfully, then they will need legal, and possibly physical, protections from the parties which they are investigating.

The issue of the safety of journalists has been repeatedly raised at the UN, and in October 2020, the UN Human Rights Council passed a resolution calling for governments to implement robust domestic legislation to protect journalists.<sup>87</sup> Reporters Without Borders are also calling for a special representative to the General Secretary of the United Nations on the safety of journalists.<sup>88</sup>

Given the fact that journalists often work across borders and the sometimes antagonistic relationship between the state and journalists, one might want to consider whether it is possible to deal with the threats to journalistic safety through international legislation or action through the UN. There are at least three possible routes this might take:

1. Limited forms of immunity from prosecution for journalists;
2. The foundation of an international body for regulating national journalism;
3. The establishment of a special observer or special representative on journalism.

#### **4.3.1 Journalistic immunity**

Just as we drew an analogy between journalism and the medical profession in 4.2, we might think that the threats to journalists (especially working outside of their country of citizenship) can be dealt with by drawing an analogy between journalists and diplomats. There is an extremely long-standing norm that diplomats ought not to be punished under local laws, which has been formalised under international legislation in the Vienna Convention on Diplomatic Relations in 1961. This legislation gives diplomats (and some of their families and associates) immunity from prosecution or punishment except if their country of origins waives this right. It is an attractive thought that journalists might similarly be protected from illegitimate persecution and prosecution by granting journalists something akin to diplomatic immunity.

This issue is more complex than it first seems. First, a good deal of the threats to journalists come from non-state actors who may not be deterred by international law. A militia group

---

87. [https://ap.ohchr.org/documents/dpage\\_e.aspx?si=A/HRC/45/L.42/Rev.1](https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/45/L.42/Rev.1)

88. <https://rsf.org/en/protector-journalists-1>



might well not be deterred from killing journalists by international legislation in the same way that a state would be.

Secondly, an unrestricted form of legal immunity for journalists would be untenable. One issue that has arisen repeatedly with diplomatic immunity is what should be done when someone claiming diplomatic immunity commits a serious crime. This issue has recently emerged in the UK with the case of Harry Dunn, who was killed in a motoring accident involving Anne Sacoolas, the wife of an American diplomat. Despite admitting to driving on the wrong side of the road, Sacoolas claimed diplomatic immunity, and at the time of writing there is an ongoing international negotiation about whether she can be prosecuted.<sup>89</sup> It would be bad if journalists had a blanket immunity from being prosecuted for any crimes, including murder and other serious crimes. It might be better to allow journalists some kind of limited liability relating to their professional work, but this would have to come with caveats. It is important for the protection of citizens from harm from malicious journalists that journalists can be prosecuted for defamation or hate speech. So even protection in the course of professional activities would have to be restricted.

Perhaps there might be better prospects for legislation to establish special legal protections for journalists (rather than excepting journalists from punishment). For example, we might want to explore the possibility of establishing special rights against unjust imprisonment of journalists under international law, or a special crime of attacking a journalist with harsher punishments or a special prosecuting body.

#### **4.3.2 International regulatory body**

A rather different strategy would be to propose international regulatory bodies for journalism, on the model of the EU Atomic Energy commission, or the International Telecommunication Union. Article 19 of the Universal Declaration of Human Rights states that:

Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.<sup>90</sup>

This article establishes both a right to hold and express opinions, and a corresponding right to receive information and ideas from anywhere. Both rights can be undermined by a poorly functioning press. If we have a right to reliable information—a right to *know* things (Watson, 2021)—then a biased or unreliable press will undermine our rights. And a press which operates in an environment where journalists are threatened or are unable to express their opinions

---

89. <https://www.bbc.com/news/world-us-canada-56246511>

90. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>

due to the financial environment is one that undermines freedom to express opinions. In 4.2 we discussed the possibility that non-state professional bodies might play an important role in regulating journalism at a national level. An international regulatory body for journalism might take on a triple role, ensuring that professional bodies support good journalism, addressing issues about journalists working internationally, and preventing national governments from stifling journalists.

#### 4.3.3 Special Representative at the UN

Reporters Without Borders' proposal for a special representative on the safety of journalists provides another way to heighten awareness and co-ordinated action at the UN around this issue. The role of special representatives to the UN Secretary General (which should not be confused with the wider category of special rapporteurs appointed under special procedure mechanisms) is to report and investigate an issue of concern and to represent the UN Secretary General on human rights issues. Thinking about the safety of journalists as a human rights issue, we might propose a special role to represent this issue at the UN level.

#### 4.4 Solutions

In this section we have covered a number of problems which prevent the news media from playing their proper role in democratic society. We can briefly outline the proposals that we have canvassed.

