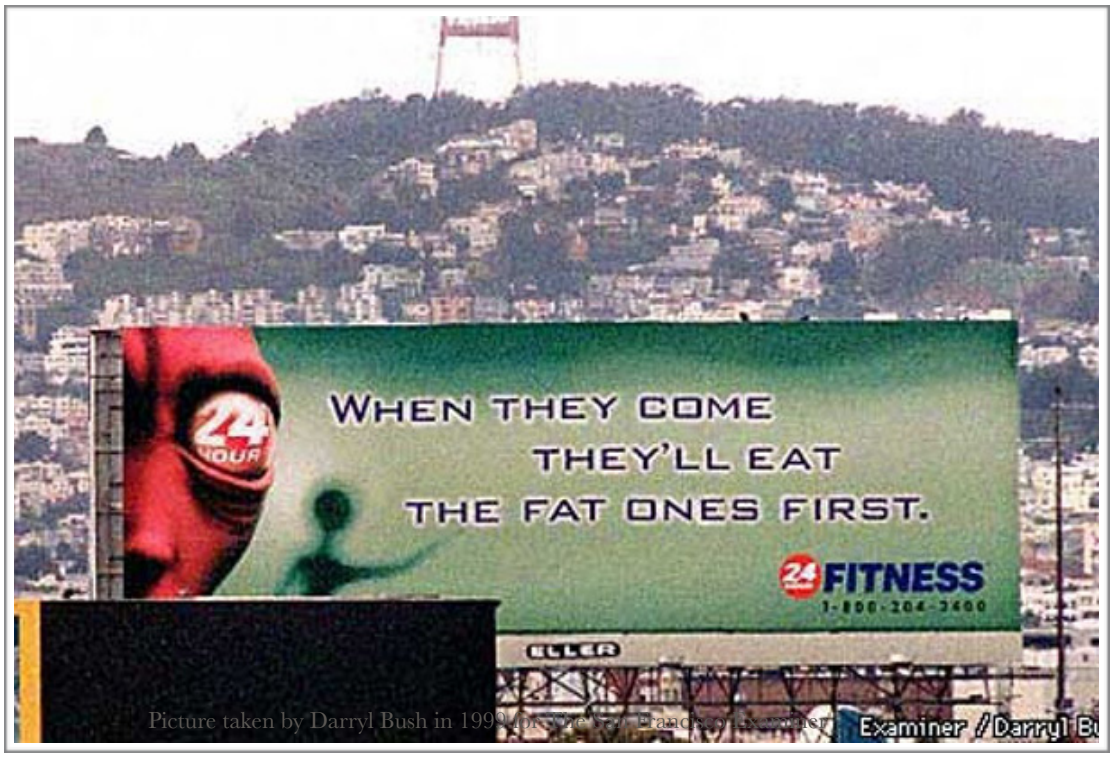
Epistemology in the Cloud

On Fake News and Digital Sovereignty[1]



Henry Story[2],

Web Science, University of Southampton

Final version: 13 August 2018

[1] The picture of the 24-Hour Fitness ad was published in the San Francisco Examiner on 16 February 1999 in an article entitled "Persons of heft protest health club's ad"
https://www.sfgate.com/news/article/Persons-of-heft-protest-health-club-s-ad-3305257.php
The 20th Century Photographic Archives of the SF Examiner were given to the Berkley Bancroft Library in the summer 2006. http://www.lib.berkeley.edu/give/bene-legere/bene70/bene70story5.html

[2] mail: henry.story@co-operating.systems

Abstract[3]: The web is an open platform that allows anyone to publish anything, and so raises anew many epistemological questions: how can one distinguish what is true, what is fake or what is fictional on the web? Indeed how can one know anything at all? We start from an analysis of knowledge that makes space for radical skepticism and which allows us to locate the essential problem with the current web application architecture. This allows us to propose a set of criteria that explicate and justify the decentralised architecture of the internet and the web, and the need for that to be extended to the data and application layer. The proposed architecture is socio-technical, recognising the roles of individuals, institutions and nations in our epistemic makeup. We illustrate this by proposing an architecture of trust that ties these institutions into browsers in a decentralised and open way, allowing them to make the web a more trustworthy space. As a side effect we gain the tools to make some serious inroads in helping combat Phishing.

[3] The content of this paper was presented informally in many talks over the past 7 years. The ideas of an institutional web of trust, was first presented in rough form at the European Association for e-identity and security (EEMA) conference held in March 2011 in Switzerland (Slides here https://bblfish.net/blog/2012/04/30/). The epistemological angle was first presented at the Social Machines (https://sociam.org/) all hands meeting in March 2017. It was then later presented to a small group at the Oxford E-Research Center (OeRC) end of November 2017 whilst also being written up as part of my first year PhD reports in Southampton. Finally it was presented as a non peer reviewed paper at The Web Conf 2018 in Lyon, in the Research Centric Scholarly Communication Workshop (https://bblfish.net/blog/2018/04/21/).

1. Epistemology

"Only The Paranoid Survive"

— Andy Grove, Intel Chairman[4]

"These illustrations suggest four general maxims[…].

The first is: remember that your motives are not always as altruistic as they seem to yourself.

The second is: don't over-estimate your own merits.

The third is: don't expect others to take as much interest in you as you do yourself.

And the fourth is: don't imagine that most people give enough thought to you to have any special desire to persecute you."

— Bertrand Russell, The Conquest of Happiness

The web is an open platform that allows anyone to publish anything, and so must raise issues of knowledge of what is true, false, fake or fictional that are as old as philosophy.

To be able to reason about different points of views, fictional truths and statements of knowledge in a precise mathematical way, only became possible relatively recently, in the 1960ies and 1970ies, with breakthroughs in modal logic which clearly formalised the work started by Gottfried Wilhelm Leibniz three centuries earlier. The transformations brought about by Arthur Prior's work in the 1960ies on temporal logic, Kripke's on necessity, and especially David Lewis' 1973 work on Counterfactuals[5], allowed concepts to be clarified in many different areas of philosophy that previously were immersed in mystery. For example, the concept of causation which had been discussed since Hume took on a new precision with David Lewis' counterfactual analysis of it[6], and has since then grown in importance as attested by a recent book "Counterfactuals and Causal Inference - Methods and Principles for Social Research"[7].

---

4 Grove, A. S. (1996). *Only the paranoid survive: How to exploit the crisis points that challenge every company and career*. Broadway Business.

5 Lewis, David. *Counterfactuals*. John Wiley & Sons, 2013.

6 David Lewis starts his article on causation with the sentence " Hume defined causation twice over. He wrote "we may define a cause to be an object followed by another, and where all objects , similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been the second never had existed" ". David Lewis points out that these are not equivalent definitions, but that the unclarity of what could possibly be meant by the counterfactual statement hid that option from most philosophers until his day.
Lewis, David. "Causation." The journal of philosophy 70.17 (1974): 556-567.

7 Morgan, Stephen L., and Christopher Winship. *Counterfactuals and causal inference - Methods and Principles for Social Research*. Cambridge University Press, 2014.

