episteme

ARTICLE

Knowledge and Disinformation

Mona Simion ⬦

Cogito Epistemology Research Centre, University of Glasgow, Glasgow, UK Email:
Mona.simion@glasgow.ac.uk

Abstract
This paper develops a novel account of the nature of disinformation that challenges several widely spread theoretical assumptions, such as that disinformation is a species of information, a species of misinformation, essentially false or misleading, essentially intended/aimed/having the function of generating false beliefs in/misleading hearers. The paper defends a view of disinformation as ignorance generating content: on this account, X is disinformation in a context C iff X is a content unit communicated at C that has a disposition to generate ignorance at C in normal conditions. I also offer a taxonomy of disinformation, and a view of what it is for a signal to constitute disinformation for a particular agent in a particular context. The account, if correct, carries high stakes upshots, both theoretically and practically: disinformation tracking will need to go well beyond mere fact checking.

1. Introduction

This paper develops a full account of the nature of disinformation. The view, if correct, carries high stakes upshots, both theoretically and practically. First, it challenges several widely spread theoretical assumptions about disinformation – such as that it is a species of information, a species of misinformation, essentially false or misleading, essentially intended/aimed/having the function of generating false beliefs in/misleading hearers. Second, it shows that the challenges faced by disinformation tracking in practice go well beyond mere fact checking.

I begin by an interdisciplinary scoping of literature in information science, communication studies, computer science, and philosophy of information to identify several claims constituting disinformation orthodoxy. I then present counterexamples to these claims, and motivate my alternative account. Finally, I outline the view put forth in this study: disinformation as ignorance-generating content.

2. Information and disinformation

Philosophers of information, as well as information and communication scientists have traditionally focused their efforts in three main directions: offering an analysis of

information, a way to measure it, and investigating prospects for analysing epistemic states – such as knowledge and justified belief – in terms of information. Misinformation and disinformation have traditionally occupied the backseat of these research efforts.[1] The assumption has mostly been that a unified account of the three is going to become readily available as soon as we figure out what information is. As a result, for the most part, misinformation and disinformation have received dictionary treatment: for whatever the correct analysis of information was taken to be, misinformation and disinformation have either been taken to constitute the false variety thereof (misinformation) and the intentionally false/misleading variety thereof (disinformation) – by theorists endorsing non-factive accounts of disinformation – or, alternatively, something like information minus truth (misinformation) or information minus truth spread with an intention to mislead (disinformation), in the case of theorists endorsing factive accounts of information.

This is surprising in more than one way: first, it is surprising that philosophers of any brand would readily and unreflectively endorse dictionary definitions of pretty much anything – not to mention entities with such high practical stakes associated with them – such as mis/disinformation. Second, it is surprising that not more effort on the side of information, communication, and computer scientists has been spent on identifying a correct account of the nature of disinformation, given increasingly high stakes issues having to do with the spread of disinformation threatening our democracies, our trust in expertise, our uptake of health provision, and our social cohesion. We are highly social creatures, dependent on each other for flourishing in all walks of life. Our epistemic endeavours make no exception: due to our physical, geographical, and psychological limitations, most of the information we have is sourced in social interactions. We must inescapably rely on the intellectual labour of others, from those we know and trust well, to those whose epistemic credentials we take for granted online. Given the staggering extent of our epistemic dependence – one that recent technologies have only served to amplify – having a correct account of the nature of mis/disinformation, in order to be able to reliably identify it and escape it, is crucial.

Disinformation is widespread and harmful, epistemically and practically. We are currently facing a global information crisis that the Secretary-General of the World Health Organization (WHO) has declared an 'infodemic'. Furthermore, crucially, there are two key faces to this crisis: two ways in which disinformation spreads societal ignorance: One concerns the widespread *sharing of disinformation* (e.g., fake cures, health superstitions, conspiracy theories, political propaganda, etc.) especially online and via social media, which contribute to dangerous and risky political and social behaviour. Separately, though at least as critical to the wider infodemic we face, is the prevalence of resistance to evidence: even when the relevant information available is reliably sourced and accurate, many information consumers fail to take it on board or otherwise resist or discredit it (Klintman 2019; Simion and Kelp 2023), due to the rise in lack of trust and scepticism generated by the ubiquity of disinformation. An important payoff, then, of a correct analysis of the nature of disinformation is an understanding of how to

---

[1]While fully fledged accounts of the nature of disinformation are still thin on the ground, a number of information scientists and philosophers of information have begun to address the problem of disinformation (Calvert 2001; Fallis 2009; Hernon 1995; Karlova and Fisher 2013; Lynch 2001; Piper 2002; Rubin and Conroy 2012; Skinner and Martin 2000; Walsh 2010; Whitty *et al.* 2012).

help build and sustain more resilient trust networks. It is urgent that we gain such answers and insights: according to the 2018 Edelman Trust Barometer, UK public trust in social media and online news has plummeted to below 25%, and trust in government is at a low 36%. This present crisis in trust corresponds with a related crisis of distrust, in that the dissemination and uptake of disinformation, particularly on social media, have risen dramatically the past few years (Barclay 2022; Levinson 2017; Lynch 2001).