Solutions for media organisations:

1. **Public funding for reliable journalism:** we have seen that lack of funding for journalism is an important barrier to news media acting as a trustworthy source on a range of democratically important topics. This issue could be addressed by establishing public funding for trustworthy and reliable journalism.
2. **Public funding for minority-interest journalism:** we have also seen that media diversity is important for journalism to play its role of filtering the diverse opinions in society. Minority-interest journalism faces special challenges in an environment in which scale is a precondition for commercial viability, so it may be important to establish special public funding for minority-interest reporting both in special venues and in venues with a general readership.
3. **Legislation addressing media consolidation:** an overly consolidated media environment in which a small group of individuals or companies controls the

media will be less likely to represent a broad range of opinion, and will be correspondingly more likely to put out false information which is in the interest of owners. This issue provides an example of the importance of market regulation in the media, and can be addressed by legislation at the national or super-national level which mandates diverse ownership of media.

Solutions for reader-side issues:

1. **System for certifying journalists and media companies:** the analogy between medical professionals and journalists suggested the possibility of non-state professional bodies to accredit journalists. While we have seen that there might be some concerns about making this accreditation a condition of working as a journalist, empowering existing accreditation schemes with more social and legal significance would help journalists to promote reliable reporting, and provide ways to hold unreliable reporters to account.
2. **Watermarking system for journalists:** one of the important issues with reading online news is the difficulty of understanding how seriously to take the source of this news. It would be interesting to think about extensions of the watermarking system proposed in 3.4 to address this issue by displaying information about the accreditation of the journalist who has written a piece and the news organisation that they work for.

Solutions for safety of journalists:

1. **Journalistic immunity:** the safety of journalists is a complex issue, and it will not be susceptible to simple solutions. Although there is some appeal to the proposal to establish diplomatic immunity, this proposal faces considerable issues in implementation relating to the difficulty of establishing which kinds of prosecution journalists ought to be immune from, and it also fails to address the problem about non-state violence towards journalists.
2. **Registration and protections model:** A more promising avenue is to pursue special protections for journalists who are registered with a professional body. This would allow journalists to have special protections under the law without giving them a problematic blanket immunity to prosecution.
3. **Global regulatory body:** Another policy possibility would be to establish an international regulatory body which could regulate nation-level professional registration bodies, address issues about journalists working internationally,

and take action around government interference with journalistic freedom of expression.

4. **Special Representative to the UN:** these legislative and institutional solutions to the problem of journalistic safety might be unrealistic in the short term, creating the need for more practical interventions. RSF's proposal of a UN special representative working on the safety of journalists is one such proposal, fitting into the established system of special representatives.

## 5. Data Privacy and Institutional Trustworthiness

Our discussions of social media and journalism have focused on issues about the provision of useful and knowledge-yielding information to people. Concerns about privacy have a somewhat different shape: they are not about the way institutions provide information to people, but rather about the way institutions collect, process, and publicise information about people. We can see this issue as an extension of political issues about the trustworthiness of the institutions we live with and within. We rely on institutions not just to provide us with reliable information about the world, but also to responsibly manage and use information about us.

Following Helen Nissenbaum's *contextual integrity* approach to privacy (Nissenbaum, 2010; 2011), I will take it that trustworthy information-management is a matter of an institution collecting, processing, and using information about people in ways that accord with reasonable social norms about information usage. Our social norms for information-use are complex, distinguishing between different *kinds* of information, different *uses for* information, different *people* who might be appropriate targets for information acquisition, and different rules for the transmission of information. One vital point which Nissenbaum makes is that the reasonable social norms for information use are context-relative. When I disclose to my friend that I am having mental health problems, I may expect them to keep that information to themselves, or perhaps to tell only other close friends. If I were to tell a teacher or academic mentor the same thing, our informational norms might require that they keep that information to themselves, except if I am at risk of harm if they don't disclose that information. If I tell my doctor (or another mental health professional) the same information in the context of an appointment, the doctor might be flouting informational norms if she kept that information to herself, depending on the clinical relevance of that information.

An important point about Nissenbaum's approach to thinking about privacy is that appropriate information use is determined by social norms at the collective level, rather than by individuals' expectations about privacy. Our norms about the use of information in clinical

contexts determines the right way for a doctor to process information about patients' mental health, not a patients' individual expectations about how information should be used. A doctor would still be obligated to record information about a patient's mental health, even if the patient expected or even demanded that that information be kept private.

