(Fake?) News Alert: Intellectual Virtues Required for Online Knowledge!

Paul R. Smart, University of Southampton


Short url: https://wp.me/p1Bfg0-3Uq (provided by WordPress)
Richard Heersmink’s (2018) article, *A virtue epistemology of the Internet: Search engines, intellectual virtues, and education*, provides an important and timely analysis of the Internet from the standpoint of virtue epistemology.¹ According to Richard, the Internet is an important epistemic resource, but it is one that comes with a range of epistemic hazards. Such hazards, he suggests, motivate a consideration of the ways in which individuals should interact with the Internet.

In particular, Richard appeals to a specific branch of virtue epistemology, known as virtue responsibilism, arguing that certain kinds of cognitive trait (e.g., curiosity and open-mindedness) are useful in helping us press maximal epistemic benefit from the Internet. Given the utility of such traits, coupled with the epistemic importance of the Internet, Richard suggests that educational policy should be adapted so as to equip would-be knowers with the cognitive wherewithal to cope with the epistemic challenges thrown up by the online environment.

There is, no doubt, something right about all this. Few would disagree with the claim that a certain level of discernment and discrimination is important when it comes to the evaluation of online content. Whether such ‘virtues’ are best understood from the perspective of virtue responsibilism or virtue reliabilism is, I think, a moot point, for I suspect that in the case of both virtue responsibilism and virtue reliabilism what matters is the way in which belief-forming informational circuits are subject to active configuration by processes that may be broadly construed as metacognitive in nature (Smart, in pressa). That, however, is a minor quibble, and it is one that is of little consequence to the issues raised in Richard’s paper.

For the most part, then, I find myself in agreement with many of the assumptions that motivate the target article. I agree that the Internet is an important epistemic resource that is unprecedented in terms of its scale, scope, and accessibility. I also agree that, at the present time, the Internet is far from an epistemically safe environment, and this raises issues regarding the epistemic standing of individual Internet users. In particular, it looks unlikely that the indiscriminate selection and endorsement of online information will do much to bolster one’s epistemic credentials.

We thus encounter something of a dilemma: As an epistemic resource, the Internet stands poised to elevate our epistemic standing, but as an open and public space the Internet provides ample opportunities for our doxastic systems to be led astray. The result is that we are obliged to divide the online informational cornucopia into a treasure trove of genuine facts and a ragbag collection of ‘false facts’ and ‘fake news.’ The information superhighway, it seems, promises to expand our epistemic power and potential, but the road ahead is one that is fraught with a dizzying array of epistemic perils, problems, and pitfalls. What ought we to do in response to such a situation?

¹ This work is supported under SOCIAM: The Theory and Practice of Social Machines. The SOCIAM Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1 and comprises the Universities of Southampton, Oxford and Edinburgh.
It is at this point that I suspect my own views start to diverge with those of the target article. Richard’s response to the dilemma is to focus attention on the individual agent and consider the ways in which an agent’s cognitive character can be adapted to meet the challenges of the Internet. My own approach is somewhat different. It is borne out of three kinds of doubt: doubts about the feasibility (although not the value) of virtue-oriented educational policies, doubts about the basic validity of virtue theoretic conceptions of knowledge, and doubts about whether the aforementioned dilemma is best resolved by attempting to change the agent as opposed to the environment in which the agent is embedded. As always, space is limited and life is short, so I will restrict my discussion to issues that I deem to be of greatest interest to the epistemological community.

Reliable Technology

Inasmuch as intellectual virtues are required for online knowledge—i.e., knowledge that we possess as a result of our interactions and engagements with the Internet—they are surely only part of a much broader and richer story that includes details about the environment in which our cognitive systems operate. In judging the role of intellectual virtue in shielding us from the epistemic hazards of the online environment, it therefore seems important to have some understanding of the actual technologies we interact with.

