Patterns

CellPress
OPEN ACCESS

Perspective

Thinking about Trust: People, Process, and Place

Stephen Marsh,[1,*] Tosan Atele-Williams,[1] Anirban Basu,[2] Natasha Dwyer,[3] Peter R. Lewis,[4] Hector Miller-Bakewell,[5] and Jeremy Pitt[6]

[1]Faculty of Business and IT, Ontario Tech University, Oshawa, ON L1G 0C5, Canada
[2]Foundations of Software Systems Group, Department of Informatics, University of Sussex, Falmer, Brighton, East Sussex BN1 9QJ, UK
[3]College of Arts and Education, Victoria University, PO Box 14428, Melbourne, VIC 8001, Australia
[4]Aston Lab for Intelligent Collectives Engineering (ALICE), Department of Computer Science, Aston University, Birmingham B4 7ET, UK
[5]Department of Computer Science, University of Oxford, Wolfsson Building, Parks Road, Oxford OX1 3QD, UK
[6]Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, UK
*Correspondence: stephen.marsh@ontariotechu.ca
https://doi.org/10.1016/j.patter.2020.100039

THE BIGGER PICTURE We discuss the implications of thinking about trust in different disciplines, including artificial intelligence (AI), data science in general, decision making, and user interaction. In particular, the key point to take away is that trust is a vital component to the computational system when it interacts with humans (which is always) and that understanding it better allows us to craft better systems and interactions. We also extend the debate about AI and trust/trustworthiness.

Production: Data science output is validated, understood, and regularly used for multiple domains/platforms

This brief paper is about trust. It explores the phenomenon from various angles, with the implicit assumptions that trust can be measured in some ways, that trust can be compared and rated, and that trust is of worth when we consider entities from data, through artificial intelligences, to humans, with side trips along the way to animals. It explores trust systems and trust empowerment as opposed to trust enforcement, the creation of trust models, applications of trust, and the reasons why trust is of worth.

First Thoughts as an Introduction

It is not always obvious. Indeed, there is often little in the way our societies work that would suggest the importance of trust, until one starts to pay attention, and then, there it is. Luhmann[1] points out that the complexity of everyday life might indeed be reduced by trust, since it allows us to take certain things as given: that people do not usually try to harm us, or that the news is accurate (our examples, not Luhmann's), for example. As Bok[2] notes, societies without trust will not be successful (see also Lagenspetz[3]). It is something most of us use every day, often without even thinking too hard, and much of the time within split seconds of needing to use it. It can be influenced by seemingly unrelated things happening around us, often without our realizing, and it matters.

Trust matters because we use it to make decisions about things. At the most prosaic level, we might use it to make a decision about buying something, online or in person—something that research of trust in eCommerce addresses (e.g., see Noorian et al.[4] and McKnight et al.[5]) We might use it to decide whether to use Uber or Lyft, and when we have done that, we might use it to decide which driver we would like to pick us up.[6] We might use it in ubiquitous systems,[7] and we might use it in health care.[8] We might use it to decide whether or not we believe what we read on Facebook, Twitter, online newspapers, or what we hear from cable news.[9] We might use it when meeting people for the first time, or hiring a babysitter, or building and rebuilding relationships or collaborations.[10,11] We might use it when thinking of our leaders.[12,13] We might use it to determine who to work with to manage emergencies, or even to determine how much supervision to give in differing circumstances for different people or organizations. We might use it in security models.[14] We may (or indeed may not, cf. Cook et al.[15]) use it simply to determine who, and how much, to cooperate with in a specific context.[16] To summarize, trust matters. For the data scientists among us, we might use it to rate and rank the data and information we are using, its quality and veracity, or its provenance, for instance. If we design or use data collection methods and tools, we may rank these in terms of their efficacy and trustworthiness. We may use it in the systems we design and use it to scan and sense, for instance, in the worlds of the Internet of Things or sensor networks.

This paper explores, in an unstructured way, some of the ways in which trust does matter, where it can be and is applied, how in some instances it can be measured, subjectively or otherwise, and where and why it may be useful to consider in the sociotechnical world we live in. It is unstructured because it does not have a single story to tell, or a particular message that is universal (other than trust is something to which we should be paying attention). It is further unstructured because trust is such a fascinating and multifaceted phenomenon that is so highly contextual

as to be different even from moment to moment for the same person (or agent). There is much more that can be and is being said about trust, and there is no doubt that some of what is here is incomplete, some is missed, and some, as our own work (and in some instances speculation), will not be agreed with by all readers. Nevertheless, we write with the aim that the reader will gain some new insights, questions, or inspiration from what is here and explore further in the literature we cite as well as the extensive literature that exists.

Thoughts about Why Trust Matters

The questions about why trust matters and how it works have been asked and studied for millennia. That given, you may be forgiven for thinking that the problems are all solved now. But here is the thing: people are not predictable, the things they do sometimes surprise us. Moreover, the way in which we work together is changing constantly and is nothing like it was only a few years ago.