This century these advances in modal logic have been given even stronger mathematical underpinning by a clean integration into Category Theory[8] where they have been shown to be as essential to reasoning with coalgebras[9] as equational thinking is to reasoning with algebras. Coalgebras are the Category Theoretic duals of algebras having been found to describe the structure of infinite streams, processes, object oriented programming[10], and thought even to form a general theory of Systems.[11]

Modal logics has found its way into the technical standards that form the world wide web, such in the RDF Semantics spec which clarifies the concept of interpretations in terms of possible worlds[12]. And as we will show in this paper, modal logic will provide an essential tool for thinking about security and knowledge.

To show how modal logic has impacted epistemology let us consider Robert Nozicks' ground breaking initial analysis of knowledge in "Philosophical Explanations"[13].

Armed with the new logical tools developed by David Lewis in his 1973 book "Counterfactuals" Nozick gave the following initial definition of knowledge conceived as a relation between a subject S and a proposition P[14]:

S knows P →

1. P
2. S believes P
3. ¬P □→ ¬ S believes P

Where φ □→ ψ is to be read as "If φ were the case then ψ would be the case", which is interpreted to mean: in the closest possible worlds to the actual world in which the proposition φ is the case, the proposition ψ is the case (see the illustration below from David Lewis' book). This requires one to have a distance relation on possible worlds where worlds are closer if they require less change from actuality to occur, and to think of propositions in terms of sets

---

[8] Cîrstea, Corina, et al. "Modal logics are coalgebraic." *The Computer Journal* 54.1 (2009): 31-41.

[9] Kurz, Alexander. "Specifying coalgebras with modal logic." *Theoretical Computer Science* 260.1-2 (2001): 119-138.

[10] For many links to this see my answer to the Quora Question "Why is functional programming seen as the opposite of OOP rather than an addition to it?" : https://www.quora.com/Why-is-functional-programming-seen-as-the-opposite-of-OOP-rather-than-an-addition-to-it/answer/Henry-Story

[11] Rutten, Jan JMM. "Universal coalgebra: a theory of systems." *Theoretical computer science* 249.1 (2000): 3-80.

[12] https://www.w3.org/TR/rdf-mt/#interp

[13] "*Philosophical Explanations*" 1981, Robert Nozick, Harvard University Press

[14] Nozick keeps his definition in terms of an "if and only if" analysis, which he refines in stages to deal with counter examples, as per well known philosophical tradition. He ends up with a definition that includes clauses relating knowledge to methods of knowing. The part of the definition given here subsists across those changes.

of possible worlds, or ways the world could be. I.e. all the ways the world could be to make the sentence that describes them true.

The definition above is then to be read as follows. If S knows P then

1. P is the case in the actual world (ie. P is true)
2. the Subject S has a cognitive relation to the proposition P as one of belief (and so given the right other beliefs and desires would act on that belief)
3. if P were not the case then S would not believe that P.

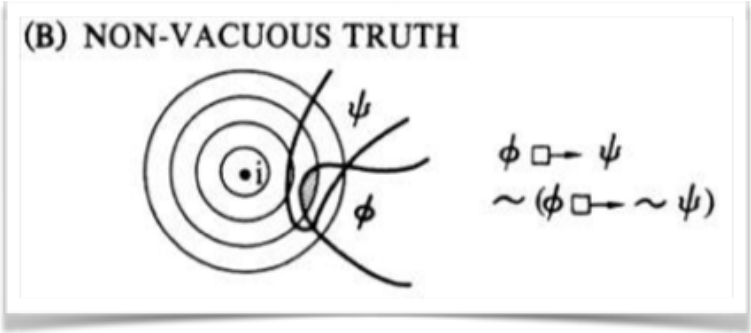


Illustration taken from David Lewis' "Counterfactuals". At the centre of the system of spheres lies the actual world **i.** Proposition ϕ is not true at the actual world, but would be true at the closest ϕ world. Note that not all ϕ worlds are ϕ worlds. Indeed if one could not distinguish the outer spheres then the counterfactual would no longer be true.

This is what is known as the truth tracking definition of knowledge[15] — S has to track the truth across counterfactuals possibilities.

The thought process that lead to this first analysis starts by considering the limits of knowledge with a thought experiment that updates the famous one in Philosophical Meditations where René Descartes summoned the Evil Genius, a spirit that could lead anyone to doubt everything except as it turns out, the thinkers own existence. Nozick is less concerned in establishing the phenomenological certainty of his own existence, than he is in asking how we can know anything empirical at all given the following possibility: it is clearly logically possible that an alien civilisation from say Alpha Centauri could have discovered our planet Earth and for some reason be interested in capturing humans at night, taking them off to their spaceship and then simulating their brain with the advanced computers they have in order to synthesise a completely realistic and coherent perception of the world indistinguishable to the perceiver to the real one. Similar thought experiments have found themselves into movies, most well-known of which is The Matrix[16]. In any case the argument

---

15 Kelly Becker and Tim Black (eds.) The Sensitivity Principle in Epistemology, Cambridge University Press, 2012, ISBN 9781107004238.
See review: http://ndpr.nd.edu/news/40514-the-sensitivity-principle-in-epistemology/

16 The Matrix, by the Vachowski Brothers, 1999

goes, if we cannot phenomenologically distinguish the situation we are in from the one we believe ourselves to be in, as is the case with the captive being experimented on by the Alpha Centaurians, how can we know anything at all? The captured brain in the vat believes many things now as being fact that are simply not true. He thinks he is going to work, but all his nerves have been unplugged from his body and it is really his avatar who is going to a simulation of a work environment. Some may argue that the meaning of his terms have then just changed to refer to objects in the simulated world, and so that they are still true, of the simulated world that is. Still it is easy to see that this won't do: if at some point the Alpha Centaurians decide to let him go, plug his brain back into his body, and send him back to earth, he will be confronted with incomprehension whenever he tries to refer to conversation that he thought he had with people but in fact only had with simulations.

So why is that a problem for knowledge? Well, one criteria one may suggest for *S knowing that P* is that S should be able to exclude situations in which P is not true. And yet there is no way anyone can subjectively exclude the Alpha Centaurian case: whatever you think is the case is also compatible with you being deceived by such a powerful alien force.

This is where the counterfactual statement 3 above kicks in: it is meant to save us from this radical doubt by allowing us to put a modal distance between us and the skeptical possibility, without needing to deny its coherence. It works as follows:

Imagine a situation in this world where I believe that I have 50 pounds in my pocket, because I went to the cash machine an hour ago. Is that knowledge?

Consider the world where it is true that I have that money in my pocket and I believe I do. The closest world in which I don't have the 50 pounds in my pocket, are pretty mundane worlds such as ones where I would have spent some of it in a shop, for a bus ticket, or given it to someone,... Ie. we can imagine many situations where had I not had 50 pounds in my pocket something would have happened that is a lot less outrageous than the Alpha Centaurians scenario. And in those circumstances, assuming I am somewhat conscious of what I am doing with my money, I would then not believe that I have 50 pounds in my pocket. Thus I satisfy the conditions for knowledge.

Now, if I live in a part of the world full of active pickpockets, and I don't take any precautions against them, then one could argue that I don't actually know that I have 50 pounds in my pocket, even if I do happen to have them and believe I do, because with just small changes to the actual world I may not have had. From the point of view of people who want to work with me, this makes me unreliable[17], and for them my claim that I have some money will always require some extra verification.