Many approaches to appropriate information-usage are based around consent, with the idea being that individuals have a default right to control information about themselves. Individuals can then choose to give rights to access and use that information when they make formal and informal agreements with others. In online contexts, the consent-based approach is often associated with end-user agreements (EUAs) which present individuals with extremely detailed and complex information about the use of their information, in the expectation that users will agree (or at least click 'agree') to those conditions without reading them.<sup>91</sup> Morally significant consent to these kinds of opaque and lengthy agreements is at best shaky. One might have thought that the problem with EUAs is that they are too long in order to ensure that individuals don't read them. Less verbose technology companies could simply describe their data use in fewer words. Nissenbaum argues that the problem with consent-based approaches to privacy is deeper, identifying what she calls the *transparency paradox* (Nissenbaum, 2010; 2011). Any reasonably complex institution—such as a medical establishment—will have a complex system for the use of information about the people it interacts with. For an agreement to generate legitimate consent to information-use, it must describe this system in a way that is both accurate and clear to non-specialists. This generates a dilemma: short and clear descriptions (for example resembling the nutritional information on a good packet) will fail to accurately describe information usage, whereas accurate and compendious descriptions will be too long and specialist for generalists to be able to use them. Nissenbaum's proposal is to ground legitimate information-usage not in individual acts of (dubious) informed consent, but rather in social norms about information usage by different kinds of institutions.

For this approach, what matters to well-functioning systems of privacy is the establishment of reasonable and socially beneficial systems of information usage. This is not a relativist approach according to which any set of social expectations can determine appropriate usage of information: we ought to assess sets of informational norms by thinking about their effects on individuals and wider society. On this approach privacy is not an intrinsic value which we should care about for its own value, but rather an instrumental value that we should care about insofar as it contributes to the pursuit of other values. Improper use of information can lead to grave harms, including: 1) informational harms (such as a victim being found by a

---

91. Actually reading the EUAs that would be required to use normal internet services is a gargantuan task  
<https://www.theguardian.com/technology/2015/jun/15/i-read-all-the-small-print-on-the-internet>  
<https://www.visualcapitalist.com/terms-of-service-visualizing-the-length-of-internet-agreements/>

stalker, or public shaming based on socially disapproved actions), 2) informational inequalities (such as differential treatment based on race, gender, or disability), 3) the undermining of autonomy through the absence of a private sphere in which we can steady our views and selves, 4) the undermining of close relationships which rely on intimacy and disclosure of information, and 5) the undermining of democratic processes such as voting which presuppose privacy.

Nissenbaum's contextual integrity approach presents us with an account of what a well-functioning and trustworthy institutional system would look like; it would be one governed by reasonable social norms about the collection, maintenance, and use of information. Measured against this ideal, the institutions which currently govern our lives online fall short in a dramatic way. Nissenbaum suggests that technology companies—here we might think particularly of Google's use of cookies, and Facebook's Like button—have developed systems for information use which are governed by technological possibility rather than reasonable social norms. In *The Age of Surveillance Capitalism* Shoshana Zuboff dubs the system which these systems implement *Surveillance Capitalism* (Zuboff, 2019). She argues that this system has transformed personal information into a novel source of extractive value, which can be packaged by and sold on to advertisers, who believe that personalised adverts are a more effective way of selling their goods. We might also add that technology companies also have a substantive financial interest in producing surveillance systems for the state, who are another important customer for personal information. We have ended up in a system characterised by extreme informational inequalities (Brunton & Nissenbaum, 2015): technology companies know a great deal about our lives online, whereas those same companies obfuscate information about their financial dealings, and the way their sites work. For example, most social media sites (Twitter being the notable exception) have EUAs which forbid academics from carrying out research on social media users, meaning that most of the research on social media sites is done internally by companies themselves, and never published.

In this context, it is important to be realistic about the actual consequences of this informational asymmetry. Technology companies having private information about their users opens users up to a range of informational harms, from the public disclosure of a dating account to discrimination based on DNA test results. However, the public discussion of privacy—notably in documentaries like *The Great Hack* and *The Social Dilemma*<sup>92</sup>—presents an extremely misleading picture of what technology companies are actually able to do with

