Digital civilisation: manifesto for a trustworthy, well regulated world.

For the SSC Blockchain initiative and the digital civilisation conference

Bill Roscoe
University College Oxford Blockchain Research Centre,
TBTL and OxHainan

Abstract

I present an idealistic but hopefully practical vision of how technology can become more trusted than any individual who uses it as a discussion paper for the Digital Civilisation conference. I concentrate on three aspects of this: control of personal data, the automation of regulation, and improving the trust in the supply chain of hardware and software components that are required to implement it. It is written to indicate the directions a blockchain-based collaborative environment can work, indicating some intriguing solutions based on the combination of blockchain and trusted computing.

The topics of Chapters 2 and 3, namely the regulatory blockchain or Regchain, and using blockchain to remove mistrust from supply chains, can stand alone as well as forming part of a comprehensive digital civilisation.

For details of the Digital Civilisation Conference, see
https://mobile.dcc.global/m/

Chapter 1: What is Digital Civilisation?

Digital technology is advancing at a fantastic, some would say alarming, rate. It is taking over so much of how we interact with the world and each other. However it more reflects commercial opportunities and disconnected attempts to streamline government services than a proper digital civilisation. It offers the chance for a smoother and easier life while

threatening a loss of personal privacy. We believe that society should develop ways of making all this technology more cohesive, more ethical and more respectful of the individual. We think that society should come together and organise the rules by which it expects technology to play, and the technological tools to allow a cohesive digital civilisation to grow. These two things are the core of our digital civilisation initiative.

The user should be presented with questions in the same way by different providers of services, probably after the questions have been subjected to automated analysis.

The internet and World Wide Web are the fabric of this technology and must be at the core of digital civilisation. The needs of digital civilisation will set much of their future agenda. The communications implementing digital civilisation must flow freely and efficiently.

This manifesto sets out the sort of definitions and rules of digital civilisation that we expect the conference and its committees to work on. We would expect, under the overall steering committee, there to be committees on ethics, technical matters, privacy and big data, health data, commerce and trade. We seek a digital commonwealth based on wide participation, wide consensus that achieves a stable equilibrium that enforces total integrity and ethics.

What is digital civilisation?

Civilisation is the coming together of people and peoples to create an ordered and functioning society that builds a safe and generally peaceful life for its inhabitants. It enables services and culture for the people to emerge. Civilisations have well organised government that keeps society together. In a harmonious society the rights of the individual are protected in balance with others' rights and to some degree the need to keep a stable society. In the modern world, Governments should exist and function only by the consent of the society they govern.

Digital civilisation provides structures through which we interact with governments, companies and each other, guaranteeing transparency, uniformity and adherence to common principles and rules.

Civilisation is too important for us to allow Big Tech companies to design it for their own benefit. They have a huge role to play but should not be allowed to design its rules or to gather huge amounts of data on us without a much greater degree of informed consent than we see at present.

Civilisation is thus a combination of stable government, the tools and components that enable society and the people in it to function, plus the people and organisations that exist in it. The components are things like healthcare, schools and universities, the trading of money and assets, and taxation. By tools here we mean things the establishment of laws and the means such as courts and regulators by which these function Civilisation provides

the platforms for commerce and trade. In the modern world, civilisations have to function in harmony, and while individual cultural and national groups may develop distinctive aspects to their civilisations, there needs to be a global concept of civilisation within which all operate and a global infrastructure that supports it.

Civilisations tend to break down when the way they operate is contrary to the consent of the people, when there are too many opportunities for unfairness or corruption, or where the clashes between interest groups or neighbouring civilisations become too stressed.

In the modern world, digital technology enables rapid interactions and takes over, for example in social media, trading and the holding of data. It is also providing undreamed of possibilities for activities through "big data": some of these are plainly civilised, some plainly not, and some open to debate. In order for it to fit into civilisation, this technology must itself be civilised: an expression in which "civilised" can equally be read as a verb or an adjective. In other words, such functions should be rule-based where the rules maintain the principles of civilisation.

Digital technology can replace or supplement mechanisms from traditional civilisation, where these can be made more efficient or offer ways to avoid the fallibility of humans, for example in the integrity and transparency of procurement processes, the running of elections, or the keeping of public records.

We argue, however, that to serve such purposes in a truly civilised manner, the processes and programs implementing such functions of society need to be made transparent and subject to inspection and review; and if necessary, modification.

The present civilisation we live in developed over thousands of years and is far from perfect. The rapid development of technology that is driving change at the moment does not give us time to reflect and is too vulnerable to technology companies and governments driving it to suit their own ends such as making money or suppressing dissident opinion. Digital civilisation must respect the values of traditional civilisation and must look like evolution rather than revolution to those whose lives it affects. Ethics and agreed ethical standards must govern it.

Participating in the digital civilisation should become a commercial necessity for governments and companies large and small for access to trade, data and markets. Its success can be encouraged if given special status in the management of identity and personal data. For example, it might implement uniform standards and technology for the collection of data that could be mandated by governments or insisted on by public demand. It should not be acceptable to have to sign up to opaque "Terms and Conditions" to gain access to services without these being scrutinised automatically by utilities provided by digital civilisation and compared against personal limits.

Civilisation as a set of rules

To ensure that society is ordered and well behaved, we need accepted rules about how its participants act. Both regular and digital civilisations need mechanisms to enforce such

rules, such as legal processes and to disable illegal acts. We want players to follow rules, perhaps because they want to be seen to do the right thing, or because they know they will be found out if they do not.

It follows that in digital civilisation we need a clear way to express rules so that they are understandable and checkable efficiently. Furthermore, they must be unambiguous except in rare cases where there is a resolution mechanism that might well involve human courts. We note that the legal profession thrives on perceived ambiguity of the rules of traditional civilisation and that this can cause vast costs and delays, so we must try to learn our lesson.