Similarly if I am a sloppy person and have not fixed a holes in my trouser pocket then it is arguable that had circumstances been just a little different, the money might easily have slipped out of my trousers. And so we should also say that even though I believe that I have the money and even if I do have it, that this belief still does not constitute knowledge. The

[17] that seems to be an opening to the reliabilist epistemology

way the world is, the way I act, the political and the ethical space I am in determines in many ways whether I know or not. "Knowledge is not just in the head", one may say to echo Hillary Putnam[18]. It is a relation between the mind, others, tools, the way things are and because propositions are in this view sets of possibilities, a relation to counterfactual states too — how things could have been.

We can start seeing many interesting applications of this idea: from a coding perspective this helps explain why automated tests are important, be they in the form of unit tests that run automatically after each major commit to test that the functions give the right result on hard coded or automatically generated particular cases, or a compiler that checks that the types of functions line up[19]. Keeping the tests up to date, well documented, and having a policy that is enforced, makes it easier for the team to know where they are in the coding process[20], and if code refactorings take place — these are code rewritings that don't change the logic of the program, but change its elegance, readability, etc… — to know the invariants have been kept. So this is the equivalent of not having holes in one's pocket, the state not allowing pickpockets to thrive, or for that matter for operating systems to be hardened so as not to allow infiltration, changes to the operating systems, and of course to teach people not to give access rights to just anyone.

What we can take back from this is that knowledge and security in one's knowledge is therefore something that brings the larger whole into the picture, including not just the world as it is, but how it could have been, or might have been. A good imagination therefore is an important aspect of knowledge: to know is to learn to imagine what could go wrong, and setting up processes to alert one when it does.

But at the same time a too powerful imagination can end up stifling all action if the distance from the actual world of imagined possibilities is not taken into account, as an encounter with a hard core skeptic will reveal. Hard core skepticism can indeed lead to a rejection of any type of process for knowledge, as the skeptic will have deemed that all knowledge is impossible in advance. Similarly it is easy in security matters to go from the thought that since any security measure can be bypassed (eg. all current cryptography can be broken by advanced enough quantum computers) there is no point in security at all. Or to attack a reasonable security measure, because one can imagine a way to break it, without taking into account that the measure makes things more secure than they were before, given a certain reasonable assumption (nobody yet has such quantum computers).

---

[18] "Meaning just ain't in the head"
Putnam, H., 1975. "The Meaning of Meaning." In Mind, Language and Reality; Philosophical Papers Volume 2. Cambridge: Cambridge University Press, 215-271.

[19] As made possible when using a language such as Scala, or by adding proofs of algorithms in code for languages that allow this such as Idris or Scala enhanced with Leon.

[20] I wrote up an initial thought of this in a blog post from Sept 2006 https://web.archive.org/web/20110601232015/https://blogs.oracle.com/bblfish/entry/the_fifth_dimension

So a modal analysis of knowledge helps us answer the radical skeptic. It can help us understand how we can know things without being able to discount the radical skeptical possibility. We can accept the limits of knowledge without abandoning reason itself.

It also allows us to answer the dogmatist, who as the mirror image of the skeptic, believes his dogma is secure and unquestionable, and does not need confrontation with larger spaces of possibility, a sure step to failure. For the modal analysis of knowledge does not deny that the outlying situation is possible.

2. The Cloud

Indeed if we were living in a world where Alpha Centaurians with the type of technology Nozick imagines were to be close to our planet, then we would be in what has to be thought of as an epistemological war.

We would either have to find a way to fight the Aliens, counteract their technology, or live with them in submission. An epistemological war against such an advanced civilisation would be nearly impossible to fight. One would not know if one is giving orders to one's army or just to a simulated one. Vice versa a soldier would not know if it was receiving an order from a superior or just receiving a fake order[21]. One would no longer know if one was installing a new operating system patch or an enhanced virus.

That would seem to leave us as only option to live with the Aliens. To make this possibility vivid, take the advertisement for 24 Hour Fitness that ran in San Francisco in 1999, showing a billboard of Aliens invading in Flying Saucers, with the caption "When they come they will eat the fat ones first"[22]. This advertisement was in a humorous way (humorous

Picture taken by Nathalie Harvey on 1 April 2016

[21] The experiment "Face2Face: Real-time Face Capture and Reenactment of RGB Videos" actually shows that even live conferencing with a famous person can already now be faked very realistically. https://web.stanford.edu/~zollhoef/papers/CVPR2016_Face2Face/paper.pdf See the video https://www.youtube.com/watch?v=ohmajJTcpNk

[22] The advertisement was rerun in 2016 and the picture below taken and posted by Nathalie Harvey on Twitter on April 1, 2016, https://twitter.com/natharvey77/status/715902927177695232. The story was covered by Stacey Ritzen in an Uproxx article "People Were Offended By This Gym Billboard Threatening Alien Abduction: 'They'll Take The Fat Ones First'" https://uproxx.com/viral/gym-billboard-alien-abduction/

because the presupposition is evidently so far-fetched, and there should be many better reasons for us to try to be fit if we can) asking us to project ourselves in that possible space[23] where in order to avoid being eaten we need to adapt our behaviour by going to the gym. Of course one can see the financial interest the gym has in spreading that thought, and those with a conspiracy theorist bent of mind would use this as an opening for a reason to avoid fitness at all costs — perhaps aliens really like fit people! The point here is that the truth of a statement, and especially a counterfactual one, depends on where one thinks the actual world is in the space of possibilities. One can thus learn to understand people by determining where they think they are in that space, and one can also try to influence people by increasing their awareness of certain possibilities, for profit, as with the above mentioned ad, or for the greater good, as is the role of educational institutions.

All of this is good fun someone may say, but nobody sane seriously believes the aliens are about to attack from the clouds. Perhaps, ... but should we therefore not be worried? Well, I think we can agree that there are more realistic things to be worried about than Alpha Centaurians, but that is precisely why Nozick's analysis of knowledge is so attractive.

Still, there is a metaphorical truth to the alien scenario that remains which can and should give one very much reason to be worried. Computing now, as writing before, is, if we follow thinkers such as Bernard Stiegler, another stage in the exosomatic evolution of thought that started with tools such as the spear, and then accelerated at exponential rate with the advent of alphabetical writing. The laws of the Polis (πόλις) were carved in stone on the walls of the Athens of Ancient Greece 2500 years ago, for everyone to read. As such these writings were difficult to change without attracting attention: one would have needed to be in front of the wall with a hammer and chisel and hammer away a fake new law in full hearing of everyone! Things are completely different now: we have now externalised our thinking in cloud computers that are looked after by agencies that often have very different aims to ones that make use of those services, and these could easily change the information without it being noticed, since information is now just constituted of differences at minuscule near atomic layers of matter. If we think of these computers and data centres as a part of our extended mind[24] then we are not that far from what now looks more like a parable of the Alpha Centaurians. Are our externalised minds not already in the hands of foreign entities, many of which are run by aliens[25] with offshore homes and money stored in tax havens!?

Many may be aware of this situation but reason that the value of their information does not make their information "fat" enough, that is: valuable enough to be abused. They may also reasonably think that they would notice information that had been changed in their writings, from the traces left in their own wetware memory located in their skull. And so indeed such changes would not be a good attack vector to use often.

---

[23] Lewis, D. (1978). Truth in fiction. American Philosophical Quarterly, 15(1), 37-46.

[24] Clark, Andy, and David Chalmers. "The extended mind." *analysis* 58.1 (1998): 7-19.