3. Against disinformation orthodoxy

In what follows, I will scope the scientific and philosophical literature and identify three very widely spread – and rarely defended – assumptions about the nature of disinformation, and argue against their credentials.

1. Assumption §1: *Disinformation is a species of information* (e.g. Carnap and Bar-Hillel 1952; Cevolani 2011; D'Alfonso 2011; Dinneen and Brauner 2015; Fallis 2009; Floridi (2007, 2008, 2011); Frické 1997; Shannon 1948).

These theorists take information to be non-factive, and disinformation to be the false and intentionally misleading variety thereof. On accounts like these, information is something like meaning: 'The cat is on the mat', on this view, caries the information 'the cat is on the mat' in virtue of the fact that it means that the cat is on the mat. Disinformation, on this view, consists in spreading 'the cat is on the mat' in spite of knowing it to be false, and with the intention to mislead.

Why think in this way? Two rationales can be identified in the literature: a practical and a theoretical rationale:

The practical rationale: Factivity doesn't matter for the information scientist: In the early days of information science, the thought behind this went roughly as follows: For the information scientist, the stakes associated with the factivity/non-factivity of information are null: after all, what the computer scientist/communication theorist cares about is the quantity of information that can be packed in a particular signal/channel. Whether the relevant content will be true or not makes little difference to the prospects of answering this question.

True: when it comes to how many bits of data one can pack into a particular channel, factivity doesn't make much difference. However, times have changed, and so have the questions the information scientist needs to answer: the 'infodemic' had brought with it concerted efforts to fight the spread of disinformation online and through traditional media. We have lately witnessed an increased interest in researching and developing automatic algorithmic detection of misinformation and disinformation: e.g. PHEME-project (2014), Kumar and Geethakumari's 'Twitter algorithm' (2014), Karlova and Fisher's diffusion model (Karlova and Fisher 2013), and the Hoaxy platform (Shao *et al.* 2016) – to name a few. Interest from developers has also been matched by interest from policy makers: the European Commission has brought together major online platforms, emerging and specialised platforms, players in the advertising industry, fact-checkers, research and civil society organisations to deliver a strengthened Code of Practice on Disinformation (June 2022). The American Library Association (2005) has issued a 'Resolution on Disinformation, Media Manipulation, and the Destruction of Public Information.' The UK Government has recently published a call for evidence into how to address the spread of disinformation via employing trusted voices. These are, of course, only a few examples of disinformation-targeting initiatives. If all of these and others are to stand any chance at succeeding, we need a correct analysis of disinformation. The practical rationale is false.

The theoretical rationale: Natural language gives us clear hints to non-factivity of information: we often hear people say things like 'the media is spreading a lot of fake information'. We also say things like 'The library contains a lot of information' – however, clearly, there will be a fair share of false content featured in any library (Fallis 2009). If this is correct, the argument goes, natural language suggest information is not factive – there can be true and false varieties thereof. Therefore, disinformation is a species of information.

One first problem with the natural language rationale is that the cases in point are underdeveloped. Take the library case: I agree that we will often say that libraries contain information in spite of the likelihood of false content. This, however, is compatible with information being factive: after all, the claim about false content, as far as I can see, is merely an existential claim. There being some false content in a library is perfectly compatible with it containing a good amount of information alongside it. Would we still say the same were we to find out that this particular library contains only false-hoods? I doubt it. If anything, at best, we might say something like: this library contains a lot of fake information.

Which brings me to my more substantial point: natural language at best cannot decide the factivity issue either way, and at worst suggests information is factive; here is why: First, it is common knowledge in formal semantics that when a complex expression consists of an intentional modifier and a modified expression, then we cannot infer a type–species relation – or, indeed, to the contrary, in some cases, we might be able to infer that a type–species relation is absent. This latter class includes the so-called privative modifiers such as fake, former, and spurious, which get their name from the fact that they license the inference to 'not x' (McNally 2016). If so, the fact that 'information' takes fake as modifier suggests, if anything, that information is factive, in that fake acts as privative: it suggests it is not information to begin with. As Dretske well puts it, mis/disinformation is as much a type of information as a decoy duck is a type of duck (1981). (See also Floridi (2004, 2005*a*, 2005*b*), Sequoiah-Grayson (2007), Mingers (1995) for defences of factivity.) If information is factive and disinformation is not, however, the one is not the species of the other. The theoretical rationale is false: meaning and disinformation come apart on factivity grounds. As Dretske well puts it,