---

92. <https://www.theverge.com/interface/2020/9/16/21437942/social-dilemma-netflix-review-orlowski-sarah-zhang-memo-facebook-buzzfeed>  
<https://librarianshipwreck.wordpress.com/2020/09/17/flamethrowers-and-fire-extinguishers-a-review-of-the-social-dilemma/>

user data. These films present a picture in which the possession of information about a user's psychology or their social media usage, fed into targeted adverts allows them to be extensively manipulated, making voting or purchasing decisions that they would not otherwise have made. While it is true that companies from Cambridge Analytica to Google do present themselves as having this extreme power of manipulation, the empirical evidence about the effectiveness of political manipulation and online advertisements does not support such strong conclusions. These kinds of apocalyptic views about the problems posed by social media and online advertising are of a piece with the 'post-truth' narrative that we described in section 1, in the sense that both paint the problems created by our online lives as creating completely new and distinctive issues. Political manipulation through social media is an important issue, but the form the problem takes is the old one of government propaganda which is not properly regulated by social media sites.<sup>93</sup> Advertising undermining individuals' autonomy is an important issue, but the problem is just the old one that under capitalism, companies have a financial incentive to cultivate consumer desires for their products, no matter whether those products benefit or harm the people that buy them. Some writers have already begun to describe online advertising as a bubble that relies on hyperbolic claims about the effectiveness of targeted advertising.<sup>94</sup>

What would a better online system of privacy look like? Nissenbaum stresses that our online lives are heterogeneous, and that many parts of our offline lives intrude on online spaces. This means that there is no such thing as a general system for online privacy: what we are after is a set of online institutions which are governed by reasonable social norms about information usage for the kinds of activities which they engage in. Online banking ought to be governed by the same norms which govern physical banks. Dating apps ought to be governed by the same norms that govern offline dating and personal adverts. Messaging apps ought to be governed by the same norms which govern everyday face-to-face conversation. There may be institutions that are unprecedented—social media being an important case in point, given the way it merges personal and political conversations—in which case it will not be possible to extend existing norms. In this case, the goal will be to have a public discussion about what the aims of this institution ought to be, and what system of information-usage would further those goals.

---

93. Although Facebook and Twitter took a stricter line with propaganda in the 2020 US election, Facebook's international record on allowing political propaganda is terrible <https://www.theguardian.com/technology/2021/apr/12/facebook-fake-engagement-whistleblower-sophie-zhang>

94. <https://thecorrespondent.com/100/the-new-dot-com-bubble-is-here-its-called-online-advertising/13228924500-22d5fd24>  
[https://www.amazon.co.uk/gp/product/0374538654/ref=ox\\_sc\\_act\\_title\\_2?smid=A3P5ROKL5A1OLE&psc=1](https://www.amazon.co.uk/gp/product/0374538654/ref=ox_sc_act_title_2?smid=A3P5ROKL5A1OLE&psc=1)

Our goal is not simply to reduce the amount of information which is gathered about us by technology companies. We want to build trustworthy institutions which we can rely on to deal with our information in a way that doesn't harm us and achieves socially beneficially outcomes. It is possible that although we will want to reduce the amount of information held by many companies, a reasonable situation might involve *increasing* the amount of information certain institutions (such as national health bodies) hold about us. Achieving this move to trustworthy institutions will require a mix of legislation about data use, contestation about the goals of institutions (including the goals of technology companies), and public debate about specific examples of information usage.

I will discuss three policy areas which provide opportunities to pursue the goal of institutions we can trust to deal with our information.

1. Ameliorating informational harms;
2. Ending the commodification of data;
3. Allowing users to obfuscate their online activity.

We will focus on policy-level interventions, for some discussion of actions which can be taken by individuals, see Chapter 6 of *Privacy is Power* by Carissa Véliz (Véliz, 2020).

## 5.1 Ameliorating Informational Harms

In making the move from the currently existing untrustworthy—and largely untrusted—institutions our first concern ought to be in harm reduction. This means that a focus of legislation ought to be on ensuring that technology companies are no longer able to handle the data of individuals in ways that causes informational harms.

It will be particularly important to ensure that marginalised people are not vulnerable to informational harms, especially since marginalised groups may use the internet in different ways than non-marginalised groups. For example, many trans and non-binary people use pseudonymous public accounts on social media, both to avoid being outed, and to shield themselves from harassment.<sup>95</sup> People who use anonymity to protect themselves from harms occupy a special context with distinctive norms, and in order to avoid serious harms to these groups, it would be beneficial to have informational norms which ensured that the offline identities of people from these groups are not publicly available.<sup>96</sup> Because the design of technology has historically focused on the needs and interests of majority groups, it is

---

95. <https://www.vox.com/culture/21432987/trans-twitter-reddit-online-anonymity>

96. <https://privacyinternational.org/sites/default/files/2018-11/From%20oppression%20to%20liberation-reclaiming%20the%20right%20to%20privacy.pdf>



important to centre and involve marginalised groups in the design process, as Sasha Costanza-Chock argues in their book *Design Justice* (Costanza-Chock, 2018).<sup>97</sup>