[25] In the sense of Sting's song "I'm an Englishman in New York" whose refrain is "I'm an alien, I am a legal alien..."

But for ex-organisms that are constituted in very large part of externalised memory, such as larger companies or dispersed groupings, such attacks would be a lot more difficult to trace. It has been argued that the size of current civilisations was made possible through transformations in mnemo techniques and technologies[26], from the alphabetical writing in stone, to parchment, to the printing press, to computer based information technology. Each of these made more complex group interactions possible[27] and so could lead to larger societies. These externalised writings are key to the existence of these organisations.

Controlling and securing the externalised memory in such ex-organisms, is the equivalent to making sure that no entity was interfering with one's brain. As such it becomes important to be able to set up control measures for changes of information and to control at least part of the hardware that supports the memory, as well as be able to inspect code that makes such changes to make sure it does not hide trojans or enable them.

This thought experiment then has lead some to think of ways of reducing the memory support needed to verify changes, by for example hashing content, and keeping the hash of the content in some secure place. That is a good way to notice changes to externalised memory if used in a disciplined way[28] (but it does require at least some part of the memory to be trusted — namely the support where the hashes are written). If the content can be made completely public and is popular then the wide distribution of the content reduces the likelihood that all copies will be lost. This can also work for keeping track of collections of encrypted content.

But the attack of changing externalised information on which an organism relies is not the only means of influencing it surreptitiously. If instead of thinking of information here we think of the propositions — which in modal logic we identify with sets of possible worlds — encoded in the information, then the filtration mechanism that will choose amongst the stream of propositions to present some and hide others, will in just as reliable a way direct a person to thinking of possibilities as being closer or more desireable than others, and so affect their behaviour. This is indeed how advertising works. The difference is that with advertising on public one-to-many channels such as Newspapers, Radio or Television, the advertising is visible to everyone and so could be discussed, criticised and regulated for the political good of

---

[26] This has been made clear through Thomas Thwaites art project, TED Talk and Book "How I build a toaster from Scratch" https://www.ted.com/talks/thomas_thwaites_how_i_built_a_toaster_from_scratch
and also by Matt Ridely in his TED Talk "When Ideas Have Sex" where he starts by showing evidence for how reductions in the sizes of civilisations have in the past lead to loss of technological know how.

27 Those around in the 1980ies may remember how terrible telephonic client service was, as one got redirected from agent to agent, each one having no memory of the previous interaction, and requesting all the same information again.

[28] Linus Torvalds, the originator and key maintainer of the Linux Operating System, and author of the versioning system Git, that uses such hashes to build a secure distributed version control system. Linus keeps the hashes of changes he makes in a logbook at home as he explained in his presentation of Git at the Google Tech talk in May 2007 and still available on YouTube at the moment https://www.youtube.com/watch?v=4XpnKHJAok8

the Polis. In a completely personalised medium where the criterion of selection of the filtration machine is furthermore sold to the highest bidder, it is no longer possible for the body politic to easily tell how filtration is happening and how it is affecting individuals, until perhaps it is too late. This influencing of individual's beliefs may affect minor issues such as what soap people buy, or it could influence an election, as some articles have claimed may have happened in the UK Brexit vote and in the US elections through targeted advertising using psychometric technology developed by Cambridge Analytica[29] developed with information harvested from social networks. Ironically these election whose theme was Sovereignty, may in fact have demonstrated how the emergence of the large social networks had in fact lead to a massive loss of digital sovereignty[30]. What good is it to control the walls of your city if the minds of your citizens who elect those in power can be taken over?

Loss of digital sovereignty could make one think that this loss may have been to the benefit of another nation. But things have become more complex with the introduction of robot generated content[31], which is responsible for a huge portion of content on the internet. As these usually try to optimise clicks leading to add sales, clicks that may have been generated by other robots, by humans being paid to do so, intentionally broken links, and just simply steered by the base unchecked lust lurking in dark corners under the cover of (pseudo) anonymity. James Bridel shows in a blog post from Nov 6, 2017 "Something is wrong on the internet"[32], how this bot-human process is creating very unpleasant content. He considers in particular all the content aimed at children under the age of five, showing how a lot of it makes no sense at all, and worrisome amounts of it is laced with violence.

Finally, there is a long tradition that rationality requires openness and transparency, a tradition going back very far, as Peter Szendy reminds us in "Kant in the Land of Extraterrestrials: Cosmopolitical Philosofictions"[33], where he points out that the famous German philosopher had imagined that a universal rationality would require one to think of

---

[29] "Data Firm Says 'Secret Sauce' Aided Trump; Many Scoff" New York Times, By Nicholas Confessore and Danny Hakim, March 6, 2017
https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html
"Did Cambridge Analytica influence the Brexit vote and the US election?", Saturday 4 March 2017, the Observer. https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump

[30] The case for the importance of Digital Sovereignty is made forcefully by Pierre Bellanger who shows that it will affect every dimension of industry.
Bellanger, Pierre. La souveraineté numérique. Stock, 2014.

[31] See the July 2016 issue of the Communication of the ACM entitled "The Rise of Social Bots" https://cacm.acm.org/magazines/2016/7

[32] https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2

[33] Szendy, Peter. Kant in the Land of Extraterrestrials: Cosmopolitical Philosofictions. Oxford University Press, 2013.

the possibility of alien thinkers and who had imagined that these would due to their having reached such an ultimate stage of civilisation never be able to say a single thought in secret, but have to pronounce each one out loud, making each thought evaluable to their peers.

But that does not somehow feel quite right, certainly not in antagonistic situations, which one could argue will always exist, and may even be essential to thinking itself[34]. We need only remind the reader of the advantage that the UK won in being able to decipher the german encryption produced by the Enigma machine during the second world war — a project on which Alan Turing worked. But this is also evident in the huge amount of money that is being poured by traders in getting the fastest access to trading data. If any trader could know before the others what bids they are making, it would then be easy to make a better bid, and so an easy profit, and in the long term win without difficulty. On the other hand games such as chess or Go which seem to be completely transparent show that facts are not the only thing needed to gain an advantage. The facts — in this case the position of the stones on the board — may be visible to all, but only be correctly interpreted by a few. But there one could argue that if one had full access to the thought of the Go master, then it would also be an unfair game (though that would very likely create a meta-interpretation problem of even bigger complexity). In any case Go and chess show that even if everything is public, not everyone is in the same position to understand what is going on. Assuming that those looking at the data won't necessarily have the concepts to interpret it correctly, might procure some temporary relief until one realises that if those reading information are not aware of their interpretative bias, if they are not aware that they may be misreading the information, and if those people are in positions of power in a security apparatus that has aims to be all encompassing, one may soon find oneself in a Kafkaesque world of misunderstood and arbitrarily applied rules.

[34] The importance of controversies to the evolution of academic thinking has been stressed by Bernard Stielger in a number of talks in the past 10 years.

3. Co-Operating Systems

So this train of reasoning leads us to a few important conclusions.