Things must be done, seen to be done, and seen to be done fairly. The conditions including about use of data that users are expected to sign up to obtain services should be subject to automated analysis against general standards and against a user's preferences. They should probably be formulated in a suitable formal language.

I see this as hugely important not only as practical solution to achieve proper consent to use of personal data, but as a means to move the balance of power from data collector to consumer since it will provide a far better and indeed collective way to resist unreasonable terms. It is almost as though consumers are organised into a trade's union.

In a single country the population might in general terms trust their government but would far prefer it if their government was following its own transparent rules. In an international setting there is probably no single trusted party, as in many areas of commerce. Digital civilisation should solve these problems: it must be created to achieve complete trust and trustworthiness. The fact that a given government or tech company exists within the digital civilisation should increase the citizen's trust in it. We might well cede the right to a government or agency to perform aspects of administration, but only if that agency could prove it had not overstepped its own rights and had followed the rules.

At the moment trust in the use of data by many parties including Big Tech is low. Initiatives such as GDPR are good, but need to be understood, obeyed and seen to be obeyed. Automated analysis of consent conditions must play a part.

Digital civilisation should ensure fundamental rights in areas like data and property and make it much easier to ensure integrity of processes, and of interactions between parties. The rules that are created and mechanisms preventing the rules from being breached should be the means to achieve this.

We see digital civilisation as having global rules and resources that must be agreed on an international forum rather than imposed by a single government or commercial company.

The role of Blockchain

For digital civilisation to be *civilised,* we want to enable bad behaviour to be almost impossible and unprofitable. We want its basic structure for records and transactions to be unimpeachable and incorruptible. This suggests something like a *Blockchain* in which strong regulation and the identifiability of participants apply. Such a Blockchain would ensure

correctness and integrity of records, where breaches of rules may well simply not be permitted or at least will be made apparent to all.

A Blockchain ensures integrity and transparency, as well as adherence to rules. For no record will be admitted into it and parties attempting to engage in corruption will be blocked and possibly penalised and exposed.

Rather than an energy wasting means of trying to magic money out of nothing, we mean blockchains designed to be green and essentially mechanisms for implementing and enforcing public systems of rules.

We think of a blockchain as a forum for establishing, implementing and automating rules and laws. It should make breaking the rules of civilisation either impossible or simply not worth doing.

Properly implemented, I believe that a Blockchain is fundamentally a civilising influence, essentially preventing corruption and the falsification of records. We may well form a global Blockchain, whether public or permissioned, out of ones that define and enforce the rules of particular jurisdictions. The global chain would enforce global standards on the component parts. In order to scale up to the role we are anticipating, it is unlikely that a simple linear blockchain will cope, we will need to consider more developed and decentralised ideas, though these might well have a conventional blockchain at the core.

There is much research to be done to make this work, but both we and others have made a good start. To create this ledger of digital civilisation we require technical, cryptographic, network and systems research, as well as research on the expression and automation of regulation, the implementation and logic of KYC, legal, ethical and economic research, and probably many other things as well.

The components of civilisation

A civilisation is held together by the spirit of the bulk of its members, but also by systems and conventions covering particular aspects of life. As much as possible these should be consistent with each other and work together. So, it should be for digital civilisation.

We should make a start by implementing components of digital civilisation to bring its benefits, while contemplating the big picture of how these and a wide rule base can build the basis of a better society, and how such a society can be constituted and operate. We must make all corners of society believe in it: the people, commerce, trade and governments, because it both promises and delivers a more transparent, better behaved and efficient world. It must enforce the principles of good order and good government, without being dominated by governments. Different rules and laws will apply to different parts of it, but these rules and laws must be clear to everyone and applied without bias.

The core components of a digital civilisation are an immutable archive of what has happened in the past, and a system for verifying the identities of players in it. This must be, or be as close as reasonably possible to, a way of entities outside the system proving who

they are and connecting them to their presence in the digital world. We would also expect it to define the assets and data owned by each of the people or organisations.

Having a single Blockchain or similar structure underlying the components of our civilisation will be a tremendous aid to their cooperation, as will common data standards or at least committed relationships between them.

There is no assumption that everything in the system is public, and the boundaries between public and private data and operation need to be carefully drawn, as do issues such as the jurisdiction and accountability of the parties. We think it unlikely that completely secret (even from legal authorities) assets will be acceptable: experience in both the digital world (anonymous digital currencies) and non-digital ones (tax enclaves and overzealous banking secrecy laws) illustrate this. These last examples also demonstrate that the rights of different subsystems, whether they are banks or nations, to create their own rules and privacy standards, should not be unlimited.

Initial components of the digital society will be access to and in many cases delivery of e-Government, defining the relationship between an individual and levels of government. Supply chains including specification, tendering, subcontracting, verification of supply and settlement represent an identified demand that can easily be incorporated into digital society, contributing to the elimination of corruption. Similarly compliance of tenders can be built into the rule base rather than being left to subjective judgement.

The Blockchain enables all sorts of distributed decision making, whether e-voting in any number of contexts, auctions, tenders and lotteries. It also allows us to create many types of markets and exchanges. Property registers can be created both for commodities and identifiable items such as individual diamonds and real estate.

We can implement licensing systems for people and organisations in performing roles; licensing software, digital media and IP. Such systems can monitor the use of licences to ensure compliance.

In many cases such a digital civilisation will be able to provide services more efficiently than traditional means, including the oft quoted examples of currency clearing, financial products and insurance. But such big changes will sometimes need to include roles for traditional players because of the need for expertise, financial resources and trust. In these areas disruptive providers may emerge, or established providers may succeed in adapting. Regulators such as central banks may encourage the adoption of the new digital technology for reasons of efficiency.

There are many potential applications in medicine, such as health records and research such as clinical trials.

Ethics

The applications above frequently require high ethical standards if they are to gain the trust of the public. We see now how individual data-based industries are criticised for their

actual use of data, their lack of security with it, or simply the opacity of their use of it. We hope that such issues can be addressed by building ethical principles into the collection and use of data. The digital civilisation can provide a common regulatory framework and thus greater understanding and confidence to all.