The range of informational harms which are being perpetuated at present is huge, and we don't have enough space to give a complete account of them. Some particularly egregious examples of technologies which are likely to cause informational harms: 1) The use of facial recognition technologies by police forces<sup>98</sup>, especially the general unreliability of facial recognition software and its specific unreliability in recognising the faces of people with darker skin<sup>99</sup>, and the historical failures of predictive policing.<sup>100</sup> 2) The use of social classification data in the algorithms which determine who is shown job adverts, which has the consequence that women are not shown suitable job adverts.<sup>101</sup> 3) Credit scoring algorithms which use an applicants' social identity (or proxies for their social identity) to determine whether they should get a loan of mortgage (see Chapter 8 of *Weapons of Math Destruction* (O'Neill, 2016) for a discussion of the discriminatory consequences of these systems).<sup>102</sup> More generally, we might worry that commercial and government surveillance creates a panoptic social system, in which peoples' knowledge that they either are or could be watched causes them to avoid socially sanctioned actions through a process of self-surveillance (Campbell & Carlson, 2002).

## 5.2 Ending the Commodification of Data

An important part of the explanation why currently existing companies that process information about people are untrustworthy is the fact that they have pursued financial models which are conditional on collecting large amounts of data about their users. This is obviously true of companies like Google and Facebook, who sell targeted adverts which are supposed to be effective because of their micro-targeting of particular kinds of users (although we've seen some reasons to be sceptical about this marketing). But it is also true of companies that are not such central participants in the economy of Surveillance Capitalism, such as banks and supermarkets, which collect a huge amount of behavioural data about their customers through financial records and loyalty cards. If we want to have institutions which we can rely on, it is crucial to remove, or at least severely curtail the market in personal data.

---

97. <https://design-justice.pubpub.org/>

98. <https://www.wired.co.uk/article/uk-police-facial-recognition>

<https://www.nytimes.com/2020/06/09/technology/facial-recognition-software.html>

99. <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

100. <https://reallifemag.com/broken-windows-broken-code/>

101. <https://www.technologyreview.com/2021/04/09/1022217/facebook-ad-algorithm-sex-discrimination/>

102. <https://www.theatlantic.com/technology/archive/2016/12/how-algorithms-can-bring-down-minorities-credit-scores/509333/>

At a European level, the *General Data Protection Regulations* introduced in 2018 (GDPR for short) go some way towards de-commodifying personal data.<sup>103</sup> This legislation<sup>104</sup> was introduced by the European Union to harmonise regulation of how companies use individuals' personal information across Europe. This legislation presents seven principles for how individuals' information can be used: lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality (security); and accountability. Fully explaining all of these principles would take us rather far afield, but it is worth highlighting the principle of data minimisation, which states that personal data must be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes."<sup>105</sup> It is also worth noting that GDPR legislation establishes quite serious fines for breaches, and at the time of writing there are currently on-going cases against Google<sup>106</sup> and Facebook<sup>107</sup> with Google being threatened with a fine of \$120 million, and Facebook standing to be fined over \$1.6 billion.

What GDPR legislation establishes in principle is a set of procedural regulations on how information is used, and a regulation system for companies that fail to comply with these regulations. However, it doesn't establish any limitations on the purposes for which information can be collected, besides clarifying that they be 'legitimate'. If we are really invested in de-commercialising the use of personal data, it might be worth considering whether it would be desirable to specify that financial gain is not a legitimate reason to gather information, or require that the purposes for which information is gathered will benefit the people whose information is being gathered.

### 5.3 Protecting Obfuscation

A theme throughout this research overview—and in social epistemology more generally—is that social epistemology ought to be concerned both with how to build better institutions, and how individuals can deal with actually existing flawed institutions. If we can't build companies which are trustworthy guardians of our information, what can we do? One option is opt out, by refusing to use technologies and sites that are particularly hungry for your data. Taken strictly this would mean abstaining from most search engines, from all large social

103. <https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018>

104. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1552662547490&uri=CELEX%3A32016R0679>

105. <https://gdpr-info.eu/art-5-gdpr/>

106. [https://techcrunch.com/2020/12/10/france-fines-google-120m-and-amazon-42m-for-dropping-tracking-cookies-without-consent/?guccounter=1&guce\\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce\\_referrer\\_sig=AQAAAKFENhPvw0eqHC4Xsi-lwXFA1WZxHPjBgnXvK4rqx54nHd7OZK3IFLLS\\_20HOuXnFJmGNLyFFouXdE6f4KfJn62WjrmPk-UlcXlMFwg-ElQ1B\\_nqPCC3rgcBv5XgyfezOP4GP5K6mPwcb0GC3LCLn-pQB0yHtfQ0xU\\_xwTfqsWIC](https://techcrunch.com/2020/12/10/france-fines-google-120m-and-amazon-42m-for-dropping-tracking-cookies-without-consent/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAKFENhPvw0eqHC4Xsi-lwXFA1WZxHPjBgnXvK4rqx54nHd7OZK3IFLLS_20HOuXnFJmGNLyFFouXdE6f4KfJn62WjrmPk-UlcXlMFwg-ElQ1B_nqPCC3rgcBv5XgyfezOP4GP5K6mPwcb0GC3LCLn-pQB0yHtfQ0xU_xwTfqsWIC)