> signals may *have* a meaning, but they *carry* information. What information a signal carries is what it is capable of telling us, telling us truly, about another state of affairs. […] When I say I have a toothache, what I say means that I have a toothache whether it's true or false. But when false, it fails to carry the information that I have a toothache. (1981: 44)

Natural language semantics also gives us further, direct reason to be sceptical about disinformation being a species of information: several instances of dis-prefixed properties that fail to signal type/species relations: disbarring is not a way of becoming a member of the bar, displeasing is not a form of pleasing, and displacing is not a form of placing. More on this below.

2. Assumption §2: *Disinformation is a species of misinformation* (e.g. Fallis (2009, 2015), Floridi (2007, 2008, 2011))

Misinformation is essentially false content, the mis- prefix modifies as badly, wrongly; unfavourably; in a suspicious manner; opposite or lack of; not. In this, misinformation is essentially non-information, in the same way in which fake gold is essentially non-gold.

As opposed to this, for the most part, dis- modifies as deprive of (a specified quality, rank, or object); exclude or expel from. In this, paradigmatically,[2] dis- does not negate the prefixed content, but rather it signals un-doing: if misplacing is placing in the wrong place, displacing is taking out of the right place. Disinformation is not a species of misinformation any more than displacing is a species of misplacing. To think otherwise is to engage in a category mistake.

Note, also, that disinformation, as opposed to misinformation, is not essentially false: I can, for instance, disinform you via asserting true content and generating false implicatures. I can also disinform you via stripping you of justification via misleading defeaters.

Finally, note, also, that information/misinformation exists out there, disinformation is us-dependent: there is information/misinformation in the world, without anyone being informed/misinformed (Dretske 1981), while there is no disinformation without target: disinformation is essentially second personal, audience-involving.[3]

3. Assumption §3: *Disinformation is essentially intentional/functional* (e.g. Fallis 2009, 2015; Fetzer 2004*a*, 2004*b*; Floridi 2007, 2008, 2011; Mahon 2008)

The most widely spread assumption across disciplines is that disinformation is intentionally spread misleading content (where the relevant way to think about the intention at stake can be quite minimal, as having to do with content that has the function to mislead (Fallis 2009, 2015). I think this is a mistake generated by paradigmatic instances of disinformation. I also think it is a dangerous mistake, in a world in which automated spread of disinformation that has little to do with any intention on the part of the programmer, to operate with such a restricted concept of disinformation. To see this, consider a black-box artificial intelligence (AI) that, in the absence of any intention to this effect on the part of the designer, learns how to and proceeds to widely spreading false claims about the Covid vaccines in the population, in a systematic manner. Intention is missing in this case, as is function: the AI has not been designed to proceed in this way (no design function), and it does not do so in virtue of some benefit or another generated for either itself or any human user (no etiological function). Furthermore, most importantly, AI is not the only place where the paradigmatic and the analytic part ways: I can disinform you unintentionally (where, furthermore, the case is one of genuine disinformation rather than mere misinformation). Consider the following case: I am a trusted journalist in village V, and, unfortunately, I am the kind of person who is unjustifiably very impressed by there being any scientific disagreement whatsoever on a given topic. Should even the most isolated voices express doubt about a scientific claim, I withhold belief. Against this background, I report on V TV (the local station in V) that there is disagreement in science about climate change and the safety of vaccines. As a result, whenever V inhabitants encounter expert claims that climate change is happening and vaccines are safe, they hesitate to update accordingly. A couple of things about this case: First, this is not a case of false content/misinformation spreading – after all, it is true that there is disagreement on these issues (albeit very isolated). Second, there is no intention to mislead present at the context, nor any corresponding function. Third, and crucially, however, it is a classic case of disinformation spreading

---

[2]Not essentially. Disagreeable and dishonest are cases in point, where the dis- prefix modifies as not-. The underlying rationale for the paradigmatic usage, however, is solidly grounded in the Latin, and later French source of the English version of the prefix. (Latin prefix meaning 'apart', 'asunder', 'away', 'utterly', or having a privative, negative, or reversing force).

[3]See Grundmann (forthcoming) for an audience-oriented account of fake news.

– indeed, I submit, if our account of disinformation cannot accommodate this case, we should go back to the drawing board.

4. Knowledge and disinformation

In what follows, I will offer a knowledge-first account of disinformation that aims to vindicate the findings of the previous section.