While the onset of things like big data and artificial intelligence (AI) give us potentially deep insights into the ways in which people behave in differing circumstances, these are still people and they still work, play, interact, and share in ways that matter to them. And they use trust to do it. Working, shopping, playing, and building relationships online simply adds a different context to the ways in which they do things. The fact that they are removed from their "monkey brain" comfort zones in terms of what they see, how they interpret it, and being unable necessarily to tell if they are speaking to a dog or a person online simply adds new dimensions to the ways in which trust matters.

However, being taken out of the "monkey brain" comfort zone only works if the necessary cognitive skills for "chimp management"[17] and critical thinking have been properly developed, with corresponding "defense mechanisms" to resist confirmation bias[18] and digital sortition.[19] Otherwise, the opportunity to exploit basic human psychology has been seized and is still unregulated: see, for example, how social media platforms have used habit-forming techniques to increase usage and screen time,[20] or softening up intended recipients of a message with preliminary attention-catching hooks that predispose the recipient to agreeing with the message when they finally encounter it.[21]

This manipulation becomes absolutely critical as one moves from commerce and marketing to, for example, politics and public health. So, theories concerning the social construction of reality have been formulated in sociology for years;[22] however, how that social construction occurs also needs to be understood. A new theory from social psychology, RTSI (Regulatory Theory of Social Influence),[23] suggests that as well as there being social influence from source to target, there is also social influence by target from source—that potential targets of social influence actively seek out potential sources by whom to be influenced.

Given the systematic and willful denigration of scientific institutes, educational establishments, and independent press,[24] those institutions that could act as guarantors of trustable information and mediators of new information have been overwhelmed by a flood of misinformation. This has left people free to pick their own facts—to seek out their preferred sources by whom to be influenced; RTSI therefore partially explains how it been possible for populist politicians to gaslight entire nations, which has led to both a collapse in, and displacement of, trust, i.e., away from traditional institutions requiring engagement and thought, and toward whoever can commandeer attention with the foghorn of social media. This has had severe negative consequences in both public health[25] and political health.[26]

Moreover, trust matters in the digital society and the information economy because of their grounding in data and the transparent and equitable use of data. The need to democratize big data, and not let it be the preserve of corporate, scientific, or political elites, has been identified for some time. But the responsible (or even ethical) use of big data is still an open issue and has become increasingly important. Some of the issues to address here include a mutual agreement not to weaponize data politically (e.g., to subvert the democratic process by gerrymandering[27]) and to collect all the data, not just because some of it may be ideologically inconvenient or out of an "obtuse desire to remain in ignorance."[28]

So, trust still matters. But how can we understand trust in this context?

Thoughts about How We Can Understand Trust

For all that we have studied it for so long, there are still important questions that we can ask about how it might be useful to us now. These might include looking more closely at the questions we alluded to above. Can we trust these data? Can we trust this system? What does that even mean, and what are we trusting it for?

Many years ago, Marsh asked the question: how might we represent trust in a way that computational systems (back then, specifically, autonomous agents) might be able to reason both about and with trust as a concept.[16]

Since then, we have built on the simpler understanding of trust, and the accompanying formalization, to include consideration of distrust, mistrust and untrust,[29] forgiveness[30] and regret,[31] and the application of trust in things from Computer Supported Cooperative Work[32] to digital government[33] and computer security.[14]

Thoughts about Trust Systems

Our work has led us to the concept of trust systems, because we have begun to perceive that technology is in many cases an integral part of the way in which people work and play, and moreover that technology can often help people do these things. A trust system is a system in which trust is an enabler, but it means slightly more than that. Because systems are systems, not standalone applications, it is where computational tools, humans, and trust reasoning capabilities come together to accomplish something where trust is the enabling factor. We consider it a triad of three Ps: people, process, and place.

People

We consider that people are central to any endeavor in which computational systems are deployed. Further, since we are considering the notion of trust here, we necessarily must also consider how it relates to humans. At any point in a trust system, then, people are involved: in any or indeed all of the decision making of the system itself, guidance of the system, or the impact of decisions made and actions taken as a result of the trust deliberations, and potentially more.

Any trust system, then, requires that people are considered, even if not physically present, as part of the system itself.

Process

A Trust System works with trust in order to achieve something: a decision, a recommendation, the prelude to or necessary requirements for an action, and so on. In order to work with trust, it is necessary to understand it in some way: Trust models and algorithms are, in their own contexts, a way of doing this. Thus, for want of a better word, process describes the way in which the system makes its deliberations: its algorithm, model, and representations.

What Does This Look like? There are many different models of trust in existence, and more arrive seemingly on a daily basis, such that there is little utility in listing them here. Indeed, while there have been excellent review papers in the past (see for instance Jösang et al.[34]), the field is fast moving, and trust is, as we have noted, so highly contextual that such reviews are useful for a limited time and as useful historical records. The interested reader is referred to the many online sources for academic work that exists for models that matter to their own field. Some (many) apply to eCommerce and buying or selling online, with or without agents to help. Some apply to the user interface between human and machine. Some are deployed to determine how much credit you might get, or if a specific purchase is being made by you. Others may be more esoteric and apply in online stock trading. Some, such as the Chinese Social Credit system or online models of trust and reputation for everyone around us, are being deployed to calculate the behavior and thus reputation of others as if it is the root of all trust (it is not, but that is a topic for another time).