The same goes for other aspects of administration and trading. In developing digital civilisation we would expect a process of developing ethical standards and indeed demand for these. Of course the objective, as far as possible, would be to build these into the automated regulatory structure.

In defining the digital civilisation, one of the most important aspects is distinguishing between global principles that must be adopted by all participants, and ones that can vary from jurisdiction to jurisdiction. For example a method of data collection or analysis, such as face recognition technology, might be legal in one place but not another. A global principle must be to make such rules transparent, so in any place one must know just how this technology can be used. Even if facial recognition technology is not being used, is it legal to record photos and videos on which it might be used *later?* If facial recognition is illegal in San Francisco, then we might expect it to be illegal for images recorded in San Francisco to be processed in Dallas, even if facial recognition is generally legal there.

What does digital civilisation look like?

If you are a real citizen, what is it like to be digital one in the world we imagine?

- Your link to it is provided by your identity, which is a closely controlled collection of information that you and possibly heavily regulated authorities have provided about you. Using it you will be able to prove who you are to other parties in digital civilisation without giving them this information.
- You can also share such information as you want to, but for protection of the individual there may be limits on this.
- In any context where you are giving such information or allowing parties to see or collect information about you, the permissions and possibly the collection itself will be moderated through a system that compares requested use with your own preferences, and hopefully records actual uses. Thus you are not constantly confronted with different formats and needing to read terms and conditions.
- It thus acts as a guide and helper for your digital interactions involving your identity or personal data, as well as policing the use by others of these things.

Principles of digital civilisation

1. **Digital civilisation is technology in the service of humanity. Innovation is required in technology, economics, government and ethics. It should respect society and the rights of the individual.**

2. **Technically speaking, it is implemented on an enormous collection of organised data, contracts, transactions and assertions, to which any participant can contribute. The evolution is governed by rules.**
3. **Rules are public, clear and ethical. The rights of organisations and nations to draw up and modify rules are themselves public, clear and ethical. Thus rules are themselves subject to rules, such as not contradicting basic rights.**
4. **Adherence to rules should be easy to verify: where complex calculations on state are required, the Blockchain (i.e.. collective decentralised architecture) maintains sufficient of these calculation to make individual checks easy.**
5. **Transactions are permitted if and only if they satisfy all rules that apply to them, which may vary with the jurisdictions of the parties and items involved.**
6. **Data and its provenance are readily verified. Our aspiration is that all organisations collecting and dealing in data adhere to the principles of digital civilisation.**
7. **Personal data other than basic identity is private, and the owner can control and monitor its use: we believe in the principles of GDPR and truly informed consent.**

In the next two chapters we expand on two other wonderful applications of blockchain as a vehicle for trust and integrity that build on the above.

Chapter 2: Making Regulation Sexy

Introduction: the need for regulation

Properly regulated finance, government, companies, healthcare, and both service and manufacturing sectors are the core of society. But regulation is normally seen as boring and cumbersome, an unwelcome encumbrance on everyday life. In this paper we argue that good regulation is essential, and that efficient automation in a trusted environment can overcome many of the drawbacks. I further argue that by moving regulation into a coherent technological space and making them intellectually satisfying we can attract new and positive attention to the subject.

The present chapter is a continuation of my digital civilisation concept, giving more substance to what it says about the implementation of regulations and laws in a world run by a decentralised digital framework. The regulatory framework described here would make sense in less ambitious projects as well.

Some well known examples

First let us discuss some failures of regulation, and places where more and better regulation was required. Note that in the following I am giving my personal opinions rather than claims of definite facts.

737 Max flight control.

It is widely asserted that the 2018 and 2019 crashes of this aircraft were closely linked to an automated anti-stall feature of the flight control activating because of erroneous data from a sensor. It seems quite remarkable to me that a system in which the failure of a single sensor could cause a crash was certified. I cite

1) The well documented crash of an Air France flight in 2009, where a faulty sensor was a major factor, ought to have alerted regulators and manufacturers to this danger.
2) Fault tolerance, including multiple redundant processors, is a well established principle in avionics: it was obviously not applied properly here. I was involved in projects with the UK and US defence sectors from the early 1990's on involving the creation, specification and analysis of fault tolerant systems, and the analysis of legacy systems and combinations thereof for fault tolerance. In many cases these exercises were undertaken to establish that the systems concerned met regulations for release to service. Of course my specialisation was proving safety (under defined fault assumptions) by formal verification, but if that had not been available the analysis would have been dome by rigorous code inspection and testing regimes. I am at a loss to understand why similar verifications or systematic studies were not done here, or included in regulations. An outline of the procedures and regulations under which much of this work can be done can be found at **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/27552/WhiteBookIssue3.pdf**.
3) A document which I identified on structuring assurance cases treats no-single-point-of-failure as its first example. **https://scsc.uk/r141B:1?t=1**

It is alleged
https://amp.theguardian.com/business/2020/mar/06/boeing-culture-concealment-fatal-737-max-crashes-report
that there was too cosy a relationship here between manufacturer and regulator and a greater emphasis on costs than on safety by the manufacturer. This example demonstrates that regulations themselves, such as fault tolerance and the principle that no single (or perhaps greater) sensor failure should have a significant effect on the aircraft, should both be public and publicly certified (with full audit trail) as part of certification. While such work may be done by a manufacturer, we should not have to trust the manufacturer to know we may depend on the results. (See the section on Personal reward below.)

The lessons to learn from this case are that the public need to trust the standards imposed for safety critical infrastructure, whether aircraft, cars, bridges, self-driving vehicles or medical equipment, and that the said standards are actually being imposed accurately.

As with other cases of regulated systems there is a clear tension between the legitimate need to keep many aspects of the system secret and data and proprietary, against the public interest in knowing that the said material is fully compliant. We will discuss this later.