107. <https://gdpr.eu/the-gdpr-meets-its-first-challenge-facebook/>

media companies, and probably from using a smart phone. There are some cases in which it is easy to switch to less invasive software or hardware options (for example Apple operating systems over Google operating systems), but completely opting out of corporate surveillance is only a realistic option for a privileged few.

A more realistic option is to use invasive software whilst running a system to *obfuscate* your activity (Brunton & Nissenbaum, 2015). Rather than trying to minimise the amount of data collected, the obfuscation strategy is to create *more* data in a way that makes it difficult to draw meaningful inferences about you.<sup>108</sup> The goal is not to prevent data from being gathered, but to minimise the amount of useful knowledge which can be garnered from the available information by ensuring that surveillance systems gather a lot of unreliable information. For example, the browser extension *AdNauseum*<sup>109</sup> runs a script in the background of the Chrome browser which clicks on every advert on every webpage, limiting advertising companies' ability to personalise adverts, and *TrackMeNot*<sup>110</sup> runs randomised web searches to hide users' actual searches in a cloud of noise or chaff. In *Obfuscation: A User's Guide* Finn Brunton and Helen Nissenbaum (2015) contextualise obfuscation practices, and defend the practice of obfuscation under conditions of extreme information asymmetry.

Given our discussion of social epistemology in section 2, it might come as a surprise that we are thinking about addressing issues around data privacy by taking measures to *reduce* the amount of knowledge that can be acquired. Brunton and Nissenbaum are aware of this issue, and argue that on grounds of justice, in conditions of information asymmetry it can be legitimate to prevent knowledge being acquired about you. At present, these pieces of obfuscation-enabling software are small academic projects, but they establish an important ethical principle about the right to resist information-gathering by companies and the state. It might be interesting to explore whether governments might establish a legal right to obfuscation to prevent technology companies from banning these extensions, or whether they might directly invest resources into creating obfuscating systems to protect their citizens' rights.

## 5.4 Solutions

Our discussion of data privacy has focused on three proposals which aim to move us in the direction of trustworthy institutions which are governed by reasonable social norms about information-acquisition and usage:

---

108. <https://www.technologyreview.com/2021/01/06/1015784/adsense-google-surveillance-adnauseam-obfuscation/>

109. <https://adnauseam.io/>

110. <https://trackmenot.io/>

1. **Ameliorating informational harms:** Establishing legislation which punishes companies that cause harms to users due to the way they collect, process and use their information.
2. **Ending the commodification of data:** Although GDPR legislation establishes that data must be collected for a legitimate purpose, we might want to explore interventions to deconstruct the market in personal data by establishing that financial gain is not a legitimate aim of data collection.
3. **Protecting obfuscation:** In the real world in which we have to interact with untrustworthy institutions and stopping data collection is unrealistic, it is important that individuals have tools available to them to minimise the amount of knowledge that can be acquired about them.

## Conclusion

If we face a crisis of truth, then its solution lies neither in individual action, nor in attempts to painstakingly recreate the institutions of the past. We need to take on this challenge collectively, and imaginatively, thinking about how to tweak existing institutions and social practices, and how to design new ones that will support public discourse and help us to generate knowledge together.

In this research overview, we have seen that post-truth narratives offer at best a limited way to think about the crisis of truth, and that conceptual tools from social epistemology can offer us a clearer way of thinking about our intellectual troubles. Recognising the importance of social factors for knowledge-acquisition, and the importance of social trust for knowledge, we can start to see how social media impedes public discourse, the ways in which journalism fails to play its filtering function, and how technology companies fail to live up to the ideal of institutional trustworthiness. With these tools in hand, we can start to think about how to improve the institutions which we rely on so heavily in our intellectual lives.

This paper is an intervention at the level of conceptual plumbing, rather than a straightforward policy proposal, and it focuses on understanding the goals of our knowledge-generating institutions, and the options which we have available to ameliorate them. Getting clear on the exact benefits and costs of these proposals and working out their fine details will require experimentation and empirical evidence that we can only acquire from experience.