As we noted above, in a short paper such as this, there is little benefit to listing the models that do exist, not least because we would also not do them credit. The reason that there are so many has less to do with reinventing the wheel every time a new one comes out (although to be fair, there is a degree of this) and more to do with the fact that trust is so contextual, as we will shortly see, that generic trust models are too vague to be immediately useful while more detailed specific ones are too specifically aimed at a particular context to be useful in others (getting your agent to bid for a camera online is not the same as choosing a babysitter).

Place

Trust models are necessarily imperfect representations of the rich phenomenon humans understand as trust (to be honest, since trust is rather challenging to isolate, and its workings in humans are hard to define properly, what we can see from the computational models is a representation of what trust might look like, and since we can refine or adapt this contextually, imperfect is perhaps a little strong). This does not make them less useful, but we need to consider when and how they work, and what limits they have. In addition, trust is a highly contextual phenomenon in humans as well as computational systems. It is necessary to consider the environmental context that the system finds itself in when considering trust, because this context changes all the time and thus so does the potential information that the system uses.

Place refers to the context of the system. This context defines when the system is considering its options, where the system can work, what assumptions are made about the information or actions it considers, and the limits of its abilities.

Thinking about Questions

Given that we can think of a trust system in this way, there remains the consideration of how a trust decision is made. It is worth mentioning here that this is how it works for us, but that does not mean that it works this way for anyone who has studied trust in specific contexts. That said, we refer the interested reader (and watcher) to Onora O'Neill's work regarding trust in public life.[12,35] One of the examples O'Neill uses in her talk is that of an elementary school teacher being trusted to teach a class but not to drive the minibus (for different reasons).

This is important, so it bears repeating again (and again): trust is a highly contextual, subjective judgment that is made in the light of available evidence and also need, among other things. The school teacher example is informative. We may trust individuals in one context, for a specific task, but not in others. More prosaically, we use rather imprecise language to talk about trust, let alone to define it properly.[36,37]

We come then to the two most important questions about trust: How much do we have? and How much do we need? The first is answered by our process, and the second, in combination, by person and place. Clearly, there needs to be a subject we are considering, and this trustee is considered in process and place. The end result of these calculations, if we can call them that, is two measures. If the How much do we have is more than the How much do we need, we have a trusting relationship for that context (I trust my brother to drive me to the airport), but if what we have is less than we need, we do not (I do not trust him to fly the plane).

It is particularly important to note here that just because I do not have enough trust in the context of flying the plane, it does not imply that I distrust, or do not trust, my brother.[31]

As Christian Jensen has pointed out (personal communication), there is actually a third consideration we can get to simply because we use measures in our computational trust mechanisms: Who or which is best? That is to say, since we can measure how much we might trust someone or something, we are now able to compare them and determine which we trust the most, or even which is more trusted in context, which may sometimes give counter-intuitive results but nonetheless indicates something of value.

Thinking about Empowerment and Enforcement

Trust Systems are most useful in determining and supplying foreground trust data,[38,39] whereby the technology helps provide data that the human can use to make trusting decisions, without necessarily making those decisions itself. We refer to this as trust empowerment rather than trust enforcement.[38] This is an important distinction, so we take a moment to discuss it further.

Trust Enforcement: "Trust This Much because We Say So"

Reputation is surely a good thing to have, but it is not the only thing that matters in trust. In trust enforcement, we, as traditional knowledge engineers or user interface designers, or simply professionals or knowledge aggregators putting together a list of reputation attributes, do this: grab reputation statistics from the world around the object in question, put them together

through some form of algorithm or formulas, which we protect as much as possible from the outside world, and present it to the subject. We do this and say "this is the trust value/reputation/rank of the object. You can trust it this much. These are the same values for other objects, you can trust them this much."

The result is that we the users are given the things we should trust in the order that they are trustable. Examples of this include almost every reputation system that exists, including Google's own Pagerank,[40] which, while constantly rejigged, essentially does this thing for web objects. It also includes things like social credit systems (for a particularly engaging version of which, Whuffie, see Doctorow[41]) and the growing list of websites that seek to tell us how trustworthy other people around us are.

Needless to say, there are problems here. The first and most obvious is that the systems are not telling us anything about trust per se, although they claim to, at least in some cases. What they are giving us, if we pay attention, is a measure of reputation that can, in most circumstances, be translated somehow to trustworthiness for a given context (a web search, a dentist's capabilities, the publication prowess of an academic, or how far one can cycle in the case of the Eddington number). While interesting, and even useful when one is considering trusting the object for some purpose, it is not how much you should trust that object. Unfortunately, it stands the chance of becoming exactly that number. Given that trust is something that individuals give, being told what to give in any circumstance is not particularly informative, especially since the reasons for that instruction are hidden in the algorithms used, which may be inherently biased in any number of ways, nefarious or innocent.