Both this case and the Deepwater Horizon oil spill
https://en.m.wikipedia.org/wiki/Deepwater_Horizon

where there were also issues of regulation and compliance demonstrate perfectly how not paying proper attention to regulatory processes, supposedly in pursuit of profit, can lead not only to human tragedies but to corporate meltdown. Note my analysis below on **personal reward.**

Sudden failure of audited public companies

In recent years a number of large UK companies (e.g. Carillion, Patisserie Valerie) have failed
https://en.m.wikipedia.org/wiki/Carillion
https://en.m.wikipedia.org/wiki/Patisserie_Valerie

Of course this has gone on at many times in many places, and often as in these cases investigations call into question the quality of the statutory audits that had happened for the periods leading up to the collapse.
https://www.theguardian.com/business/2019/feb/01/decline-in-quality-auditors-face-scrutiny-over-string-of-scandals
It would be better if audits and the rules applied could themselves be audited, particularly when problems surfaced. Hopefully if this were possible it would place more pressure on audit professionals to conduct audits thoroughly, and allow better rules for risk assessment to evolve. Transparency and adequacy of process are vital.

The suspicion always is that audit firms develop too cosy relationships with their clients, so the more transparent and externally verifiable an audit is the better. Of course exactly the same was said about the relationship between FAA and Boeing. We can be sure that the only cases that reach the public consciousness are those where these failings lead to catastrophe, and that there must be many more cases where unreasonable risks are taken but these fail to materialise.

It must be demonstrable that the auditing of compliance with regulations is done fairly and for the benefit of the world at large, rather than simply for the benefit of the entity being audited, or indeed (in the example of when an outside party imposes an audit) for anyone else specific. Again we need to be conscious of the need for confidentiality.

Financial crashes because of wrongly assessed or disguised risks (e.g. 2008)

Banking and market regulations, on a macroscopic scale, clearly failed badly in the great market crash earlier this century. Others have analysed the background to this in great detail, for example
https://corpgov.law.harvard.edu/2010/11/20/the-financial-panic-of-2008-and-financial-regulatory-reform/
https://www.thebalance.com/2007-financial-crisis-overview-3306138
The main causes here seem to have been

1. The willingness of US institutions to lend money that the recipients could not afford. In doing so they breached what seems to be a basic regulatory requirement: quite aside to the damage done to the mortgagors, they were not pricing in risk.

2. The fact that the said institutions were then able to package up and securitise such loans and derivatives based on them in such a way that disguised the underlying risk from the purchasers of the securities.
3. The role of the US government-sponsored companies Fannie Mae and Freddie Mac in creating false confidence **https://www.investopedia.com/articles/economics/08/fannie-mae-freddie-mac-credit-crisis.asp**
4. False assumptions about the independence of risks: if you toss a thousand coins the laws of large numbers give you high confidence in the distribution, but a thousand mortgages lent to stretched people sensitive to the same dip in the economy, leading to a crash in property values destroying mortgage security, will not behave independently.

There appears to have been failure on the part of regulators to ensure that risk was properly modelled or communicated to purchasers of securitised loans, aside from an overall over adventurousness in lending and insufficient capital reserves. Certainly banking regulations and oversight changed rapidly after the crash.

I regard it as a failure of the regulation process that the risks involved were not apparent to either the central banks or the purchasers of securities.

Discussion

These examples and more have given rise to discussion of the nature of regulatory regimes. **https://www.regulation.org.uk/key_issues-regulatory_failure.html**
The above link also quotes further examples. Of course it would have been far better if such failures had been avoided or discovered before they caused disasters: a cynic might regard fixes subsequent to them as fixing the stable door after the horse has bolted.

The risk of personal reward

A possible factor in all the cases I have quoted is a misalignment between the interests of a company and the personal interests of the people in the companies who make crucial decisions. If traders and CEOs can make enormous personal gains from increments in profits but no personal losses if they go horribly wrong then it is understandable that some of the situations I have described have arisen. An analogy might be a hedge fund manager whose contract gives him 10% of his customers' profits but 0% of their losses. If he then puts his whole fund on the favourite in a horse race, and perhaps other funds on other horses in the same race, he has a wonderful strategy for himself, but a terrible one for his investors.

Similarly, if a company's CEO can double its profits and hence his salary by paring back safety measures in a way permitted by patchy regulation or inadequate regulators with only a 10% chance of suffering a catastrophic failure during his tenure, there is the serious chance that he will.

It may be close to impossible to remove this asymmetric reward strategy, so effective and soundly based regulation is needed to prevent corners being cut on safety in pursuit of profit or traders using investment strategies whose risks have not been properly assessed.

Desiderata

I argue that several things must be achieved to make regulation work better:

1. Regulation must more much higher up the consciousness of the people at large and of those working at various levels. Its role must be understood and fully embraced. It should be seen as an enabler rather than a mass of red tape.
2. Regulations themselves must be understandable and justified. They should be public and open to comment and debate, just like academic research.
3. They must demonstrably be applied accurately, fairly and universally. It should not be possible to use power, money or contacts to obtain lenient treatment
4. The process of checking regulations must itself be open to audit.
5. While there must be sufficient regulations to eliminate unacceptable risks and behaviours, they should not be unnecessarily onerous or heavy handed.

Above all, regulation must be seen as important and fashionable. A career in regulation should not be seen as obscure and boring.

We can make this possible by providing a technological framework which makes all of this possible. The technology can make the fairness, transparency and audibility automatic, and by bringing the excitement technology of to the subject.

Why blockchain?

Blockchain is a technology that allows data and actions to be recorded with complete integrity: there is no doubt about what was said or done, when this happened and who did or said it. This is perfect for establishing the facts that are subject to regulation, and exactly what has been done to check regulations, by who, it can also record the regulations themselves and what process is required to verify them.

A single blockchain can straightforwardly contain this diversity of information. In some domains the things to be regulated can be completely be recorded or contained in digital form in the chain: for example the transactions in a digital bank, or data and accesses to it. Other things particularly well suited to recording in them include (stages in) processes that are performed digitally, rules, regulations and their evolution, access control policies and actual data accesses as well as edits and modifications.