This is important because it helps us understand the kinds of intellectual virtue that might be required, as well as the efficacy of specific intellectual virtues in helping us believe the truth (and thus working as virtues in the first place). Internet technologies are, of course, many and varied, and it will not be possible to assess their general relevance to epistemological debates in the present commentary. For the sake of brevity, I will therefore restrict my attention to one particular technology: blockchain.

Blockchain is perhaps best known for its role in supporting the digital cryptocurrency, Bitcoin. It provides us with a means of storing data in a secure fashion, using a combination of data encryption and data linking techniques. For present purposes, we can think of a blockchain as a connected set of data records (or data blocks), each of which contains some body of encrypted data. In the case of Bitcoin, of course, the data blocks contain data of a particular kind, namely, data pertaining to financial transactions. But this is not the only kind of data that can be stored in a blockchain. In fact, blockchains can be used to store information about pretty much anything. This includes online voting records, news reports, sensor readings, personal health records, and so on.

Once data is recorded inside a blockchain, it is very difficult to modify. In essence, the data stored within a blockchain is immutable, in the sense that it cannot be changed without ‘breaking the chain’ of data blocks, and thereby invalidating the data contained within the blockchain. This property makes blockchains of considerable epistemic significance, because it speaks to some of the issues (e.g., concerns about data tampering and malign forms of information manipulation) that are likely to animate epistemological debates in this area.
This does not mean, of course, that the information stored within a blockchain is guaranteed to be factually correct, in the sense of being true and thus yielding improvements in epistemic standing. Nevertheless, there are, I think, reasons to regard blockchain as an important technology relative to efforts to make the online environment a somewhat safer place for would-be knowers. Consider, for example, the title of the present article. Suppose that we wanted to record the fact that a person known as Paul Smart—that’s me—wrote an article with the title:

(Fake?) News Alert: Intellectual Virtues Required for Online Knowledge!

We can incorporate this particular piece of information into a blockchain using something called a cryptographic hash function, which yields a unique identifier for the block and all of its contents. In the case of the aforementioned title, the cryptographic hash (as returned by the SHA256 algorithm2) is:

7147bd321e79a63041d9b0a937954976236289ee4dc6f8e97533f86083a8532

Now suppose that someone wants to alter the title, perhaps to garner support for an alternative argumentative position. In particular, let’s suppose they want to claim that the title of the article is:

Fake News Alert: Intellectual Virtues Required for Online Knowledge!

From an orthographic perspective, of course, not much has changed. But the subtlety of the alteration is not something that can be used to cause confusion about the actual wording of the original title—the title that I intended for the present article. (Neither can it be used to cast doubt about the provenance of the paper—the fact that the author of the paper was a person called Paul Smart.) To see this, note that the hash generated for the ‘fake’ title looks nothing like the original:

cc05baf2fa7a439674916fe56611eaace55d31f25aa6458b255f8290a831dde4

It is this property that, at least in part, makes blockchains useful for recording information that might otherwise be prone to epistemically malign forms of information manipulation. Imagine, for the sake of argument, that climatological data, as recorded by globally distributed sensors, was stored in a blockchain. The immutability of such data makes it extremely difficult for anyone to manipulate the data in such a way as to confirm or deny the reality of year-on-year changes in global temperature. Neither is it easy to alter information pertaining to the provenance of existing data records, i.e., information about when, where, and how such data was generated.

None of this should delude us into thinking that blockchain technology is a panacea for Internet-related epistemic problems—it isn’t! Neither does blockchain obviate the need for

---

agents to exercise at least some degree of intellectual virtue when it comes to the selection and evaluation of competing data streams. Nevertheless, there is, I think, something that is of crucial epistemological interest and relevance here—something that makes blockchain and other cybersecurity technologies deserving of further epistemological attention. In particular, such technologies may be seen as enhancing the epistemic safety of the online environment, and thus perhaps reducing the need for intellectual virtue.