We believe that empowerment is a much more human-oriented way to manage the trust questions.

Trust Empowerment: "Here Is the Evidence You May (Want to) Use"

In trust empowerment, evidence is still gathered for trust-based decisions, but explicit recommendations are not made. It is of course the case that the evidence gathered may be implicitly biased—the systems gathering the data gather it from specified sources and may choose to discount others—but the design of a good system for trust empowerment would take into account the evidence that the potential truster requests or requires in order to make that trust decision. The end result is a technology that presents the truster with a set of evidence, potentially ranked in terms of its own trustworthiness (which we get to shortly), and simple questions: "Who do you want to trust to do this?" or "Which one do you trust more?", for example.

The difference between empowerment and enforcement in this instance is striking: the person taking on the risk (the truster) is informed and aware of what is being asked of them and has the choice of what happens next.

Foreground trust[38] is an example of this: a user online, considering other humans online, is presented with the evidence needed to make trust-based decisions about those other people in a way that matters to them. This means that the evidence given is that which matters to the person making the trust decision. A specific value for trust is not given, because in any trust decision, the trust that is given is that of the individual truster, not the evidence gathering process. Why is this important? If we consider that trust is putting oneself into a situation or risk, then it should go without saying that the person taking on the risk should be the one who makes the decision in as informed a way as possible (the technology or the algorithm, in any case, takes on no risk).

We mentioned PageRank above. How might this look in an empowered version? There is a process to be followed here. The first is that the searcher is expected to provide some kind of information about the things that matter to them (this needs only to be done rarely, but in principle can be done every time a search is made, because different searches are different contexts). The second is the search itself, and the result is presented, much as it is today, as a ranked list of possible hits, with one exception: the reasons why they are hits are provided, or how the results relate to the things that matter, including matching keywords, matches for the kinds of pages that refer to this one, or matches for relevant authors, for instance. Indeed, the priorities of such metrics can be as personalizable as the searches themselves. Is this more work for the searcher? Potentially (probably!). But informed decision making can also be achieved with a set of reasons why pages are ranked above others, or products are recommended above others, and so forth, which can come from similar sources. There is always the problem of filter bubbles here, which is a widely researched area (see for instance, Spohr[42] and Dillahunt et al.,[43] but this barely scratches the surface, and which itself is open to discussion[44]). We believe that part of the beauty of systems like this is their ability to challenge the ways in which people think. Thus, careful design is required to ensure that the searcher is not rewarded simply with compliant results, but also with results that encourage breadth of mind. This is an ongoing challenge for all of the technology that is currently helping people.

Thoughts about Applications

We have come a long way in a short time. At the start of this paper, we briefly mentioned some of the ways in which trust can be put into play by those who work with data. In this section, we explore in more detail the concepts of trust in data and what it means, and very briefly explore trust of data, which is something rather different. We also begin an exploration of the human-trust-AI issue that is becoming increasingly important.

In all of these instances, the two questions we referred to above are important. How the values of How much do you have? and How much do you need? are calculated is different in almost all cases, however, and we discuss that briefly here.

Trust in Data

In many ways, this is the simplest of our problems, and one that has been extensively addressed (see for example, Penner and Klahr[45] and Aman[46]). This is because we inherently care about the data we are using and how useful it might be. Thus, when we consider data from sensor networks[47] or shared among mobile networks,[48] we relate the validity of the data to the integrity of the sources. This is, or should be, no different from trusting information given to journalists,[49,50] on social media sites,[51] or other kinds of information.[9] The point here is that we need to examine where the data came from, the routes the data may have taken, the efficacy of the collection mechanisms, the sensors themselves (expected battery life, harsh environments, gaps in data, and so forth).

Every one of these considerations is a data point in itself for our two questions, and in particular for our first: How much do you

have? In every single case, a system can calculate the answer to this question based on what matters to the individual because the individual can tell the system what is important to them about the data (hence trust empowerment). To be fair, for human beings it is often hard to decide what matters, how much, and when, although in some of our work, we have explored different ways of eliciting this information.[38,52]

The second question, How much do you need?, is answered as straightforwardly by considering what the data are to be used for, what is at risk, how much it matters, and so on. Thus, for data being used to determine the amount of range left on an electric vehicle in the summertime, we can be a little less careful than with data from intensive care systems (or even the same vehicle in the wintertime when temperatures are −30°C or so and running out of power is a little more problematic!). Context matters, and the second question aims to capture it.

Trust in data can change because the answers to our two questions can change. This changes how we perceive the data, how we might use it, when we might use it differently, and what we might use it for, even if the data are the same from one moment to the next.

Trust of Data

A slightly more complex question puts into play the fact that a single piece of data can be perceived as an entity in and of itself. This is the concept of smart data,[53,54] as well as, ultimately, the premise behind information security (although generally with less autonomy of data). The question is simple: does the data trust the requester enough to share itself?