Traditionally, in epistemology (e.g. Dretske 1981) and philosophy of information alike, the relation between knowledge and information has been conceived on a right-to-left direction of explanation: i.e. several theorists have attempted to analyse knowledge in terms of information. Notably, Fred Dretske thought knowledge was information-caused true belief. More recently, Luciano Floridi's network theory involves an argument for the claim that should information be embedded within a network of questions and answers, then it is necessary and sufficient for it to count as knowledge. Accounts like these, unsurprisingly, encounter the usual difficulties in analysing knowledge (see e.g. Ichikawa and Steup 2018).

The fact that information-based analyses of knowledge remain unsuccessful, however, is not good reason to abandon the theoretical richness of the intuitive tight relation between the two. In extant work (Simion and Kelp 2022) I have developed a knowledge-based account of information that explores the prospects of the opposite, left-to-right direction of explanation: according to this view, very roughly, a signal s carries the information that p iff it has the capacity to generate knowledge that p.[4] On this account, then, information carries its functional nature up its sleeve, as it were: just like a digestive system is a system with the function to digest, and the capacity to do so in normal conditions, information has the function to generate knowledge, and the capacity to do so in normal conditions – i.e. given a suitably situated agent.

Against this background, I find it very attractive to think of disinformation as the counterpart of information: roughly, as stuff that has the capacity to generate or increase ignorance – i.e. to fully/partially strip someone of their status as knower, or to block their access to knowledge, or to decrease their closeness to knowledge. Here is the account I want to propose:

Disinformation as ignorance generating content (DIGC): X is disinformation in a context C iff X is a content unit communicated at C that has a disposition to generate or increase ignorance at C in normal conditions.

Normal conditions are understood in broadly etiological functionalist terms (e.g. Graham 2010; Simion 2019, 2021*a*, 2021*b*) as the conditions at which our knowledge-generating cognitive processes have acquired their function of generating knowledge. The view is contextualist in that the same communicated content will act differently depending on contextual factors such as: the evidential backgrounds of the audience members, the shared presuppositions, extant social relations, and social norms. Importantly, as with dispositions more generally, said content need not actually generate ignorance at the context – after all, dispositions are sometimes masked.

---

[4]My co-author and I owe inspiration for this account to Fred Dretske's excellent 1981. While Dretske himself favours the opposite direction of analysis (knowledge in terms of information), at several points he says things that sound very congenial to our preferred account, and that likely played an important role in shaping our thinking on this topic. On page 44 of his 1981, for instance, Dretske claims that 'Roughly speaking, information is that commodity capable of yielding knowledge, and what information a signal carries is what we can learn from it.'

Now, importantly, generating/increasing ignorance can be done in a variety of ways – which means that disinformation will come in diverse incarnations. In what follows, I will make an attempt at offering a comprehensive taxonomy of disinformation. (The ambition to exhaustiveness is probably beyond the scope of this paper, or even of an isolated philosophical project such as mine; however, it will be useful to have a solid taxonomy as a basis for a fully-fledged account of disinformation: at a minimum, any account should be able to incorporate all varieties of disinformation we will have identified.)[5] Here it goes:

(1) Disinforming via spreading content that has the capacity of *generating false belief*: The paradigmatic case of this is the traditionally recognised species of disinformation: intentionally spread false assertions with the capacity to generate false beliefs in hearers.

(2) Disinforming via misleading defeat: This category of disinformation has the capacity of stripping the audience of held knowledge/being in a position to know via defeating justification.

(3) Disinforming via content that has the capacity of inducing epistemic anxiety: this category of disinformation has the capacity of stripping the audience of knowledge via belief defeat. The paradigmatic way to do this is via artificially raising the stakes at the context/introducing irrelevant alternatives as being relevant: 'Are you really sure that you're sitting at your desk? After all, you might well be a brain in a vat.'; 'Are you really sure he loves you? After all, he might just be an excellent actor, in which case you will have wasted years of your life.' The way this variety of disinforming works is via falsely implicating that these error possibilities are relevant at the context, when in fact they are not. In this, the audience's body of evidence is changed to include misleading justification defeaters.

(4) Confidence-defeating disinformation: has the capacity to reduce justified confidence via justification/doxastic defeat: you are sure that your name is Anna, but I introduce misleading (justification/doxastic) defeaters, which gets you to lower your confidence. You may remain knowledgeable about p: 'My name is Anna', in cases in which the confidence lowering does not bring you below the knowledge threshold. Compatibly, however, your knowledge – or evidential support – concerning the correct likelihood of p is lost: you now take/are justified to take the probability of your name being Anna to be much lower than it actually is.