In this sense, the epistemological analysis of Internet technologies may be best approached from some variant of modal epistemology—e.g., epistemological approaches that emphasize the modal stability of true beliefs across close possible worlds (Pritchard, 2009, chap. 2). But even if we choose to countenance an approach that appeals to issues of intellectual virtue, there is still, I suggest, a need to broaden the analytic net to include technologies that (for the time being at least) lie beyond the bounds of the individual cognitive agent.

**Safety in Numbers**

“No epistemic perspective,” Richard writes, “the most salient dimension of the Internet is that it is an information space” (Heersmink, 2018, p. 5). Somewhat surprisingly, I disagree. Although it is obviously true that the Internet is an information space, it is not clear that this is its most salient feature, at least from an epistemological standpoint. In particular, there is, I suggest, a sense in which the Internet is more than just an information space. As is clear from the explosive growth in all things social—social media, social networks, social bots, and so on—the Internet functions as a social technology, yielding all manner of opportunities for people to create, share and process information in a collaborative fashion. The result, I suggest, is that we should not simply think of the Internet as an information space (although it is surely that), we should also view it as a social space.

Viewing the Internet as a social space is important because it changes the way we think about the epistemic impact of the Internet, relative to the discovery, production, representation, acquisition, processing and utilization of knowledge. Smart (in pressb), for example, suggests that some online systems function as knowledge machines, which are systems in which some form of knowledge-relevant processing is realized by a socio-technical mechanism, i.e., a mechanism whose component elements are drawn from either the social (human) or the technological realm.

An interesting feature of many of these systems is the way in which the reliability (or truth-conducive) nature of the realized process is rooted in the socio-technical nature of the underlying (realizing) mechanism. When it comes to human computation or citizen science systems, for example, user contributions are typically solicited from multiple independent users as a means of improving the reliability of specific epistemic outputs (Smart, in pressb; Smart and Shadbolt, in press; Watson and Floridi, 2018). Such insights highlight the socially-distributed character of at least some forms of online knowledge production, thereby moving us beyond the realms of individual, agent-centric analyses.
On a not altogether unrelated note, it is important to appreciate the way in which social participation can itself be used to safeguard online systems from various forms of malign intervention. One example is provided by the Google PageRank algorithm. In this case, any attempt to ‘artificially’ elevate the ranking assigned to specific contributions (e.g., a user’s website) is offset by the globally-distributed nature of the linking effort, coupled with the fact that links to a specific resource are themselves weighted by the ranking of the resource from which the link originates. This makes it difficult for any single agent to subvert the operation of the PageRank algorithm.

Even ostensibly non-social technologies can be seen to rely on the distributed and decentralized nature of the Internet. In the case of blockchain, for example, multiple elements of a peer-to-peer network participate in the computational processes that make blockchain work. In this way, the integrity of the larger system is founded on the collaborative efforts of an array of otherwise independent computational elements. And it is this that (perhaps) allows us to think of blockchain’s epistemically-desirable features as being rooted in something of a ‘social’ substrate.

All of this, I suggest, speaks in favor of an approach that moves beyond a preoccupation with the properties of individual Internet users. In particular, there seems to be considerable merit in approaching the Internet from a more socially-oriented epistemological perspective. It is easy to see the social aspects of the Internet as lying at the root of a panoply of epistemic concerns, especially when it comes to the opportunities for misinformation, deception, and manipulation. But in light of the above discussion, perhaps an alternative, more positive, take on the Internet (qua social space) starts to come into sharper focus. This is a view that highlights the way in which certain kinds of online system can work to transform a ‘vice’ into a ‘virtue,’ exploiting the social properties of the Internet for the purposes of dealing with reliability-related concerns.

**Filter Bubblicious**

Search engines form one of the focal points of Richard’s analysis, and, as with previous work in this area, Richard finds at least some aspects of their operation to be highly problematic. A particular issue surfaces in respect of personalized search. Here, Richard’s analysis echoes the sentiments expressed by other epistemologists who regard personalized search algorithms as of dubious epistemic value.