While again the answer to this question comes down to answering our regular two questions, there is sometimes a simpler way to explore this topic. Indeed, if we expect that data belong to someone, we may say that the answer is a proxy for whether or not the data owner trusts the requester. In this case, what matters to the owner is what matters to the data. As Maurer[55] points out though, data may well have interesting heredity and mixed up ownership, and this can complicate the problem somewhat. That said, what we may say the data care about is: how much, who, what for, how long, and why? The work of Behrooz[52] explored eliciting preferences from data owners on mobile devices for how health data might be shared and for what purpose, for instance.

With more autonomy of data (or the systems that police and provide it), we can have much more nuanced and interesting answers to our questions. Indeed, autonomous data can be much more useful in different circumstances, because data may be opened up for previously disallowed or questioned uses when it is able to protect itself, redact itself, and police itself and its use. We are beginning to see some examples in industry of how this might look in practice (such as DataPassports and Immuta), but it remains early days, and data autonomy is in its infancy but shows great promise.

Thoughts about Artificial Intelligence

The elephant in the room, if there is to be one, is AI. There are many definitions of what AI is, and indeed many conceptions about what it can, cannot, might, and even possibly should and should not do (see for example, Brockman[56] for a good collection of essays, as well as Broussard,[57] but the list of authors discussing AI in many different ways, positive and negative, is a long one and growing). We defer to Margaret Boden for simplicity's sake, and define AI as "computers doing the sorts of things that minds can do."[58, p. 1]

Clearly, AI exists now, according to this definition, and we believe that is apropos. Cars drive us around, medical systems search for diagnoses, social media systems mine data to predict behaviors, and so on. There is a great deal of potential for the things that such systems can do, and we are probably lucky enough to be able to see more of them. The question of thinking about trust with regard to such systems is complex, though. What, exactly, are we trusting, and to what end? For the sake of brevity, we address this quickly here, but there is a great deal more to be done in this area, and the reader is referred for example to Andras et al.[59] for both discussion and further references.

From a simple point of view, there are a couple of different ways we can address this problem. The first is about understanding (and trusting) what AI tells us, trusting the decisions, the deductions, and the actions AI might make. This particular issue comes down to one of explainability. In the literature,[1,39] a positive correlation has been found between trust and the reduction of uncertainty, either in information or in user interfaces. Being able to explain automated decision making to the user may (although there is no comprehensive study on this yet) lead to the reduction of uncertainty. However, given the complexity of machine-learning models, explaining them may even contribute to higher uncertainties. Such explainability may also bring about a certain level of trust enforcement in a "believe us because we can explain ourselves (even if you are not able to understand our explanation)" fashion. The process for empowerment for trust in AI, thus, remains an open research challenge.

Almost certainly more controversial is the question of how AI may trust us. In Marsh's work[16] for example, the concept of agents trusting each other is at the forefront, and agents are not taken to be simply artificial (so humans are in the frame). Thus, we begin to think about how artificial systems may trust, as well as be trusted by, humans. To be fair, this is not actually as controversial as we might like to think, the specter of human exceptionalism aside (we return to this shortly). Self-driving cars continually check for the attention of the human driver, for instance. Simple web browsers are most objectionable when the human wants to visit a web site that has an expired certificate. Moreover, as Kaliouby[60] points out, the AI really needs to trust humans when it cannot accomplish the task it is supposed to do (like drive straight, for instance). It is at such limit points that humans, as liminal creatures, may be better than AI, but they had better also be paying attention!

It is possible, just, to argue that the trust placed in AI by humans is a proxy trust, and that in fact we are trusting the people who coded the AI. It is also possible to explore that retribution is available for AI mistakes by punishing or holding to account the individuals or the organizations that created the AI (cf. Bryson[61]). The argument here is that AI is a tool and should be treated as such. We may humbly disagree here and are beginning to explore the way in which trust and trustworthiness, combined with principles of restorative justice,[62,63] may lead us to better interactions with such systems (and each other). Of course, this all begs questions around things like animism and (human)

exceptionalism. This is not the place to hold a much deeper exploration of these issues, which are both historically troubling in the instance of exceptionalism (the reader is encouraged to consider racism, animal rights, and the environment, for instance) and fraught with difficulties of labels of "primitive" or "uncivilized" for animism.[64,65] It is, however, worth acknowledging that much more discussion needs to take place and indeed is taking place (for instance, with regard to the robots in our midst[66]).

As we said earlier, the AI and trust question is complex, and understandably laden with a great deal of emotion. This is to be expected, and much more needs to be done to advance our understanding of ourselves as well as those who may be different from us (we do not have an amazing track record here, and we should learn better[67]) before we jump to conclusions about AI. To be more succinct, it is not enough to consider AI as a threat or a promise until we have started to consider how we should behave when more complex and capable AI comes to be.

Final Thoughts about People

In this paper, we have discussed trust and what it is, and how it might be useful. We have touched on some of the ways in which it can be more specifically useful in the areas of data and AI. We want to finish with something that could be seen as unrelated but we hope will become clear.