(5) Disinforming via exploiting pragmatic phenomena: Pragmatic phenomena can be easily exploited to the end of disinforming in all ways above: True assertions carrying false implicatures (Grice 1957, 1967, 1989) will display this capacity to generate false beliefs in the audience. I ask: 'Is there a gas station anywhere near here?' I'm almost out of gas,' and you reply 'Yeah, sure, just one mile in that direction!,' knowing perfectly well that it's been shut down for years. Another way in which disinformation can be spread via making use of pragmatic phenomena is by introducing false presuppositions. Finally, both justification and doxastic defeat will be achievable via speech acts with true content but problematic pragmatics, even in the absence of generating false implicatures.

---

[5]See Simion (forthcoming, 2021*a*, 2021*b*, 2019) and Simion and Kelp (2022, 2023) for knowledge-centric accounts of trustworthiness, testimonial entitlement, and evidence resistance. See Kelp and Simion (2017, 2021) for a functionalist account of the distinctive value of knowledge.

What all of these ways of disinforming have in common is that they generate ignorance – either by generating false beliefs, knowledge loss, or generating a decrease in warranted confidence. One important thing to notice, which was also briefly discussed in the previous section, is that this account, and the associated taxonomy, is strongly second-personal, in that disinformation has to do with the capacity to have a particular effect – generating ignorance – in the audience. Importantly, though, this capacity will heavily depend on the audience's background evidence/knowledge: after all, in order to figure out whether a particular piece of communicated content has the disposition to undermine an audience in their capacity as knowers, it is important to know their initial status as knowers. Here is, then, on my view, in more precise terms, what it takes for a signal to carry a particular piece of disinformation for an audience A:

Agent disinformation: A signal r carries disinformation for an audience A wrt *p* iff A's evidential probability that *p* conditional on r is less than A's unconditional evidential probability that *p*, and p is true.

What is relevant for agent disinformation with regard to p is the probability that p on agent's evidence. What is A's evidential probability? In my view (Simion forthcoming), A's evidence – and, correspondingly, what underlies A's evidential probability – lies outwith A's skull: it consists in probability raisers that A is in a position to know. Here is the account I have defended in previous work (where the relevant probability is evidential probability):

Evidence as knowledge indicators: a fact e is evidence for p for S iff S is in a position to know e, and $P(p/e) > P(p)$ (Simion forthcoming).

Evidence, thus, may consist of facts that increase extant evidential probability and that are located 'in the head', or in the world. Some facts – whether they are in the head or in the world, it does not matter – are available to A, they are, as it were, 'at hand' in A's (internal or external) epistemic environment. Some – whether in the head (think of justified implicit beliefs, for instance) or in the world, it does not matter – are not thus available to A. In turn, my notion of availability will track a psychological 'can' for an average cogniser of the sort exemplified. There are qualitative limitations on availability: we are cognitively limited creatures. There are types of information that we just cannot access, or process, and types of support relations that we cannot process. There are also quantitative limitations on my information accessing and processing: I lack the power to process everything in my visual field, it's just too much information.

I take this availability relation to have to do with a fact being within the easy reach of my knowledge generating cognitive processes. A fact F being such that I am in a position to know it has to do with the capacity of my properly functioning knowledge generating processes to take up F:

Being in a position to know: S is in a position to know a fact F iff S has a cognitive process with the function of generating knowledge that can (qualitatively, quantitatively, and environmentally) easily uptake F in cognisers of S's type (Simion forthcoming).

This completes my account of disinformation. On this account, disinformation is the stuff that undermines one's status as a knower. It does so via lowering their evidential probability for p – the probability on the p-relevant facts that they are in a position to know – for a true proposition. It can, again, do so by merely communicating to A (semantically, pragmatically, etc.) that not-p when in fact p is the case. Alternatively, it can do so by (partially or fully) defeating A's justification for p, A's belief that p is the case, or A's confidence in p.

One worry that the reader may have at this point goes along the following lines: isn't the account in danger of over-generating disinformation? After all, every true assertion

that I make in your presence about p being the case may, for all I know, serve as (to some extent) defeating evidence for a different proposition q, which may well be true. I truthfully tell you it's raining outside, which, unrelatedly and unbeknownst to me, together with your knowledge about May not liking the rain, may function as partial rebutting defeat for 'Mary is taking a walk' – which may well, nevertheless, be true. Is it now appropriate to accuse me of having thereby disinformed you? Intuitively, that seems wrong.[6]

Three things about this: first, note that restricting disinforming via defeat to intentional/functional cases will not work for the same reasons that created problems for the intention/function condition on disinformation more broadly: we want an account of disinformation to be able to predict that asserters generating doubt about e.g. climate change via spreading defeaters to scientific evidence, even if they do it without any malicious intention, are disinforming the audience.