In fact, I suspect the consensus that has emerged in this area fails to tell the whole story about the epistemic consequences of personalized search. Indeed, from a virtue epistemological position, I worry that epistemologists are in danger of failing to heed their own advice—prematurely converging on a particular view without proper consideration of competing positions. In my new-found role as the virtue epistemologist’s guardian angel (or should that be devil’s advocate?), I will attempt to highlight a couple of reasons why I think more empirical research is required before we can say anything useful about the epistemological impact of personalized search algorithms.
My first worry is that our understanding about the extent to which search results and subsequent user behavior is affected by personalization is surprisingly poor. Consider, for example, the results of one study, which attempted to quantify the effect of personalization on search results (Hannak et al., 2013). Using an empirical approach, Hannak et al. (2013) report a demonstrable personalization effect, with 11.7% of search results exhibiting differences due to personalization. Interestingly, however, the effect of personalization appeared to be greater for search results with lower rankings; highly ranked results (i.e., those appearing at the top of a list of search results) appeared to be much less affected by personalization.

This result is interesting given the observation that college students “prefer to click on links in higher positions even when the abstracts are less relevant to the task at hand” (Heersmink, 2018, p. 6). From one perspective, of course, this tendency looks like a vice that jeopardizes the epistemic standing of the individual user. And yet, from another perspective, it looks like the preference for higher ranked search results is poised to negate (or at least reduce) the negative epistemological effects of personalized search. What we seem to have here, in essence, is a situation in which one kind of ‘intellectual vice’ (i.e., a tendency to select highly-ranked search results) is playing something of a more positive (virtuous?) role in mitigating the negative epistemological sequelae of a seemingly vicious technology (i.e., personalized search).

None of this means that the epistemic effects of personalized search are to the overall benefit of individual users; nevertheless, the aforementioned results do call for a more nuanced and empirically informed approach when considering the veritistic value of search engines, as well as other kinds of Internet-related technology.

A second worry relates to the scope of the epistemological analysis upon which judgements about the veritistic value of search engines are based. In this case, it is unclear whether analyses that focus their attention on individual agents are best placed to reveal the full gamut of epistemic costs and benefits associated with a particular technology, especially one that operates in the socio-technical ecology of the Internet. To help us understand this worry in a little more detail, it will be useful to introduce the notion of mandevillian intelligence (Smart, in press; Smart, in press).

Mandevillian intelligence is a specific form of collective intelligence in which the cognitive shortcomings and epistemic vices of the individual agent are seen to yield cognitive benefits and epistemic virtues at the collective or social level of analysis, e.g., at the level of collective doxastic agents (see Palermos, 2015) or socio-epistemic systems (see Goldman, 2011). According to this idea, personalized search systems may play a productive role in serving the collective cognitive good, providing a means by which individual vices (e.g., a tendency for confirmation bias) are translated into something that more closely resembles an epistemic virtue (e.g., greater cognitive coverage of a complex space of thoughts, ideas, opinions, and so on). Consider, for example, the way in which personalized search may help to focus individual attention on particular bodies of information, thereby restricting access to a larger space of ideas, opinions, and other information.
While such forms of ‘restricted access’ or ‘selective information exposure’ are unlikely to yield much in the way of an epistemic benefit for the individual agent, it is possible that by exploiting (and, indeed, accentuating) an existing cognitive bias (e.g., confirmation bias), personalized search may work to promote cognitive diversity, helping to prevent precipitant forms of cognitive convergence (see Zollman, 2010) and assisting with the epistemically optimal division of cognitive labor (see Muldoon, 2013). This possibility reveals something of a tension in how we interpret or evaluate the veritistic value of a particular technology or epistemic practice. In particular, it seems that assessments of veritistic value may vary according to whether our epistemological gaze is directed towards individual epistemic agents or the collective ensembles in which those agents are situated.