When Alice shares her personal data with an organization for the purpose of tracking her steps and running, or Bob shares for the purpose of monitoring glucose levels online, or when Charlie shares for the purpose of filing taxes, and when Dennis shares his family's pictures on Facebook, they do it for a reason. That reason is sometimes very specific to them (for a particularly erudite exploration of self-tracking, see Neff and Nafus[68]). The reason is sometimes not quite what you might expect, and this can color what is shared (and how). There are, to be sure, legal and regulatory obligations in place for how the data are used, when, buy whom and so forth, and when we insert trust into the equation, this makes things much more interesting.

The key thing here is this: Alice, Bob, and company are in fact people. This is ultimately a political statement. It is sometimes easy to forget that this is the case when confronted with anonymized, pseudonymized, or otherwise anonymous data, but it is the case. It is important to remember for several reasons. Some are obvious. For example, people matter and can be hurt. Some are less obvious. If we consider for instance who owns the data, it is possible to get caught up in considerations of parentage, for instance,[55] which brings fascinating questions about when data grow up and become their own "owner." Finally, consider this: in order for people to share the data with us, there is a question of trust, both that of the person and that of the data itself. If either of those is lacking for the purpose, ultimately everyone may lose. As Jordan notes, the digital economy, on which all of this data about people is based, is "ultimately a vampire and must be staked by a democratised digital culture."[69, p. 171]

ACKNOWLEDGMENTS

The authors would like to thank the many colleagues and friends who have listened, questioned, and improved the thoughts that are expressed here. S.M. is partially supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (Information Transparency for Privacy through Trust RGPIN/005897-2018). To Our Mother, for the gift of breath.[70]

AUTHOR CONTRIBUTIONS

It is rare for work such as this to be undertaken in a vacuum. The work here is no exception, and is the happy product of several years of discussion, shared thoughts, excited emails, and messages back and forth, as well as physical visits when we were fortunate enough to be able to travel. There is a first (senior) author on the paper (S.M.) who is the happy scribe for all of the shared understandings that we have forged together on this particular journey and is responsible for writing the original draft. All of the named authors contributed to this paper. In particular, all authors were at various times involved in conceptualization. All authors contributed to review and editing. In addition, J.P. wrote much of the section on why trust matters. J.P., P.R.L., S.M., and N.D. were involved in supervision. N.D., A.B., H.M.-B., and T.A.-W. contributed software and resources for projects leading to understanding for this paper.

DECLARATION OF INTERESTS

A.B. works at Hitachi R&D within Hitachi Ltd. The views, opinions, and/or findings contained in this article are those of the authors. These are not related to A.B.'s work at Hitachi and should not be interpreted as an official Hitachi position, policy, or decision, unless so designated by other documentation. S.M. is a member of the Advisory Board for DataPassports, an unpaid position. The authors declare no other competing interests.

REFERENCES

1. Luhmann, N. (1979). Trust and Power (Wiley).
2. Bok, S. (1978). Lying: Moral Choice in Public and Private Life (Pantheon Books).
3. Lagenspetz, O. (1992). Legitimacy and trust. Philosophical Investigations *15*, 1–21.
4. Noorian, Z., Marsh, S., and Fleming, M. (2011). Multi-layer cognitive filtering by behavioural modeling. In Proc. 10th Int. Conf. on Autonomous Agents and Multiagent Systems, S. Tumer, P. Yolum, and P. Stone, eds. (AAMAS), pp. 871–878.
5. McKnight, D.H., Choudhury, V., and Kacmar, C. (2000). Trust in e-commerce vendors: a two-stage model. In ICIS 2000 Proceedings http://aisel.aisnet.org/icis2000/54.
6. Hawapi, M.W., Sulaiman, Z., Kohar, U.H.A., and Talib, N.A. (2017). Effects of perceived risks, reputation and electronic word of mouth (e-WOM) on collaborative consumption of uber car sharing service. In IOP Conference Series: Materials Science and Engineering, *215*, p. 012019, https://doi.org/10.1088/1757-899x/215/1/012019.
7. Krukow, K., Nielsen, M., and Sassone, V. (2008). Trust models in ubiquitous computing. Philos. Trans. R. Soc. A *366*, 3781–3793. http://rsta.royalsocietypublishing.org/content/366/1881/3781.full.pdf.
8. Sillence, E., and Briggs, P. (2008). Ubiquitous computing: trust issues for a "healthy" society. Soc. Sci. Comput. Rev. *26*, 6–12.
9. Atele-Williams, T., and Marsh, S. (2018). Towards a computational model of information trust. In IFIPTM 2018: Proceedings of IFIP International Conference on Trust Management, N. Gal-Oz and P. Lewis, eds. (Springer AICT), pp. 124–136.
10. Rempel, J.K., Holmes, J.G., and Zanna, M.P. (1985). Trust in close relationships. J. Pers. Soc. Psychol. *49*, 92–112.
11. Kaur P., Ruohomaa S., Kutvonen L. User interface for trust decision making in inter-enterprise collaborations, in: ACHI 2012: The Fifth International Conference on Advances in Computer-Human Interactions, 2012.
12. O'Neill, O. (2002). A question of trust: the BBC Reith lectures. http://www.bbc.co.uk/radio4/reith2002/.
13. Barber, B. (1983). Logic and Limits of Trust (Rutgers University Press).