First, it is essential to take into account the physicality of information when reasoning about it and deciding to make use of it. Where was it produced? Where is it stored? What processes of verification did it undergo? are key questions that need to be answered in order to be able to assess the content as knowledge, and so to be able to make use of it. One needs to know who gave one a piece of information to be able to evaluate it epistemically: does that agent have the required organisation to be reliable in talking about a specific topic? The meteorological office uses satellites and networks of experts[35] to come to conclusions about the weather. An organisation that simply copies the information from the expert site and republishes it, depends for their trustability on that copying being understood to be faithful to the institution with the right organisation to make such statements. Lack of domain knowledge can only be countered by faithful copy and correct attribution. Lack of such an apparatus or attribution would make those claims indistinguisheable from someone who was just consistently lucky at guessing the weather — luck that may run out the day the originating office pays back the lack of honesty by feeding fake information to the pirate.

Second, these assessments of reliability are modal notions. They require one to consider not just what happened but what could have happened: how would relevant counterfactual situations have affected our answers? This requires thinking about the processes we are using and their embedding in a larger whole: have we secured ourselves against problematic situations? Or are we relying on someone else to do that for us? Different actors are tasked with specialising at understanding different possibilities: we walk home at night in confidence that we won't be invaded, because we trust on our military to be considering and preparing for the worst possibilities so that we don't have to think of those[36]. But we have to trust that military to work for us. The Swiss way of doing that is to require every man to be in the militia every year for over a week and to make referenda very easy so as to make sure the few don't send the many to war. Such a concept could be translated into the cloud by requiring every citizen to have their piece of cloud at home, where both device and data can be controlled and maintained.

---

[35] the Heidegerian Gestell according to B. Stiegler

[36] But because counterfactuals take all of reality in consideration, the danger is that over-specialisations of different security organs that take some possibilities into account but not all of them can lead to a misreading of the terrain. This is the point made by Pierre Bellanger on 21 December 2017 at the Ministry of the Army in France under the title "Comment Gagner une guerre Perdue" ("How to win a lost war") and republished on the web here https://www.lettrevigie.com/blog/2018/01/09/comment-gagner-une-guerre-perdue%E2%80%89-p-bellanger/

Thirdly, this larger context requires us to think of ourselves as working both in cooperative and antagonistic situations: we need to work together with other agents and larger processes but we also need to be aware that there are intelligent forces that may undermine us for their perceived good.

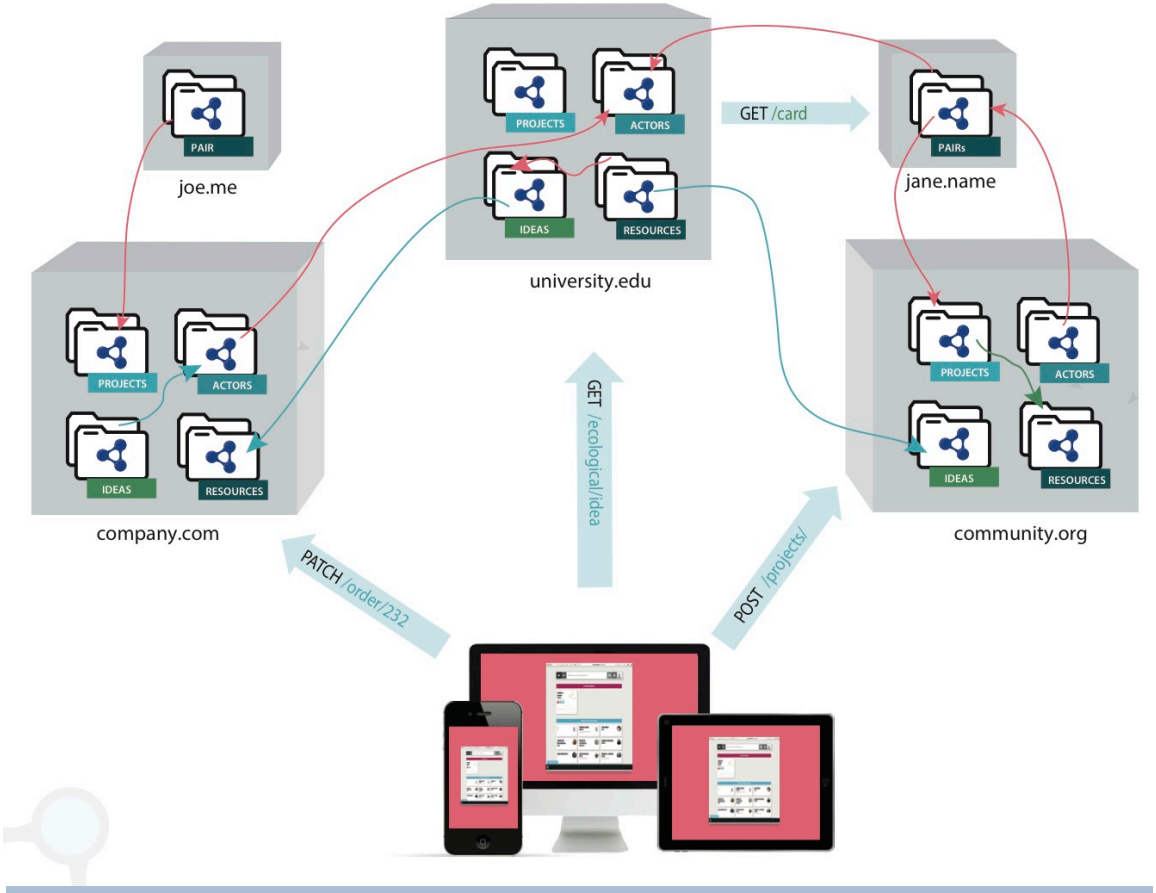
These three points do raise the question of who the "we" is that is being spoken of. As Bernard Stiegler, basing himself on Gilbert Simondon's work on the process of individuation reminds us, we are formed through our language, the skills we are taught (such as reading, which Maryanne Wolf has shown to transform our brain into a reading brain[37]), and the many people around us we interact with who are members of institutions such as schools, police, hospitals, universities, armies and other organisations. We in turn shape these by contributing in original ways to develop new ideas, objects, technologies, works of art, ... These consistent wholes in turn individuate others, such as when the various towns of Italy competed to produce ever more beautiful towns, palaces, and churches, or when the European Nations compete to produce the best universities, companies, or educational systems.

There is thus a transductive dependence of the whole on the individual and vice versa, and of larger wholes on larger wholes that suggests that the correct organisation of the information infrastructure needs — one that enables knowledge — has to be decentralised with each node acting as peers, under the ownership of individuals or organisations, yet able to co-operate with any other peer in the system. It has to be distributed, for otherwise there would be one over-arching network owner, and a loss of individual autonomy and collective sovereignty — understood as the ability to make laws that bind a group of individuals. Parts of the data must be open to all and other parts secured and accessible to only some, such as when companies need secrecy to develop a product, lovers want to communicate, or teachers want to help students in a space where these don't need to live under the weight of real or imagined potential social criticism.

We thus need a read-write web with global distributed identity, on which access control can be built, with systems of trust based on social networks at the level of individuals as well as at the level of organisations, with the ability for each agent to archive what has been read or seen for later use in potential litigational situations, or just to remember how he came to a conclusion. At the limit the architecture of the system has to be designed to allow each individual access to his publication platform where he can make sure that no hidden agency is located — ie, no hidden back door that could undermine his agency. This type of device is what Eben Moglen has coined the Freedom Box[38], a material device that someone can own, on which they can place their information and publish it to the world, and yet that is protected from search and seizure by the fourth amendment of the US constitution or

[37] Wolf, Maryanne, and Catherine J. Stoodley. Proust and the squid: The story and science of the reading brain. Findaway World LLC, 2008.