Second, note that it is independently plausible that, just like any bad deed can be performed blamelessly, one can also disinform blamelessly; if so, given garden-variety epistemic and control conditions on blame, any plausible account of disinformation will have to accommodate non-knowledgeable and non-intentional instances of disinformation.

Finally, note that we don't need to restrict the account in order to accommodate the datum that disinformation attribution, and the accompanying criticism, would sound inappropriate in the case above. We can use simple Gricean pragmatics to predict as much, via the maxim of relevance: since the issue of whether Mary was going for a walk was not under discussion, nor remotely relevant at our conversational context, flat our accusing you of disinforming me when you assert truthfully that it's raining is pragmatically impermissible (although strictly speaking true with regard to Mary's actions).

Going back to the account: note that, interestingly, on this view, one and the same piece of communication can, at the same time, be a piece of information and a piece of disinformation: information, as opposed to disinformation is not context relative. Content with knowledge-generating potential – i.e. that can generate knowledge in a possible agent – is information. Compatibly, the same piece of content, at a particular context, can be a piece of disinformation insofar as it has a disposition to generate ignorance in normal conditions. I think this is the right result: me telling you that p: 99% of black people at Club X are staff members is me informing you that p. Me telling you that p in the context of you inquiring as to whether you can give your coat to a particular black man is a piece of disinformation since it carries a strong disposition (due to the corresponding relevance implicature) to generate the unjustified (and maybe false) belief in you that this particular black man is a member of staff (Gendler 2011).

Finally, and crucially: my account allows that disinformation for an audience A can exist in the absence of A's hosting any relevant belief/credence: (partial) defeat of epistemic support that one is in a position to know is enough for disinformation. Even if I (irrationally) don't believe that vaccines are safe, or that climate change is happening, to begin with, I am still vulnerable to disinformation in this regard in that I am vulnerable to content that has, in normal conditions, a disposition to defeat epistemic support available to me that vaccines are safe and climate change is happening. In this, disinformation, on my view, can generate ignorance even in the absence of any doxastic attitude – by decreasing closeness to knowledge via defeating available evidence. This, I submit, is a very nice result: in this, the account explains the most dangerous variety

---

[6]Many thanks to Julia Staffel and Martin Smith for pressing me on this.

of disinformation available out there – disinformation targeting the already epistemically vulnerable.

5. Concluding remarks and practical stakes

Disinformation is not a type of information, and disinforming is not a way of informing: while information is content with knowledge-generating potential, disinformation is content with a disposition to generate ignorance in normal conditions at the context at stake. This way to think about disinformation, crucially, tells us that it is much more ubiquitous and hard to track than it is currently taken to be in policy and practice: mere FactCheckers just won't do. Some of the best disinformation detection tools at our disposal will fail to capture most types of disinformation. To give but a few examples – but more research on this is clearly needed: the PHEME project aims to algorithmically detect and categorise rumours in social network structures (such as Twitter and Facebook), and to do so, impressively, in near real time. The rumours are mapped according to four categories, which include 'disinformation, where something untrue is spread with malicious intent' (Søe 2016). Similarly, Kumar and Geethakumari's project (2014) develops an algorithm which ventures to detect and flag whether a tweet is misinformation or disinformation. In their framework 'Misinformation is false or inaccurate information, especially that which is deliberately intended to deceive [and] Disinformation is false information that is intended to mislead, especially propaganda issued by a government organization to a rival power or the media.' (Kumar and Geethakumari 2014: 3). In Karlova and Fisher's (2013) diffusion model, disinformation is taken to be deceptive information. Hoaxy (Shao *et al.* 2016) is 'a platform for the collection, detection, and analysis of online misinformation, defined as "false or inaccurate information"' (Shao *et al.* 2016: 745). Examples targeted, however, include clear cases of disinformation such as rumours, false news, hoaxes and elaborate conspiracy theories (which are Shao *et al.* 2016).

It becomes clear that these otherwise excellent tools are just the beginning of a much wider effort that is needed in order to capture disinformation in all of its facets, rather than mere paradigmatic instances thereof. At a minimum, pragmatic deception mechanisms, as well as evidential probability lowering potentials will need to be tracked against an assumed (common) evidential background of the audience.[7]

References

American Library Association (2005). 'Resolution on Disinformation, Media Manipulation and the Destruction of Public Information.' Available at http://www.ala.org/aboutala/sites/ala.org.aboutala/files/content/governance/policymanual/updatedpolicymanual/ocrpdfofprm/52-8disinformation.pdf.

Barclay D.A. (2022). *Disinformation: The Nature of Facts and Lies in the Post-Truth Era.* Maryland, MD: Rowman & Littlefield.