The Necessity of Virtue

As Richard notes, virtue epistemology is characterized by a shift in emphasis, away from the traditional targets of epistemological analysis (e.g., truth, justification and belief) and towards the cognitive properties of would-be knowers. “Virtue epistemology,” Richard writes, “is less concerned with the nature of truth and more concerned with the cognitive character of agents” (Heersmink, 2018, p. 2). This is, no doubt, a refreshing change, relative to the intellectual orientation of traditional philosophical debates.

Nevertheless, I assume that virtue epistemologists still recognize the value and priority of truth when it comes to issues of epistemic evaluation. Someone who holds false beliefs is not the possessor of knowledge, and this remains the case irrespective of whatever vices and virtues the agent has. In other words, it does not matter how careful, attentive and assiduous an agent is in selecting and evaluating information, if what the agent believes is false, they simply do not know.

What seems to be important in the case of virtue epistemology is the role that intellectual virtue plays in securing the truth of an agent’s beliefs. In particular, the central feature of virtue epistemology (at least to my mind) is that the truth of an agent’s beliefs stem from the exercise of intellectual virtue. It is thus not the case that truth is unimportant (or less important) when it comes to issues of positive epistemic standing; rather, it is the role that intellectual virtue plays in establishing the truth of an agent’s beliefs. An agent is thus a bona fide knower when they believe the truth and the truth in question is attributable to some aspect of their cognitive character, specifically, a cognitive trait (virtue responsibilism) or cognitive faculty (virtue reliabilism).

What then makes something a vice or virtue seems to be tied to the reliability of token instantiations of processes that are consistent with an agent’s cognitive character. Intellectual virtues are thus “cognitive character traits that are truth-conducive and minimalise error” (Heersmink, 2018, p. 3), while intellectual vices are characterized as “cognitive character traits that are not truth-conducive and do not minimalise error” (Heersmink, 2018, p. 3). It is this feature of the intellectual virtues—the fact that they are, in general, reliable (or give rise
to reliable belief-relevant processes)—that looks to be important when it comes to issues of epistemic evaluation.

So this is what I find problematic about virtue theoretic approaches to knowledge. (Note that I am not an epistemologist by training, so this will require a generous—and hopefully virtue-inspiring swig—of the old intellectual courage.) Imagine a state-of-affairs in which the Internet was (contrary to the present state-of-affairs) a perfectly safe environment—one where the factive status of online information was guaranteed as a result of advances in cyber-security techniques and intelligent fact-checking services. Next, let us imagine that we have two individuals, Paul and Sophia, who differ with respect to their cognitive character. Paul is the least virtuous of the two, unreflectively and automatically accepting whatever the Internet tells him. Sophia is more circumspect, wary of being led astray by (the now non-existent) fake news.

Inasmuch as we see the exercise of intellectual virtue as necessary for online knowledge, it looks unlikely that poor old Paul can be said to know very much. This is because the truth of Paul’s beliefs are not the result of anything that warrants the label ‘intellectual virtue.’ Paul, of course, does have a lot of true beliefs, but the truth of these beliefs does not stem from the exercise of his intellectual virtues—if, indeed, he has any. In fact, inasmuch as there is any evidence of virtue in play here, it is probably best attributed to the technologies that work to ensure the safety of the online environment. The factive status of Paul’s beliefs thus has more to do with the reliability of the Internet than it does with the elements of his cognitive character.

But is it correct to say that Paul has no online knowledge in this situation? Personally, I do not have this intuition. In other words, in a perfectly safe environment, I can see no reason why we should restrict knowledge attributions to agents whose beliefs are true specifically as the result of intellectual virtue. My sense is that even the most unreflective of agents could be credited with knowledge in a situation where there was no possibility of them being wrong. And if that is indeed the case, then why insist that it is only the exercise of intellectual virtue that underwrites positive epistemic standing?