[38] https://freedomboxfoundation.org/



Example of a hyper-app.

A hyper-Address Book (hAB) functions like a hyper-text browser but with the additional ability to write to the owner's hyper data space. Such an hAB can try to follow hyper-links to data on other SoLiD web servers where it will mostly not have write access and sometimes not even have read access. The hAB helps the user navigate the hyper-data but has no ties to any of these domains in particular.

equivalents elsewhere. The concept of the Freedom Box need not be restricted to individual humans, but can of course also be applied just as well to larger institutions such as schools, universities, churches, hospitals, law cabinets, police forces, the army and even a nation, each of which should be able to host their own servers and control their own data whilst being able to co-operate with other such individuals and the general public.

The Freedom Box concept deals with the materiality of the publication device, and the inspectability of its operating system. But to allow for fluid co-operation between individuals and organisations without requiring centralisation of information either on one server or in a distributed unique database (eg a blockchain), one needs not just an operating system, but a co-Operating System[39], one that allows individuals and groups to work together as peers, using applications that can follow links between servers owned by different groups of people without needing anything to be centralised or everything to be visible to all. The World Wide

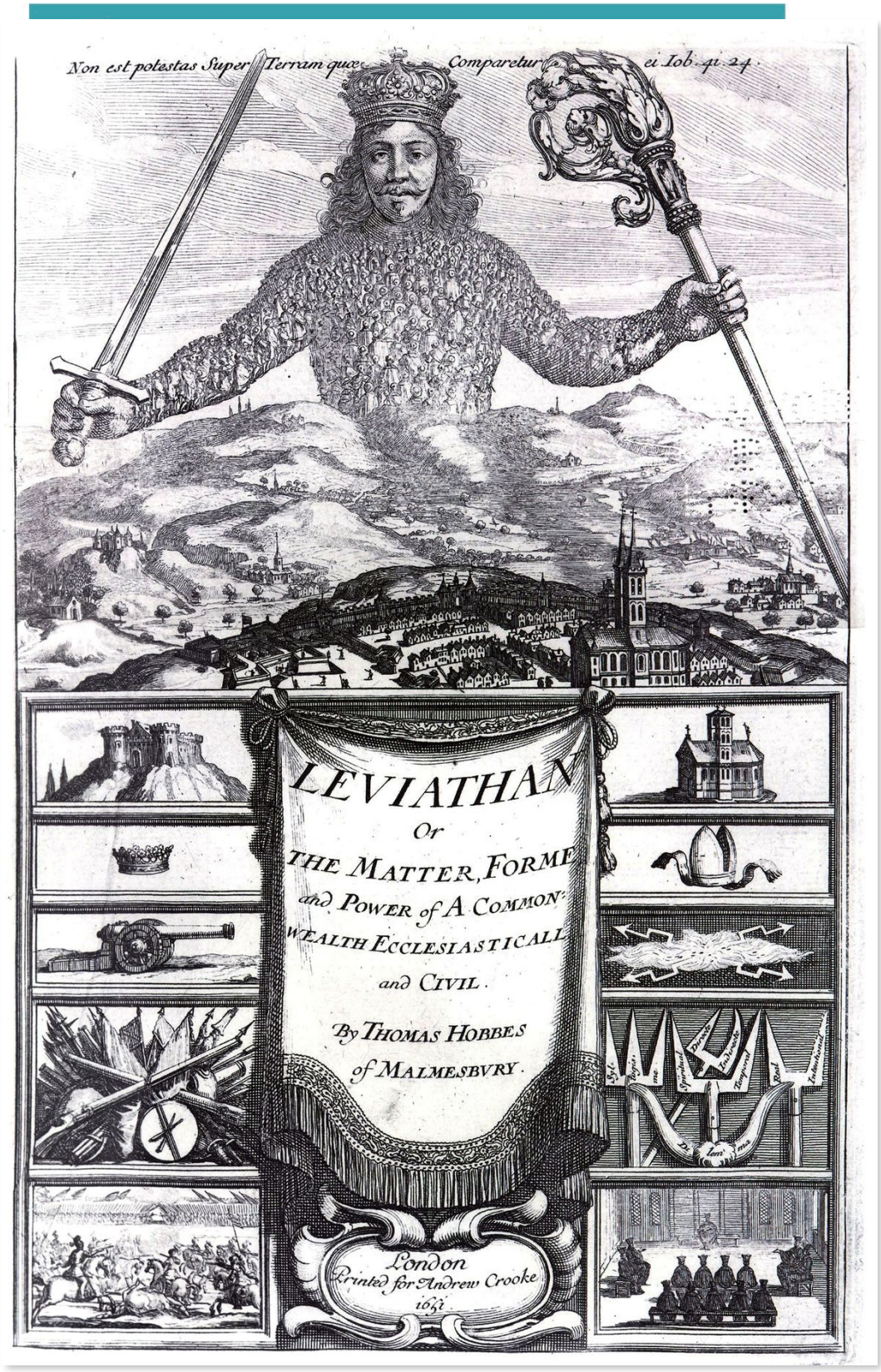
[39] And indeed this paper is part of the work on this topic whose web page is http://co-operating.systems

Web has shown how one can have applications — the Web Navigator or Browser being the first such — that can follow links between content produced and published on different servers located across the world, without the application being tied to one domain, as the current social network apps currently are. This concept needs to be extended to all applications.

By adding a machine readable document layer to the web, one can then create new hyper applications that can present hyper-data (data that can link web resources on the same server or across domains and standardised at the W3C under the banner Semantic Web). Such hyper applications need to have the ability to authenticate their users across domains seamlessly and securely, since information will in many important cases be spread across server nodes. In order to avoid bothering the user about each login, it should be easy to specify privacy policies that allow the software to make decisions as to when automatic authentication is acceptable, which identifier or verifiable claim to use.

We thus can describe a socio-technical stack that such a co-operating system must be composed of. At the bottom layer we have the inspectable hardware, and the operating system that controls it, such as the freedom box, where it is relatively easy for the owner of the box to verify or have verified that no Trojan Horses are located. Above it we need a Web Server that can authenticate incoming requests and make access control decisions based on rules or policy as to the ability to read or write depending on information linked to from each resource, so that the server and the client can make decisions based on the same information. The work the MIT Distributed Information Group has been doing under the name of SoLiD[40] (for **S**ecure s**o**cial **Li**nked **D**ata server) basing itself on various W3C standards such as the Linked Data Protocol for the Read/Write component forms such an enhanced web server. The interlinking of resources on these individual servers forms an enhanced World Wide Web that has all the same properties as the current one plus some new ones that make distributed hyper applications possible, ones where the data is not centralised in one place but linked to across the nodes of the network, as far as possible on hardware of the agency responsible for that information. These hyper-apps give people the view needed to interact with this data web. They are the modern equivalent of pens and paper to write hyper data — that is data structured in such a way that it can be distributed and located in a global space of agents working in institutions that need to co-operate safely. Initial prototypes of hyper-address books, hyper-calendars, and many more have been developed by members of the SoLiD project. These hyper-apps will then be used by individuals often in institutional roles to cooperate with other within and across institutions. This should make it possible for institutions to work in a fluid way with others in a just in time way on subsets of data relevant to the work they need to do. This would also allow hyper-agile small companies to work together by making it easy to share skills and knowledge and so reduce duplication of work, all while being respectful of each groups intellectual contribution.