Calvert P.J. (2001). 'Scholarly Misconduct and Misinformation on the World Wide Web.' *Electronic Library* **19**(4), 232–40.

Carnap R. and Bar-Hillel Y. (1952). *An Outline of a Theory of Semantic Information.* Technical Report 247. Cambridge, MA: Research Laboratory of Electronics, MIT.

Cevolani G. (2011). 'Strongly Semantic Information and Verisimilitude.' *Etica e Politica / Ethics and Politics*, **13**(2), 159–179.

D'Alfonso S. (2011). 'On Quantifying Semantic Information.' *Information*, **2**(1), 61–101.

Dinneen J.D. and Brauner C. (2015). 'Practical and Philosophical Considerations for Defining Information as Well-Formed, Meaningful Data in the Information Sciences.' *Library Trends*, **63**(3), 378–400.

Dretske F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Fallis D. (2009). A Conceptual Analysis of Disinformation, Preprint from iConference, Tucson, AZ, 331. Available at www.ideals.illinois.edu/bitstream/handle/2142/15205/fallis_disinfo1.pdf?sequence=2.

Fallis D. (2015). 'What is Disinformation?.' *Library Trends* **63**(3), 401–26.

Fetzer J.H. (2004*a*). 'Information: Does it Have To Be True?' *Minds and Machines* **14**, 223–9.

Fetzer J.H. (2004*b*). 'Disinformation: The Use of False Information'. *Minds and Machines* **14**, 231–40.

Floridi L. (2004). 'Outline of a Theory of Strongly Semantic Information.' *Minds and Machines*, **14**(2), 197–221.

Floridi L. (2005*a*). 'Is Semantic Information Meaningful Data?' *Philosophy and Phenomological Research* **LXX**(2), 351–70.

Floridi L. (2005*b*). 'Semantic Conceptions of Information.' In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, Spring 2013 ed., 46 pp. Palo Alto: Stanford University. Available at http://plato.stanford.edu/archives/spr2013/entries/information-semantic/.

Floridi L. (2007). 'In Defence of the Veridical Nature of Semantic Information.' *European Journal of Analytic Philosophy (EUJAP)* **3**(1), 31–41.

Floridi L. (2008). 'Trends in the Philosophy of Information.' In P. Adriaans and J. van Benthem (eds), *Philosophy of Information*, pp. 113–31. Amsterdam and Oxford: Elsevier B.V.

Floridi L. (2011). *The Philosophy of Information*. Oxford: Oxford Scholarship Online. Available at www.oxfordscholarship.com.

Frické M. (1997). 'Information Using Likeness Measures.' *Journal of the American Society for Information Science*, **48**(10), 882–892.

Gendler T.S. (2011). 'On the Epistemic Costs of Implicit Bias.' *Philosophical Studies* **156**, 33–63.

Graham P. (2010). 'Testimonial Entitlement and the Function of Comprehension.' In A. Haddock, A. Miller and D. Pritchard (eds), *Social Epistemology*, pp. 148–74. New York: Oxford University Press.

Grice H.P. (1957/1989/1991). 'Meaning.' *Studies in the Way of Words*, Paperback edition, pp. 213–23. Cambridge, MA and London: First Harvard University Press.

Grice H.P. (1967/1989/1991). 'Logic and Conversation.' *Studies in the Way of Words*, Paperback edition, pp. 22–40. Cambridge, MA and London: First Harvard University Press.

Grice H.P. (1989/1991). *Studies in the Way of Words*, Paperback edition. Cambridge, MA and London: First Harvard University Press.

Grundmann T. (forthcoming). 'Fake News: The Case for a Purely Consumer-Oriented Explication.' *Inquiry*.

Hernon P. (1995). 'Disinformation and Misinformation through the Internet: Findings of an Exploratory Study.' *Government Information Quarterly* **12**(2), 133–9.

Ichikawa J.J. and Steup M. (2018). 'The Analysis of Knowledge.' In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). Available at https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/.

Karlova N.A. and Fisher K.E. (2013). 'A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behavior.' *Information Research* **18**(1).

Kelp C. and Simion M. (2017). 'Comodious Knowledge.' *Synthese* **194**(5), 1487–502.

Kelp C. and Simion M. (2021). *Sharing Knowledge: A Functionalist Account of Assertion*. Cambridge: Cambridge University Press.

Klintman M. (2019). *Knowledge Resistance: How We Avoid Insight from Others*. Manchester: Manchester University Press.

Kumar K.P.K. and Geethakumari G. (2014). 'Detecting Misinformation in Online Social Networks Using Cognitive Psychology.' *Human-centric Computing and Information Sciences* **4**(1), 22.