In others, including stages in processes that are performed in the physical world or record human judgements, we can specify what representations of the regulated object have to be registered in the blockchain, be it a design, process, opinion or object, and how inspection, testing or similar are recorded.

By publishing records of what is done, the standards it is expected to satisfy, and the verification process in the immutable and timestamped blockchain, we can hugely increase confidence in what is done. By opening up regulation and the process of it to public examination we should be able to eliminate the unnecessary and spot weaknesses.

We can make it both the culture and the requirement of the blockchain that a high level of scrutiny takes place, hopefully eliminating the possibility of inadequate regulation.

There are issues of privacy that we will need to address later, but as a general rule I would expect that regulations themselves and their evolution will always be public.

Privacy with integrity: the power of hashing

Two arguments are frequently levelled against blockchains: one cannot keep data secret on them and you cannot delete things from them. In fact it is not difficult to do these things using encryption and/or cryptographic hashing.

Note that privacy and deletion of data (especially personal data) is frequently the subject of regulation itself. It would therefore be perverse to use a data storage technology that did not permit them.

A cryptographic hash gives a digest of a value v which, if it is placed in a blockchain (probably signed by the originator A) represents a commitment to v that is irrefutable and undeniable by A. However, provided the hash is made of a salted version of v (i.e. combined with a random addition), v is not revealed by its presence on the chain. It can still be made public by a link to an off chain location (or more than one if appropriate, with the salt used for it so its equality with the committed value can be verified). To delete the item, the target of that link can be removed.

The use of cryptographic hashes has the advantage that it reduces the size of what has to be recorded to a small number of words, no matter how large the things being committed is such as a scan, photo or video.

Of course encrypted data can also be stored in a blockchain, which will ensure that it is always available to someone who knows the key. Depending on the circumstances this does not necessarily commit the underlying value v: if not that is not satisfactory a cryptographic hash of v or the key is required too.

Of more interest is the combination of the two: using a compact representation as a hash in the chain, combined with a pointer to an encrypted copy being stored elsewhere. It is probably better, in fact, to store the combination of an encrypted copy and a hash of the original in the off chain location, and a hash of this combination in the chain.

Of course where data is stored encrypted either on or attached to a blockchain, it is not accessible to regulators or anyone else unless they have the key. So how encryption keys are managed is a vital part of any such use. In effect they can pass data around by communicating the relevant key, but cannot manipulate the contents of the data.

Elementary regulation in blockchain

In this section we examine what can be achieved when using a blockchain for regulation without further technology beyond what we have already discussed.

There is a great deal to be said for storing a complete regulatory environment in one blockchain: regulations themselves, including the regulations surrounding auditing,

representations of the things that are being regulated, and the records of auditing and regulatory checks being done.

Consider the case of standard financial auditing. We would expect that the company being audited would ensure that its own financial records were linked to the blockchain in a timely fashion, together with supplementary documentation required to support the audit process, including its side of interactions with the auditors and relevant authorities. The auditor would be given access to the keys and would be expected to include substantial records of what it has done and examined and how this choice meets the requirements on it. Such records (which would probably also include the key to the company's records referred to already) would themselves be encrypted with the key either known to a central regulator or in the hands of an escrow service which reveals it to the regulator in response to legitimate requests. Or the key might be split between several services to ensure only legitimate access. The key would presumably also be known to the audited company.

There might be interesting questions about who has custody of the off-chain data, encrypted or otherwise, and under what circumstances it can be, or must be deleted. Is it permitted or perhaps compulsory to replicate it, largely relating to GDPR regulations and similar? Should depersonalised copies of data be kept and what regulations pertain to these? I have never believed that keeping personal records intact but changing name, address and DoB is remotely sufficient.

The main output of a financial audit would usually be public, as annual report and accounts plus audit report, and therefore placed publicly on the blockchain or securely linked to it. (There may be a more detailed private one.) Of course audit reports for investigatory purposes might not be public. In any case an audit report could now contain links to evidence on the blockchain.

It would be natural to place other material relevant to compliance on the same blockchain such as accounting standards, audit and disclosure rules and changes to these.

Whenever data is stored, particularly long term, care should be made to record data standards and other metadata.

With this model the process of accounting and auditing would be very similar to the current world, as would be the underlying model of trust. The blockchain would add to this trust by greater data integrity.

Automating Regulations

All we considered above was a new way of handling data relevant to existing audit processes. There are clearly opportunities to define data formats for the data that is regulated, the process of deciding whether they are satisfied, and turning the regulations themselves into programs that decide them.

Many banking regulations, which define when a transaction is legal, what sort of checks are necessary on individuals to hold or transfer assets, the reserves that must be held by banks themselves, are amenable to determination by a combination of logical computation and

calculation: given the information that the bank holds about itself and its customers, the active regulations etc, it can run a computer program to decide whether or not the regulation holds.

A bank is, however, accountable to its customers, its regulators and the community at large for its adherence to regulations, so it is not enough for it to satisfy itself that it is applying regulations properly. The same goes for many other regulated entities.

Of course that is the standard function of audit: a trusted third party (usually chosen by the audited) from amongst agencies trusted by regulators to carry them out) goes in an checks that regulations have in fact been followed. That process is, however, anything but real time and, given the examples above, does not appear to be fool-proof. Banks typically have internal audit and/or compliance teams. My nephew's first function on working for a bank was to create spreadsheets that tested each day's transactions for compliance, so only just post hoc.

Of course there is no reason why a bank or other organisation's own systems and software should not continuously monitor regulations, and clearly they frequently do. What we would like is a reliable and accountable way to reassure other stake-holders and the world at large that this being done and that the outcome is satisfactory. This is a matter of the quality and completeness of the monitoring and the trustworthiness of whoever is doing it or takes responsibility for it, with clear issues arising about privacy.