## About the author

Joshua Habgood-Coote is a research fellow in the School of Philosophy, Religion and History of Science at the University of Leeds and an Honorary Research Fellow at the Department of Philosophy at the University of Bristol. He completed a PhD in Philosophy from the University of St Andrews and the University of Stirling in 2017. He works in epistemology, philosophy of language, and the philosophy of technology, has written academic papers on knowledge-how, collective inquiry, and how to think about misinformation. He has also published articles in *the Guardian*, *Real Life*, and *Aeon*.

## References

- Abrams v. United States*, 250 U.S. 616, 40 S. Ct. 17, 63 L. Ed. 1173 (1919).
- Adler, J. (2008). Epistemological Problems of Testimony. *Stanford Encyclopedia of Philosophy*.
- Anderson, E. (2006). The Epistemology of Democracy. *Episteme*. 3(1). 8-22.
- Ball, J. (2017). *Post-Truth: How Bullshit Conquered the World*. Biteback Publishing.
- Barrett, P. and Sims, J.G. (2021). *False Accusation: The Unfounded Claim that Social Media Companies Censor Conservatives*. NYU Stern Center for Business and Human Rights. [https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/6011e68dec2c7013d3caf3cb/1611785871154/NYU+False+Accusation+report\\_FINAL.pdf](https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/6011e68dec2c7013d3caf3cb/1611785871154/NYU+False+Accusation+report_FINAL.pdf).
- Briggle, A. (2008). Real friends: how the Internet can foster friendship. *Ethics and Information Technology*. 10(1). 71-79.
- Brock Jr., A. (2020). *Distributed Blackness: African American Cybercultures* (Vol. 9). NYU Press.
- Brunton, F. and Nissenbaum, H. (2015). *Obfuscation: A User's Guide for Privacy and Protest*. MIT Press.
- Campbell, J.E. and Carlson, M. (2002). Panopticon.com: Online Surveillance and the Commodification of Privacy. *Journal of Broadcasting & Electronic Media*. 46(4). 586-606.
- Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C.M., Brugnoli, E., Schmidt, A.L., Zola, P., Zollo, F., and Scala, A. (2020). The COVID-19 Social Media Infodemic. *Scientific Reports*. 10(1). 1-10.
- Costanza-Chock, S. (2018). Design Justice: Towards an intersectional feminist framework for design theory and practice. *Proceedings of the Design Research Society*.
- d'Ancona, M. (2017). *Post-Truth: The New War on Truth and How to Fight Back*. Random House.
- Davis, E. (2017). *Post-Truth: Why We Have Reached Peak Bullshit and What We Can Do About It*. London: Little, Brown.
- Dutant, J. (2015). The legend of the justified true belief analysis. *Philosophical Perspectives*. 29(1). 95-145.
- Fallis, D. and Mathiesen, K. (2019). Fake news is counterfeit news. *Inquiry*. 1-20.
- Farkas, J. and Schou, J. (2018). Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood. *Javnost - The Public*. 25(3). 298-314.
- Farkas, J. and Schou, J. (2019). *Post-Truth, Fake News and Democracy: Mapping the Politics of Falsehood*. Routledge.
- Fenton, N. and Freedman, D.J. (2017). Fake Democracy, Bad News. In: Panitch, L. and Albo, G. (eds.). *Rethinking Democracy*. 54 London: The Merlin Press. 130-149.
- Fraser, N. (1990). Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. *Social Text*. 25/26. 56-80.
- Fuller, S. (2018). *Post-Truth: Knowledge as a Power Game* (Vol. 1). London: Anthem Press.
- Geib, J. (2021). *Dynamics of instituting mini-publics for a more participatory democracy*. (Working Paper No. 1). International Youth Think Tank. Available at: <https://iythinktank.com/euro-youth-data/>.
- Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*. 438. 900-901.
- Goldman, A. (2010). Systems-oriented social epistemology. *Oxford Studies in Epistemology*. 3. 189-214.
- Goldman, A. I. (2001). Experts: Which Ones Should You Trust?. *Philosophy and Phenomenological Research*. 63(1), 85-110.
- Grim, P., Singer, D.J., Bramson, A., Holman, B., McGeehan, S., and Berger, W.J. (2019). Diversity, Ability, and Expertise in Epistemic Communities. *Philosophy of Science*. 86(1). 98-123.
- Habermas, J. (2006). Political Communication in Media Society: Does Democracy Still Enjoy an Epistemic Dimension? The Impact of Normative Theory on Empirical Research. *Communication Theory*. 16(4). 411-426.
- Habgood-Coote, J. (2019). Stop talking about fake news! *Inquiry*. 62(9-10). 1033-1065.
- Hicks, M. (2021). When did the Fire Start? In: Mullaney, T.S., Peters, B., Hicks, M., and Philip, K. (eds.). *Your Computer Is on Fire*. MIT Press.
- Hong, L. and Page, S.E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*. 101(46). 16385-16389.
- Jackson, S.J., Bailey, M., and Foucault Welles, B. (2020). *#HashtagActivism: Networks of Race and Gender Justice*. MIT Press.