14. Marsh, S., Basu, A., and Dwyer, N. (2013). Security enhancement with foreground trust, comfort, and ten commandments for real people. In Theories and Intricacies of Information Security Problems, Volume Technische Berite Nr 63, A.V.D.M. Kayem and C. Meinel, eds. (Hasso-Plattner Instituts fur Softwaresystemtechnick and der Universitat Potsdam), pp. 1–7.

15. Cook, K.S., Hardin, R., and Levi, M. (2005). Cooperation without Trust (Russell Sage Foundation).

16. Marsh, S. (1994). Formalising Trust as a Computational Concept, PhD thesis (University of Stirling). http://www.stephenmarsh.ca/pubs/Trust/PhD/Trust.pdf.

17. Peters, S. (2012). The Chimp Paradox: The Mind Management Programme for Confidence, Success and Happiness (Random House).

18. Tavris, C., and Aronson, E. (2015). Mistakes Were Made (But Not by Me) [Revised Edition] (Mariner Books).

19. Bishop, B. (2008). The Big Sort: Why the Clustering of Like-Minded America Is Tearing Us Apart (Houghton Mifflin Harcourt).

20. Eyal, N. (2014). Hooked: A Guide to Building Habit-Forming Products (Portfolio/Penguin).

21. Cialdini, R. (2016). Pre-Suasion: A Revolutionary Way to Influence and Persuade (Simon & Schuster).

22. Berger, P., and Luckmann, T. (1966). The Social Construction of Reality (First Anchor Books).

23. Nowak, A., Vallacher, R., Rychwalska, A., Roszczyńska-Kurasińska, M., Ziembowicz, K., Biesaga, M., and Kacprzyk-Murawska, M. (2020). Target in Control: Social Influence as Distributed Information Processing (Springer).

24. Oreskes, N., and Conway, E.M. (2010). Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming (Bloomsbury Press).

25. Hussain, A., Ali, S., Ahmed, M., and Hussain, S. (2019). The anti-vaccination movement: a regression in modern medicine. Cureus *10*, e2919.

26. Runciman, D. (2018). How Democracy Ends (Profile Books).

27. Daley, D. (2016). Ratf**ked: How the Democrats Won the Presidency but Lost America (W.W. Norton).

28. Lewis, M. (2018). The Fifth Risk (W.W. Norton).

29. Marsh, S., and Dibben, M.R. (2005). Trust, untrust, distrust and mistrust — an exploration of the dark(er) side. In Trust Management: Proceedings of iTrust 2005, P. Herrmann, V. Issarny, and S. Shiu, eds. (Springer Verlag), pp. 17–33.

30. Vasalou, A., Hopfensitz, A., and Pitt, J.V. (2008). In praise of forgiveness: ways for repairing trust breakdowns in one-off online interactions. Int. J. Hum. Comput. Stud. *66*, 466–480.

31. Marsh, S., and Briggs, P. (2009). Examining trust, forgiveness and regret as computational concepts. In Computing with Social Trust, Human Computer Interaction Series, J. Golbeck, ed. (Springer), pp. 9–44.

32. Jones, S., and Marsh, S. (1997). Human computer human interaction: trust in CSCW. ACM SIGCHI Bulletin *29*, 36–40.

33. Marsh, S., Patrick, A., and Briggs, P. (2006). Trust in digital government – social issues. In The Encyclopedia of Digital Government, A. Anttiroiko and M. Malkia, eds. (IGI Publishing). https://doi.org/10.4018/978-1-59140-789-8.ch224.

34. Jösang, A., Ismail, R., and Boyd, C. (2007). A survey of trust and reputation systems for online service provision. Decision Support Systems *43*, 618–644.

35. O'Neill, O. (2013). What We Don't Understand about Trust: Tedx Houses of Parliament. https://www.ted.com/talks/onora_o_neill_what_we_don_t_understand_about_trust.

36. McKnight, D.H., and Chervany, N.L. (1996). The meanings of trust. Working paper, MISRC. http://www.misrc.umn.edu/wpaper/wp96-04.htm.

37. McKnight, D.H., and Chervany, N.L. (2001). Trust and distrust definitions: one bite at a time. In Trust in Cyber-Societies, R. Falcone, M. Singh, and Y.-H. Tan, eds. (Springer-Verlag), pp. 27–54.

38. Dwyer, N. (2011). Traces of Digital Trust: An Interactive Design Perspective, PhD thesis (School of Communication and the Arts, Faculty of Arts, Education and Human Development, Victoria University).

39. Marsh, S., Noël, S., Storer, T., Wang, Y., Briggs, P., Robart, L., Stewart, J., Esfandiari, B., El-Khatib, K., Bicakci, M.V., et al. (2012). Non-standards for trust: foreground trust and second thoughts for mobile security. In Security and Trust Management. STM 2011, C. Meadows and C. Fernandez-Gago, eds. (Springer), pp. 28–39.