[40] https://github.com/solid/



The Frontispiece of Hobbes' famous Leviathan is an early depiction of how the state is composed of its people reposing on military and clerical institutions.

4. Digital Sovereignty

By enabling decentralised ownership of hyper content, the co-Operating Systems stack allows each actor to the best of his ability to gain the maximum control of his information space, and his relations to others. This makes it much more difficult to have a brain in the vat attack where an alien can inspect or change the data on which the agent relies to think. But it cannot stop people from believing fake stories that seem believable published elsewhere in the decentralised network.

A decentralised social network based on friend of a friend type ontologies such as foaf[41] allows people to connect to people they know, and through tracking of what people say of themselves and what they know about their friends in daily interaction with them, to evaluate the believability of the information they publish and to certify their identity without needing much if any institutional support. But as one extends the degree of separation between oneself and the friends of the friends ... of ones friends, one will very quickly have included all agents in the world in that group, including robots, thieves, enemies, and vendors of fake news[42].

To resolve this problem our civilisations have over time developed complex webs of specialised institutions each of which has processes in place to help evaluate different types of information. These institutions need to be made visible by integrating them into all hyper-apps including the original one — namely the web browser or Navigator — to allow users of these applications to identify which institution is behind which web site they are looking at.

That this is needed is shown by the recent crisis of fake-news organisations that create scandalising stories to attract clicks and earn advertising dollars[43]. It show that very many people using their web browser are not currently able to work out what kind of organisation they are getting a story from — which is not surprising since many of these do their best to appear like organisations their users may want to trust. The ability to work out what kind of web site one is on, is furthermore not an easy skill to acquire, and currently if done seriously would be extremely time consuming even to advanced users of the web, who cannot do much more than deploy some well informed guesswork.

---

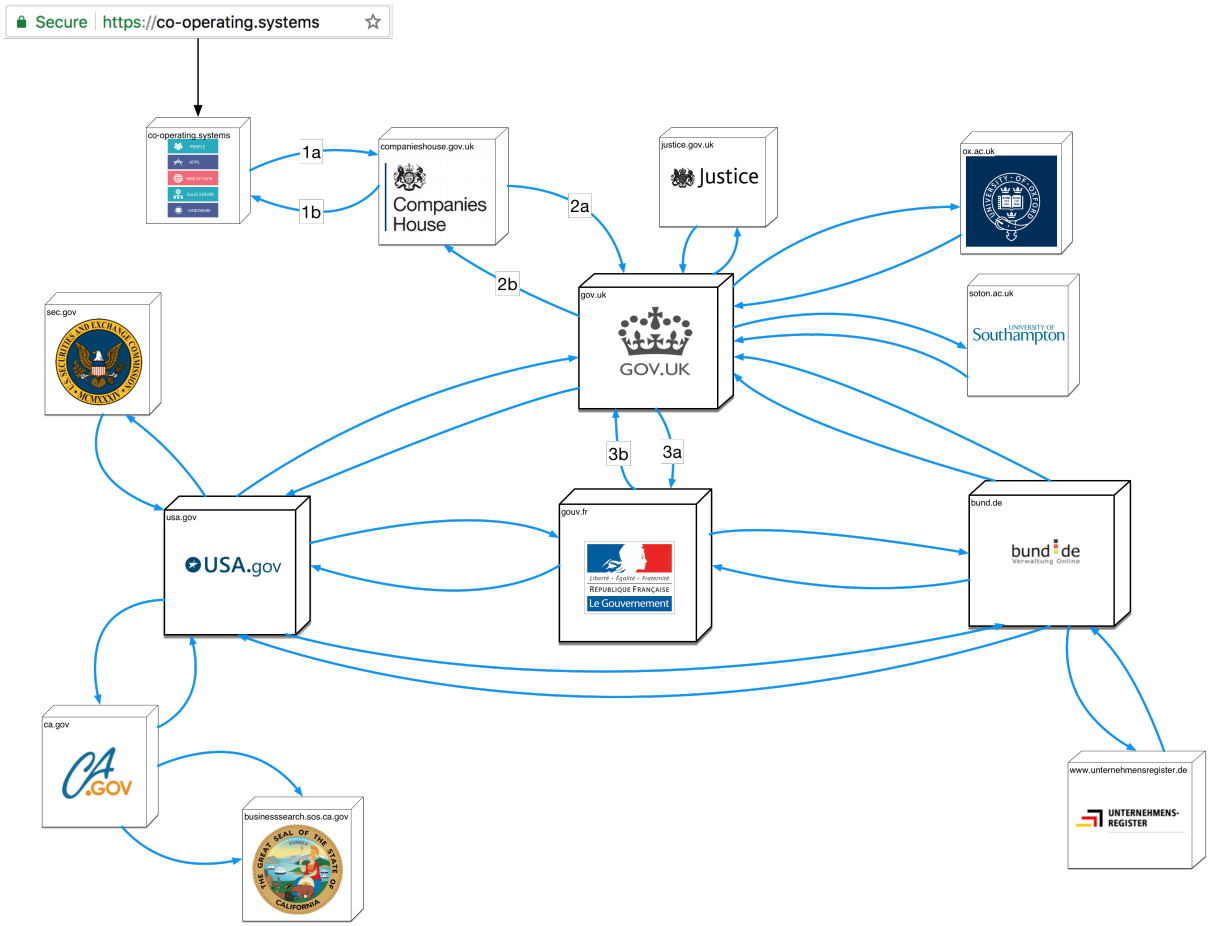
[41] http://www.foaf-project.org/

[42] This is the famous "six degrees of separation" thesis, that argued every person was separated to any other person in the world by only six degrees of separation. https://en.wikipedia.org/wiki/Six_degrees_of_separation

[43] Wired "The Macedonian Teens Who Mastered Fake News" February 2017
https://www.wired.com/2017/02/veles-macedonia-fake-news/

What would be needed for the guesswork to be removed in most situations? Well, it would require the browser to be able to use the information states have about their organs — Which are the universities? Where are they located? Which are the companies? the hospitals? the schools? and so on — and display this succinctly, elegantly and in an appealing way to the user in the browser. On first arriving on a web page the user could for example be shown a rich and up to date set of information about the company owners, the type of company it is, the legal space in which it works, etc... all collected from one of the official registries of such information as declared by a nation in a special file at an easy to remember URL (eg. for British citizens `gov.uk`, for German citizens `bund.de` or for US citizens `USA.gov`) that these citizens would be able to set in their browser. If no such information were available, this could also be made clear. This would allow people and robots to know if they are dealing with a real bank or potentially a fake one, if the newspaper is a recognised news source, and what country it is based in and so what legal system it is bound by, and following links allow one even to know who the owners of the company or the heads of the institutions are. This is not something that should or could be delegated to a world wide central agency — since there are too many rivalries globally — but it should in the end be a requirement of the states to publish these in a standard way, so that errors can be understood and corrected by citizens using it using the countries legal and diplomatics mechanisms. It would need the browser to be enhanced to allow it to work off nation based trust anchors — multinational people like me could potentially choose a number of them whereas skeptics would continue to use none (as we all do now). These trust anchors would be URLs that point to resources that describe a nations organs — an organology in Bernard Stiegler's terms. These URLs would link to a document that could link to similar but more specialised documents: one for the list of universities that constitute the nation, others for companies registered there such as `https://companieshouse.gov.uk/`, and so on with links to similar documents published by friendly states. This requires an agreement by some important enough nations on the high level ontologies to use that navigators can understand in order to show a researcher (understood in a wide sense now as anyone searching for knowledge) data about the type of institution he has landed on, their legal character and even how this was found starting from the trust anchors chosen by the user.