In terms of integrity of process and data blockchains have the same advantages as previously: the programs ensuring and checking compliance can make entries stating what they have ensured for what prior states of that chain.

Delegating trust to machines

In the previous sections we were using blockchain to make the existing regulatory system more transparent and trustworthy by ensuring that processes can be checked unambiguously and records cannot be tampered with. The various parties involved all use a reliable and trustworthy system for keeping records. One of the most important tensions is between privacy and accountability: if A is conducting a regulated business involving private data then whoever (say B) is auditing its observance must be trusted not to misuse that data and must be trusted by all to do the correct calculations. This problem grows if we expect B's work to be checked, for example by replication of its work.

This illustrates a familiar tension in security: in an environment where agents are not absolutely trustworthy, replication of calculation is good for ensuring that the correct result is returned. It is unlikely that multiple agents will all be bad and agree on the same wrong answer. On the other hand the chance that one of the agents misuses or gives away the secrets increases steadily with the number of replicated verifiers.

There is also the tension that there is frequently a suspicion that auditors may be partisan, for example towards those who choose or pay them. Depending on the jurisdiction it may or may not be permissible for auditing firms to market other services to their clients, which it is argued may well amplify conflicts of interest.

A novel technology should be able to help us to overcome this paradox, and indeed to ensure professional standards in a variety of areas.

Imagine a computer chip that can prove to its user (even one who has no physical access to it) what it is: it is a particular model of processor created by a specified manufacturer. It may well be able to prove more about itself such as its location. Furthermore it can attest that it has performed a particular calculation using a program that is known to the observer. If required it can decrypt the data it reads and re-encrypt the results (including further data it may store for itself and others to read). Its owner is unable to extract data from it. Then if it is running a program that is certified by a regulator and verified to carry out regulatory inspections and to protect data privacy, then everyone should trust its results provided it can be proved

1. That the correct data was used.
2. That the protocols for data transmission and key management used to communicate with this Trusted Execution Enclave (TEE) are secure.

The natural way to ensure the first of these is with a blockchain or similar: it can be proved that the TEE was proving the data that is agreed by everyone is the true record. The second is simply a matter of security engineering. A possible solution to the second problem is to ensure that there is a regulatory TEE (meaning one with access to the necessary programs for regulations) within the secure systems of the regulated party.

There would of course have to be very clear rules about the handling of data by the regulator TEEs (let's call them RTEEs), that would be built into the programs they run.

TEEs are a rapidly developing technology whose security is being both developed and probed by researchers. TBTL has a very rare level of expertise both in the theory and practical aspects of TEE security and of developing systems for them. Both SSC (OxHainan) and TBTL have identified RTEEs as a core topic for research and product development.

TEEs are very useful in virtually any decentralised context where mutual trust between parties is an issue, including blockchains themselves, identity management and KYC, and the execution of smart contracts. This is especially true for ones where data is intended to remain private. In this paper we concentrate on regulation,

What is risk?

Much of regulation involves attempts to limit risk. Looking at our three initial examples:

I. It is clear that regulations or their applications failed to identify that the risk of a disastrous misbehaviour was far too high. Clearly with real life machines operating in the real world there are risks of component failure, bad weather, bird strikes etc. One of the main functions of regulations in this area is to ensure that despite such risks the likelihood of a disaster is acceptably low, namely close to zero.

Thus stringent testing and verification are typically specified for software, and typically redundancy for many electronic, computing and physical systems.

II. In financial auditing there is implied risk assessment in judging the likelihood of debts being paid. There are risks in the audit process itself: sampling transactions to investigate in depth and setting "materiality" thresholds that are not disclosed to accounting personnel are both methods of limiting risk rather than eliminating it.

III. A large factor in the financial crisis appears to have been the underestimation of the correlation of risks of mortgages defaulting.

It follows that a full analysis of RegTech — Regulation Technology — and systems to support it need to have extensive support for risk management. There are a number of ways we might choose to classify risk.

There are risks that are accurately quantifiable, perhaps by bounds on the risks rather than the accurate probability on the bad outcome happening. As we have already seen, while the likelihood of an individual event may be predictable, but it may be hard to correctly anticipate the pattern by which these correlate. Of course we can learn from previous cases and provide support for this type of independence analysis.

There are risks that are traditionally classified according to a discrete scale, often in a matrix where the probability is classified on one axis into verbally described categories such as *likely, unlikely, very unlikely* or *negligible chance,* and the other axis gives a measure of the impact in similar terms. These are frequently converted into colours to help visual processing. Sometimes also one is expected to document risk mitigation activities and re-assess the risk after these. This type of analysis is frequently standardised or semi-standardised into risk evaluation processes which attempt to guide users into covering all the angles of risk.

Sometimes a distinction is made between the term *probability,* meaning a number between 0 and 1, and the term *likelihood,* meaning a categorisation such as the above. One can link the two together by placing probability bounds on the categories, but of course this is only correct if it can be justified.

Some situations more or less defy probability, such as assessing the likelihood that a complex system is secure against cyberattackers when this has not been proved. One can know that some modification makes this more or less likely without being able to justify any probability. To do so means removing all "unknown unknowns" and reducing the "known unknowns" such as the amount of computing the attacker is prepared to spend, or their ingenuity to a probability or worst case.

Regulations may cover aspects of such processes: ensuring that assessments are properly done, signed off, and reviewed. Of course they will frequently specify that the results of such assessments give satisfactory answers.

Where a risk or aggregate risk (i.e. the chance that more than some possibly weighted measure of a collection of events occur) that can accurately be at least bounded, regulations may specify some bound that must be bettered. Where they can not regulations might still

assume unjustified figures, or will need to eliminate any possibility they might occur, or simply

The perception of risk, either by the public or by "experts" may be different from actual risk. There is a strong sense in which regulations may be skewed towards perceived risk, or that there may be demand from the public to address risks that they exaggerate. The now-current Covid-19 pandemic provides a case study in how risk and the public perception of risk are managed: I am sure that people will be writing about how this and the contrasts between countries for years to cme.