- List, C. and Goodin, R.E. (2001). Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*. 9(3). 277-306.
- Marwick, A., Kuo, R., Cameron, S.J., and Weigel, M. (2021). Critical Disinformation Studies: A Syllabus. Center for Information, Technology, & Public Life (CITAP), University of North Carolina at Chapel Hill. <http://citap.unc.edu/critical-disinfo>.
- McIntyre, L. (2018). *Post-Truth*. MIT Press.
- Mejia, R., Beckermann, K., Sullivan, C. (2018). White Lies: A Racial History of the (post)truth, *Communication and Critical/Cultural Studies*, 15:2, 109-126.
- Midgley, M. (1992). Philosophical Plumbing. *Royal Institute of Philosophy Supplement*. 33. 139-151.
- Mill, J.S. (1895/1998). *On Liberty and Other Essays*. Oxford: Oxford University Press.
- Mills, C. (2007). White Ignorance. In: Sullivan, S. and Tuana, N. (eds.), *Race and Epistemologies of Ignorance*. State University of New York Press. 11-38.
- Nagel, J. (2014). *Knowledge: A Very Short Introduction*. Oxford University Press.
- Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.
- Nissenbaum, H. (2011). A Contextual Approach to Privacy Online. *Daedalus*. 140(4). 32-48.
- Noble, S.U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- O'Connor, C. and Weatherall, J.O. (2019). *The Misinformation Age: How False Beliefs Spread*. Yale University Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- O'Neill, O. (2002). *A Question of Trust: The BBC Reith Lectures 2002*. Cambridge University Press.
- O'Neill, O. (2020). Trust and Accountability in a Digital Age. *Philosophy*. 95(1). 3-17.
- Pepp, J., Michaelson, E., and Sterken, R.K. (2019). What's New About Fake News? *Journal of Ethics and Social Philosophy*. 16. 67.
- Richeri, G. and Prario, B. (2016). Media Ownership and Concentration in Italy. In: Noam, E.M. (ed.). *Who Owns the World's Media?: Media Concentration and Ownership around the World*. Oxford University Press.
- Roberts, S. T. (2019). *Behind the Screen*. Yale University Press.
- Sobieraj, S. (2018). Bitch, slut, skank, cunt: patterned resistance to women's visibility in digital publics. *Information, Communication & Society*. 21(11). 1700-1714.
- Srnicek, N. (2017). *Platform Capitalism*. John Wiley & Sons.
- Stanley, J. (2015). *How Propaganda Works*. Princeton University Press.
- Stoycheff, E., Liu, J., Xu, K., and Wibowo, K. (2019). Privacy and the Panopticon: Online mass surveillance's deterrence and chilling effects. *New Media & Society*. 21(3). 602-619.
- Sunstein, C. (2017). *#Republic*. Princeton University Press.
- Tesich, S. (1992). A Government of Lies. *The Nation*. 6 Jan.
- Tufekci, Z. (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. Yale University Press.
- Van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press.
- Véliz, C. (2020). *Privacy Is Power*. London: Penguin (Bantam Press).
- Watson, L. (2021). *The Right to Know: Epistemic Rights and Why We Need Them*. Routledge.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.
- Zuckerman, E. (2020). *Mistrust: Why Losing Faith in Institutions Provides the Tools to Transform Them*. W. W. Norton & Company.

INTERNATIONAL  
**YOUTH**  
THINK TANK

Public discourse, academic work, and newspapers are full of grave warnings about the crisis of truth faced by contemporary societies. What is this crisis and how can we address it? This research overview uses tools from social epistemology to think about how institutions support knowledge-production, considering how social media, journalism, and data privacy can be changed to improve our collective intellectual lives.

---

IYTT: INTERNATIONAL YOUTH THINK TANK

**Contact:**

Urban Strandberg  
Managing Director / Co-Founder  
urban.strandberg@lindholmen.se • +46 (0) 730-59 55 15  
<https://iythinktank.com/>

**Supporters:**

MARCUS AND AMALIA  
WALLENBERG FOUNDATION



**KVVS**  
The Royal Society  
of Arts and Sciences  
in Gothenburg



**Global  
Business  
Gate**  
GOTHENBURG / SWEDEN