Levinson P. (2017). 'Fake News in Real Context.' *Explorations in Media Ecology* **18**(1), 173–7.

Lynch C.A. (2001). 'When Documents Deceive: Trust and Provenance as New Factors for Information Retrieval in a Tangled Web.' *Journal of the American Society for Information Science and Technology* **52**(1), 12–7.

Mahon J. (2008). 'The Definition of Lying and Deception.' In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, Fall 2019 ed., 16 pp. Palo Alto: Stanford University. Available at http://plato.stanford.edu/archives/fall2009/entries/lying-definition/.

McNally L. (2016). 'Modification.' In M. Aloni and P. Dekker (eds), *Cambridge Handbook of Formal Semantics*. Cambridge: Cambridge University Press, pp. 442–66.

Mingers J.C. (1995). 'Information and Meaning: Foundations for an Intersubjective Account.' *Information Systems Journal* **5**(4), 285–306.

PHEME (2014). 'About Pheme.' Available at www.pheme.eu (accessed 3 March 2014).

Piper P.S. (2002). 'Web Hoaxes, Counterfeit Sites, and Other Spurious Information on the Internet.' In A.P. Mintz (ed.), *Web of Deception*, pp. 1–22. Medford, NJ: Information Today.

Rubin V.L. and Conroy N. (2012). 'Discerning Truth from Deception: Human Judgments and Automation Efforts.' *First Monday* **17**(3). Retrieved 26 November 2014, from http://firstmonday.org/ojs/index.php/fm/article/view/3933/3170.

Sequoiah-Grayson S. (2007). 'The Metaphilosophy of Information.' *Minds and Machines* **17**, 331–44.

Shannon C.E. (1948). 'A Mathematical Theory of Communication.' *The Bell System Technical Journal* **27**, 379–423.

Shao C., Ciampaglia G.L., Flammini A. and Menczer F. (2016). 'Hoaxy: A platform for tracking online misinformation.' WWW'16 Companion, Montréal and Québec, 11–15 April, pp. 745–750. Available at http://doi.org/10.1145/2872518.2890098.

Simion M. (2019). 'Hermeneutical Injustice as Basing Failure.' In P. Bondy and J.A. Carter(eds),, *Well Founded Belief: New Essays on the Epistemic Basing Relation*. New York: Routledge.

Simion M. (2021*a*). *Shifty Speech and Independent Thought*. Oxford: Oxford University Press.

Simion M. (2021*b*). 'Testimonial Contractarianism: A Knowledge-First Social Epistemology.' *Nous* **55**(4), 891–916.

Simion M. (forthcoming). 'Resistance to Evidence and the Duty to Believe.' *Philosophy and Phenomenological Research*.

Simion M. and Kelp C. (2022). Information, Misinformation, Disinformation. Manuscript.

Simion M. and Kelp C. (2023). 'What Is Trustworthiness?' *Nous*, Online First.

Skinner S. and Martin B. (2000). 'Racist Disinformation on the World Wide Web: Initial Implications for the LIS Community.' *The Australian Library Journal* **49**(3), 259–69.

Søe S.O. (2016). *The Urge to Detect, the Need to Clarify. Gricean Perspectives on Information, Misinformation, and Disinformation*. PhD thesis, Faculty of Humanities, University of Copenhagen.

Walsh J. (2010). 'Librarians and Controlling Disinformation: Is Multi-Literacy Instruction the Answer?' *Library Review* **59**(7), 498–511.

Whitty M.T., Buchanan T., Joinson A.N. and Meredith A. (2012). 'Not All Lies are Spontaneous: An Examination of Deception across Different Modes of Communication.' *Journal of the American Society for Information Science and Technology* **63**(1), 208–16.

Mona Simion is Professor of Philosophy and Director of the Cogito Epistemology Research Centre at the University of Glasgow. She has published widely in epistemology, philosophy of language, and moral and political philosophy. She is the author of *Shifty Speech and Independent Thought* (Oxford University Press, 2021), *Sharing Knowledge* (Cambridge University Press, 2021), and *Resistance to Evidence* (Cambridge University Press, 2023). She runs several major research projects in social epistemology, funded by e.g. the European Research Council and the Leverhulme Trust. She has received several prestigious international awards and fellowships, including the Young Epistemologist Prize (2021) and the Mind Fellowship (2018). She is managing editor at Ergo and associate editor at Philosophical Studies. She sits on the Executive Board of the Aristotelian Society, the Management Board of the British Society for Theory of Knowledge, and the Advisory Board of the Institute of Philosophy, London. She also sits on the Board of the Young Academy of Europe and is a fellow of the Royal Society of Edinburgh Young Academy of Scotland.