In an ideal world a comprehensive risk support package based on the mathematics of probability and statistics, the analysis, classification of risk according to established standards, should be included in any RegTech support environment. Research on doing this in a consistent and sound way is very important.

A vision of the future RegChain

We have just scratched the surface of the areas that need support from regulation and where good regulation benefits stakeholders and the world at large. It is clear that regulations in different domains have individual characteristics but also similarities. Our vision is to create a framework that allows unambiguous formulation of laws and data standards that underpin their decision framework, across most or all known areas.

We can look to the blockchain model to help define this, since clearly we are looking to unify regulation using it. A blockchain has participants who contribute entries which are assembled into blocks by participants which are checked by others, and a consensus process to resolve who gets to create each block and how the results of checking affect it. The blocks provide an evolving record of what is agreed to be true.

In a RegChain, a regulated blockchain, one of the main decisions is the extent to which regulations are decided before or after the blocks are formed. Does the system accept an addition that is not known to satisfy the underlying rules? Can adherence to regulation be decided absolutely unambiguously on the basis of the rules, the data, and agreed logic or do we need a resolution procedure involving human or other judges?

In a pure RegChain

a. The blockchain's basic structure, as with all blockchains, is laid out in the genesis block. So are the logic for deciding regulations, the initial regulations, and the protocol for changing regulations. We would probably also want the basic rules for deciding on the jurisdiction of an agent, asset (including things like data) or transaction. By *jurisdiction* here we mean not only the nationality or locality of the respective thing, but also other factors that affect what set of rules apply to it. So we are interpreting jurisdiction literally as the "system of laws" that applies. The genesis block would contain initial rules for determining this, but also rules by which jurisdictions themselves and perhaps restrictions on how lower level rules for classification into them can evolve.

b. It will thus be clear that neither regulations and laws, nor the definitions of jurisdictions, should be immutable. That would be very convenient but leagues away from reality. However it is vital that proper processes are followed in changing these things: rules must always satisfy the rules, or *meta rules* about what rules are permissible. Furthermore they should only be changed when the process of change itself meets *legislative rules* laid down in the blockchain. Clearly both of these categories of second-order rules can vary between jurisdictions governed with the same blockchain, subject to overall limits on what they do or do not way. While plainly meta rules and legislative rules can potentially be changed, I think that in initial experiments and implementations this might be too complex.
c. When rules change, they might make a previously legal state or transaction illegal. The resolution of the second of these is easy: the rules that apply to a transaction are those in force at the time it is adopted into the chain. A legal transaction should not be made illegal later, just as in the legal world *retrospective legislation* is frowned upon and often thrown out by courts. It is possible to imagine a rule change, however, that makes a previously legal state illegal. Examples might be location or recording of data, the distribution of assets or the structure of existing securities. Legislative rules should allow for this, for example by preventing rule changes that make anything illegal that is incapable of correction, and by allowing parties warning and the ability to correct where this is happening.
d. It will sometimes be desirable to include the state of the world outside the blockchain into regulations: *if it is raining the speed limit is reduced from 130kph to 110kph,* or *a patient's heath records may be accessed by doctors treating him or her when it is impossible to give rapid consent* or a legal case requires a change in rules. For such things there must be an agreed protocol for getting the information into the blockchain so that it is unambiguous.
e. Some questions of law and regulations may be too difficult for an automated decision process and require investigation and arbitration in "slow time". It should be perfectly possible to manage this by allowing things to be entered with a marker saying they are *subject to adjudication* or similar. The subsequent decision making process about them can also be recorded in the chain, and almost invariably the original record that is being adjudicated will not enter the official records until after it has been approved as legal.

The regulatory framework we imagine is thus logically very much like the present ones found in the real world. Defined rights to set regulations and set out their scope, governed by a legal framework, but where the need to eliminate ambiguity should force more clarity and enable arguments about timing and provenance to be avoided. We want to create a system which is transparent, and transparently fair, rather than a land of milk and honey for lawyers. The core principle is that, except in well defined exceptional circumstances, only valid entries of any sort make it into the record.

The language of regulation

We can classify the things that are regulated as

i. Processes: participants are obliged to follow: risk assessment, audit, safety assessment, managing elections
ii. Access to data such as health and financial records, for writing, reading and the right to run analytics over it.
iii. Accountancy and tax
iv. Transferring assets such as money, securities, commodities and individual items
v. Physical products such as foodstuffs and safety-critical items. Such regulations may include linkage with particular digital assets: each physical asset has a digital twin; an immensely useful concept for supply chain management.
vi. Transportation of people and goods
vii. Infrastructure: roads, energy, water, internet
viii. Media and hopefully social media.
ix. Construction and city planning
x. Insurance
xi. Identity and proof of identity, including KYC,
xii. Government and security activities requiring regulation

This is an enormous range and each of them has many variations. The principles we have developed apply to them all and more. The same basic logic applies across the board, and the essential tools of regulation will be similar such as certifications of facts by one or more authorised parties, computations over the state, assessments of risk as discussed earlier.

I imagine that we will create specific support for areas such as the above relying on core support. It is clear that this is an enormous project. Furthermore each of the above topics have many subtopics. Through things like metadata and means of accurately describing regulation and regulated things, such as *ontologies,* and the design of suitable logics and programming languages, we need to bring all the topics of discourse with the scope of our systems and digital civilisation.

Clearly this cannot all be done at once, and each and every area will require close collaboration between experts on this new RegTech and experts on regulation in the given domain.

We aspire to create a technological basis which, while it produces necessarily distinct solutions for different application domains and sometimes systems of law, will be based on a common framework that allows for interdependencies and integration, and which provides solutions that are quicker, more transparent and cheaper than the current methods.