After all, it seems perfectly possible, to my mind, that Sophia’s epistemic caution contributes no more to the minimization of error in an epistemically benign (i.e., safe) environment than does Paul’s uncritical acceptance. (In fact, given the relative efficiency of their doxastic systems, it may very well be the case that Sophia ends up with fewer true beliefs than Paul.) It might be claimed that this case is invalidated by a failure to consider the modal stability of an agent’s beliefs relative to close possible worlds, as well as perhaps their sensitivity to counterfactual error possibilities. But given the way in which the case is characterized, I suggest that there are no close possible worlds that should worry us—the cybersecurity and fact checking technologies are, let us assume, sufficiently robust as to ensure the modal distance of those worrisome worlds.

One implication of all this is to raise doubts about the necessity of intellectual virtue, relative to our conceptual understanding of knowledge. If there are cases where intellectual virtue is
not required for positive epistemic standing, then intellectual virtue cannot be a necessary condition for knowledge attribution. And if that is the case, then why should intellectual virtue form the basis of an approach that is intended to deal with the epistemic shortcomings of the (contemporary) Internet?

Part of the attraction of virtue epistemology, I suspect, is the way in which a suite of generally reliable processes are inextricably linked to the agent who is the ultimate target of epistemic evaluation. This linkage, which is established via the appeal to cognitive character, helps to ensure the portability of an agent’s truth-tracking capabilities—it helps to ensure, in other words, that wherever the agent goes their reliable truth-tracking capabilities are sure to follow.

However, in an era where our doxastic systems are more-or-less constantly plugged into a reliable and epistemically safe environment, it is not so clear that agential capabilities are relevant to epistemic standing. This, I suggest, raises doubts about the necessity of intellectual virtue in securing positive epistemic status, and it also (although this is perhaps less clear) encourages us to focus our attention on some of the engineering efforts (as opposed to agent-oriented educational programs) that might be required to make the online world an epistemically safer place.

Conclusion

What, then, should we make of the appeal to virtue epistemology in our attempt to deal with the epistemic hazards of the Internet. My main concern is that the appeal to virtue epistemology (and the emphasis placed on intellectual virtue) risks an unproductive focus on individual human agents at the expense of both the technological and social features of the online world. This certainly does not rule out the relevance of virtue theoretic approaches as part of our attempt to understand the epistemic significance of the Internet, but other approaches (e.g., modal reliabilism, process reliabilism, distributed reliabilism, and systems-oriented social epistemology) also look to be important.

Personally, I remain agnostic with regard to the relevance of different epistemological approaches, although I worry about the extent to which virtue epistemology is best placed to inform policy-related decisions (e.g., those relating to education). In particular, I fear that by focusing our attention on individual agents and issues of intellectual virtue, we risk overlooking some of the socio-epistemic benefits of the Internet, denigrating a particular technology (e.g., personalized search) on account of its failure to enhance individual knowledge, while ignoring the way a technology contributes to more collective forms of epistemic success.

In concluding his thought-provoking paper on virtue epistemology and the Internet, Richard suggests that “there is an important role for educators to teach and assess [intellectual] virtues as part of formal school and university curricula, perhaps as part of critical thinking courses” (Heersmink, 2018, p. 10). I have said relatively little about this particular issue in the present paper. For what it’s worth, however, I can see no reason to object to the general
idea of Internet-oriented educational policies. The only caveat, perhaps, concerns the relative emphasis that might be placed on the instillation of intellectual virtue as opposed to the inculcation of technical skills, especially those that enable future generations to make the online world a safer place.

No doubt there is room for both kinds of pedagogical program (assuming they can even be dissociated). At the very least, it seems to me that the effort to resolve a problem (i.e., engineer a safer Internet) is just as important as the effort to merely cope with it (i.e., acquire a virtuous cognitive character). But, in any case, when it comes to education and learning, we should not lose sight of the fact that the Internet is itself something that is used for educational purposes. Perhaps, then, the more important point about education and the Internet is not so much the precise details of what gets taught, so much as the issue of whether the Internet (with all its epistemic foibles) is really the best place to learn.

Contact details: ps02v@ecs.soton.ac.uk

References


Smart, P. R. (in pressb). Knowledge machines. The Knowledge Engineering Review.