40. Brin, S., and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems *30*, 107–117.

41. Doctorow, C. (2003). Down and Out in the Magic Kingdom (TOR).

42. Spohr, D. (2017). Fake news and ideological polarization: filter bubbles and selective exposure on social media. Bus. Inf. Rev. *34*, 150–160.

43. Dillahunt, T.R., Brooks, C.A., and Gulati, S. (2015). Detecting and visualizing filter bubbles in Google and Bing. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '15, B. Begole and J. Kim, eds. (Association for Computing Machinery), pp. 1851–1856.

44. Borgesius, F.Z., Trilling, D., Moeller, J., Bodó, B., de Vreese, C.H., and Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review. J. Internet Regul. *5*, https://doi.org/10.14763/2016.1.401.

45. Penner, D.E., and Klahr, D. (1996). When to trust the data: further investigations of system error in a scientific reasoning task. Mem. Cogn. *24*, 655–668.

46. Aman, M.N., Chua, K.C., and Sikdar, B. (2017). Secure data provenance for the Internet of Things. In Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security, IoTPTS '17, R. Chow and G. Saldamli, eds. (Association for Computing Machinery), pp. 11–14.

47. Lopez, J., Roman, R., Agudo, I., and Fernandez-Gago, C. (2010). Trust management systems for wireless sensor networks: best practices. Comput. Commun. *33*, 1086–1093.

48. Cho, J.-H., Swam, A., and Chen, I.-R. (2011). A survey on trust management for mobile ad hoc networks. Communications Surveys & Tutorials IEEE *13*, 562–583.

49. Klaidman, S. (1987). The Virtuous Journalist (Oxford University Press).

50. Peters, C., and Broersma, M.J. (2013). Rethinking Journalism: Trust and Participation in a Transformed News Landscape (Routledge).

51. Turcotte, J., York, C., Irving, J., Scholl, R.M., and Pingree, R.J. (2015). News recommendations from social media opinion leaders: effects on media trust and information seeking. J. Comput. Mediat. Commun. *20*, 520–535.

52. Behrooz, S., and Marsh, S. (2016). A trust-based framework for information sharing between mobile health care applications. In IFIPTM 2016: Proceedings of IFIP International Conference on Trust Management, S.M. Habib, J. Vassileva, S. Mauw, and M. Mühlhäuser, eds. (Springer AICT), pp. 79–95.

53. Marsh, S., Ghorbani, A.A., and Bhavsar, V.C. (2003). The ACORN multiagent system. Web Intelligence and Agent Systems *1*, 65–86.

54. Harvey, I., Cavoukian, A., Tomko, G., Borrett, D., Kwan, H., and Hatzinakos, D. (2013). SmartData: Privacy Meets Evolutionary Robotics (Springer).

55. Maurer, B. (2015). Principles of descent and alliance for big data. In Data, Now Bigger and Better, T. Boellstorff and B. Maurer, eds. (Prickly Paradigm Press), pp. 67–86.

56. Brockman, J. (2015). What to Think about Machines that Think (Harper Perennial).

57. Broussard, M. (2018). Artificial Unintelligence: How Computers Misunderstand the World (MIT Press).

58. Boden, M. (2018). Artificial Intelligence: A Very Short Introduction (Oxford University Press).

59. Andras, P., Esterle, L., Guckert, M., Han, T.A., Lewis, P.R., Milanovic, K., Payne, T., Perret, C., Pitt, J., Powers, S.T., et al. (2018). Trusting intelligent machines: deepening trust within socio-technical systems. IEEE Technology and Society *37*, 76–83.

60. Kaliouby, R.E. (2019). How do we build trust between humans and ai?. https://www.weforum.org/agenda/2019/08/can-ai-develop-an-empathetic-bond-with-humanity/.

61. Bryson, J. (2018). AI & global governance: no one should trust AI. https://cpr.unu.edu/ai-global-governance-no-one-should-trust-ai.html.

62. Braithwaite, J. (2000). Restorative justice: critical issues. Saskatchewan Law Review *63*, 185.

63. Ness, D.W.V., and Strong, K.H. (2014). Restoring Justice (Routledge).

64. Tylor, E.B. (1871). Primitive Culture, *vol. I* (John Murray).

65. Tylor, E.B. (1873). Primitive Culture, *vol. II* (John Murray).

66. Asma, S. (2020). Ancient animistic beliefs live on in our intimacy with tech. https://aeon.co/ideas/ancient-animistic-beliefs-live-on-in-our-intimacy-with-tech.

67. Gladwell, M. (2019). Talking to Strangers (Little and Brown).

68. Neff, G., and Nafus, D. (2016). Self-Tracking, Essential Knowledge Series (MIT Press).

69. Jordan, T. (2020). The Digital Economy (Polity Press).

70. Kimmerer, R.W. (2013). Braiding Sweetgrass (Milkweed Editions).