Since there are a lot of institutions and companies in the world, it would be impractical and unnecessary for a web browser to download the descriptions of all of them. Rather each web site could link through its TLS certificate or in other standard ways to be agreed on, to one or more official descriptions of its type, owner, ... each of which would be linked to a larger organisations (local authority web site to regional authority) and the browser could then follow these links in reverse and confirm the pages as being tied to one of the trust anchors the user selected. This would then complement the TLS and DNS-SEC (RFC 6698) based security standards with much richer information coming from decentralised trust networks built by the large organisations that form us and that we form, known as nation-states. Just as individuals build trust networks of friends by linking to friends and



acquaintances, so each Leviathan's trust anchor can link to similar trust anchors of other Leviathans. The UK trust anchor could link to all the other nations trust anchors, and those could link in return to the trust anchors of states they recognise. This would then allow institutions of knowledge to be embedded in a decentralised manner into the architecture of the web, without compromising the digital sovereignty of any of the nations. These trust anchors can then be the foundation on which statements of provenance (PROV[44]) can build to help keep track in a much more fine grained way on what basis statements are made.

To make this more practical we can illustrate this with a simple example. Imagine that there is an agreement to allow a company to add a `Link` header[45] on pages it serves of type `companyRegistrationUrl` pointing to its registry as shown on the working example below:

---

[44] The standards for Provenance have already been developed at the W3C, but the following book should be helpful to find one's way around.
Moreau, Luc, and Paul Groth. "Provenance: an introduction to prov." Synthesis Lectures on the Semantic Web: Theory and Technology 3.4 (2013): 1-129.

[45] The `https` part of the URL only guarantees that the server reached is indeed the one named by that URL. The `X509 Certificate` that comes with such a connection may contain a little more official information such as the country of origin of the company and the address of its headquarters. But that information is so minimal as to make the relation to the legal situation of the company completely opaque for anyone other than security experts.

```
$ curl -I https://co-operating.systems/
HTTP/1.1 200 OK
Date: Mon, 02 Jul 2018 07:45:18 GMT
Server: Apache/2.4.25 (Debian)
Last-Modified: Mon, 02 Jul 2018 01:29:11 GMT
ETag: "22fb-56ffa1fec7282"
Accept-Ranges: bytes
Content-Length: 8955
Vary: Accept-Encoding
Link: <https://api.companieshouse.gov.uk/company/09920845>; rel="companyRegistrationUrl"
Content-Type: text/html
```

The browser knowing it has reached the site it wanted to and the connection is secure, would then notice the above `LINK` header and asynchronously fetch that document. The `json(-ld)` result would only need to be enriched with a link back to the domain(s) owned by the company as shown below. Pay attention especially to the domain attribute that I had to add to current the result produced by `companyhouse.gov.uk`[46]

```
$ curl -u token https://api.companieshouse.gov.uk/company/09920845
{
  "registered_office_address": {
    "locality": "London",
    "address_line_1": "2, Harlequin Court",
    "address_line_2": "6 Thomas More Street",
    "country": "United Kingdom",
    "postal_code": "E1W 1AR"
  },
  "undeliverable_registered_office_address": false,
  "has_insolvency_history": false,
  "company_number": "09920845",
  "jurisdiction": "england-wales",
  "company_status": "active",
  "has_charges": false,
  "type": "ltd",
  "company_name": "CO-OPERATING SYSTEMS LTD.",
  "date_of_creation": "2015-12-17",
  "domain": "co-operating.systems",
  "accounts": {
    "next_due": "2018-09-30",
    "accounting_reference_date": {
      "day": "31",
      "month": "12"
    },
    "last_accounts": {
      "period_start_on": "2015-12-17",
      "made_up_to": "2016-12-31",
      "period_end_on": "2016-12-31",
      "type": "dormant"
```

[46] (The `-u token` is required at present by CompaniesHouse to access that page—which is odd, given that the information is openly available in human readable form at the parallel `beta.companieshouse.gov.uk` web site. However it is straightforward to register and get a token.)

```
    },...
}
```

After this request, the web browser would have verified the first link in the chain shown in the diagram as arrows **1a** and **1b**.

But why would the browser trust `api.companieshouse.gov.uk`? After all that could also be a fake website. Or perhaps it once was the right place to look things up, but later the hostname was changed — as it is likely that it will be — and the data is still hanging around there because someone forgot to turn off the machine. We don't want these servers hardcoded in the browser. The way to solve this intelligently is to use the same technique and have api.companieshouse.gov.uk point in a `Registry` header or in the content to the root of the UK trust which would be `gov.uk`. That would in turn link to the registry root domain by specifying that CompaniesHouse was the official source describing companies for the UK. Developing the right high-level ontologies for this would, of course, require a W3C Working Group with technical representatives from the nations involved in setting this standard. With that standardised the browser could verify the second link **2a** and **2b** in the trust chain from `co-operating.systems`. For a UK citizen's browser where `gov.uk` has been set as the root trust anchor, the verification would stop there.

But what about a browser owned by a German, Japanese, Russian, US, Chinese, … citizen? Why would they trust `gov.uk` to state what is the case about a random company? If that sounds implausible, think of it the other way around: why would a UK citizen trust the statement of one of these other countries root authorities? Indeed, how would the browser actually know that `gov.uk` is a root authority, and not just a fake website? Here we continue the process but in a peer to peer mode. We need the states involved to create a web of nations, where each having described itself, links to those it trusts to keep such information up to date. Links need not go both ways, nor be complete, and indeed at the beginning, they won't. This part is illustrated in the diagram by the link formed by the two arrows **3a** and **3b** which would the link followed by a French citizen.

These three links form a chain of trust in an institutional web of trust that is easily verifiable by browsers, but one that is not necessarily globally coherent or complete.

Having such system would make it much easier for small companies to have a place on the web. Currently only very large companies such as Amazon, Apple, etc... can gain people's trust, because they have spent a huge amount of money building it through other channels by branding. People feel comfortable giving Apple their credit card because they know they can trust that company, and they know they are on the right web site — and that Apple would have the money to shut down any fake. But smaller companies that may be just as trustworthy, cannot spend that money on building such awareness. An institutional web of trust would provide the necessary infrastructure for them to be able to have a trusted presence too. I should be able to go to a specialist chocolate shop in a foreign country (to take a random

example) and know that the web site I am at is the web presence of a recognised shop by the local authority in which it is based, which local authority knows the owner to perhaps have a recognised chocolate chef certificate by a recognised institution. This is what is needed to allow the local to flourish in a globalising world.

This clearly would not get rid of all fake news, fake shops, or other fake information. But it would allow the trust systems we have built over the centuries to be put to use in the online world, and because in use to be improved, and then for institutions to grow that can further help citizens in evaluating statements. This would allow us to move from an information society — one where we can receive information verifiably from any place in the world secure that what was sent matches what is received — to a more powerful knowledge society where we can start making and even automating knowledge claims, by making actors making them accountable.