The typical regulation will take one of the forms

- An assertion about the legal status (e.g. jurisdiction, ownership, marriage status, authorship, licensing, physical properties) of something. Such assertions may be mandated at certain junctures or form components of other regulations.
- An action (transaction, data access, rule change…) proposed by a party is legal, meaning that it meets the current regulations about the class of thing it is.

- The same is legal in the present system state (for example is not performing a data action that the access control defined by previous actions forbids, or spends an asset that has already been spent.
- Some fact has been asserted by a sufficient collection of agents (where "sufficient" may mean many different things depending on context).
- Some fact about the overall state is true. This may be encapsulated in the state of a smart contract that monitors state.
- Some procedure (i.e. linked series of things) has completed properly.
- A regulation about the nature of software (as a product, as a smart contract, as a new regulation or anything else). This may be a specification coupled with a claim or guarantee that the software meets it. That may include testing, verification or a procedure.

The combination of TEE execution for trust and privacy, and blockchain or similar for integrity, immutability and trust should provide solutions to many of the security related problems.

I hope that this vision makes regulation seem worthwhile, an amazing technological challenge, and appealing. A *sexy* thing to work on indeed.

Next steps

A number of exercises must be initiated.

A. The development of blockchain(s) in which regulation is integral: direct support for layered regulation and rules, with rule satisfaction and maintenance of regulatory state built into the consistency relations for new blocks.
B. Serious research into the structure of regulation in a general high level sense: what logics are to be built in; how do we mix automated and human-based judgements and processes?
C. Research into the nature of jurisdiction and the relationship to blockchain structure.
D. Research into the relation between TEEs and blockchain structure, including blockchains and operations including smart contracts over private data. For example executing smart contracts on TEEs should solve the privacy problem that limits their use. For similar reasons they will enable regulations involving private data to be checked. By reducing or eliminating the need for replication to ensure correctness and authenticity, they have the prospect of improving the efficiency of all sorts of calculation.
E. We need to engage with several joint user/regulatory communities (e.g. bankers and banking regulators) to discover how well they will fit into our framework, and develop prototypes.
F. Creating regulatory languages in support of the above, hopefully ones that can be brought within a single overall notation with dialects.

Privacy versus accountability

Those posting material into digital civilisation must be accountable for what they put there, whether it is a financial transaction, medical diagnosis, news posting or judicial/electoral decision. In some of these and other cases we would expect everyone to be able to identify them from the posting. In others the agent A doing the posting might have the wish to remain anonymous. So someone's bank transaction might be private to them unless it was illegal; a posting on something like Facebook might be anonymous unless proved to be malicious or "fake news". A vote should remain anonymous unless it is proved to be fraudulent, and so on.

The regulations surrounding each domain of digital civilisation should make it clear just what the rights to this accountable anonymity are, what transgressions can lead to exactly what lowering of anonymity, and to whom they should be reported. Depending on circumstances this might be: the regulatory authority, another interested party, the legal authorities, or everyone.

In some cases a breach of regulations warranting loss of anonymity might well be immediately apparent, so that the offending transaction or information never gets posted. In others the offence will only become apparent later.

It is clearly vital that the mechanisms for both protecting anonymity and providing accountability are extremely robust, completely objective, and absolutely trustworthy.

Quis custodiet custodies

That is a well-known latin expression meaning *who guards the guards.* It is not an unreasonable proposition that the regulator R of domain X of society for jurisdiction J should have an important role in the governance of a coalition regulation chain C implementing its dictats, subject to a suitable degree of scrutiny of the rules it lays down and their implementation. But how do the general population trust that the limitations on its powers are being respected. For example how do we know it is not abusing its permission to break anonymity when it has no justification to do so?

The mechanism for the last of these would depend on how the anonymity was implemented (e.g. through TEE, MPC or both) but might involve gaining approval for an illegitimate rule. It would be much harder — hopefully effectively impossible — when the operation of the chain C is guaranteed to be monitored by independent parties. There may well be a strong argument for making rule changes subject to approval by these independents and delegating anonymity breaking to them.

This could be achieved by ensuring suitable access, duties and independence of the blockchain(s) higher up a hierarchy of blockchains. The ultimate guarantor of trustworthiness and correctness will be a well structured international public blockchain, but to be effective the relationships between various levels will need care and research.

Chapter 3: Trust through technology

Introduction

Our vision is of a world operating on a technological infrastructure which is itself reliable and trustworthy, and which helps to overcome malicious behaviour and subversion which may be present among the people and states using it. Our vision requires international cooperation, many software and hardware based solutions to how to implement it effectively and efficiently, and a fair and open playing field for those involved in building it, supplying it and using it. We want it to be inclusive, both because this is fair to all and because making it exclusive will drive potential participants away.

In my previous papers I have written about the overall logical structure of such a system, protecting integrity and resolving arguments through blockchain structures, and about how legal and regulatory frameworks can both be maintained and implemented naturally within it. In the present paper I look at how we can ensure that digital civilisation cannot be undermined by incorrect and untrustworthy hardware and software and how it can itself support trust in traded IT products.. Equally, I want to stop belligerent politicians undermining trust in it by inventing reasons why the constituent hardware should be condemned.

I first consider the question of how trust in technology can be separated from political prejudice, using pharmaceuticals as an interesting counterpoint, before analysing roles for things like verification, diversity and testing in the creation of reliable and trustworthy components for digital civilisation.

Trust in technology: where did it all go wrong?

As I write this (July 2020), trust in technology, particularly information technology, has become a huge political issue. The higher tech something is, and the closer it is to the core of civilisation, the more of a political football it becomes. No nation should risk its core infrastructure, but equally, nor should it conjecture a "risk", that it either knows or could easily know is fallacious, to justify a policy. The fact that such technology is hard to understand for the layperson makes it easy for politicians and others with agendas to create suspicions. Such agendas typically arise in the context of political, geopolitical and economic rivalries. I do not attempt to judge such disputes, merely argue that it would be far better if IT products were not so easy to condemn without technically legitimate reason.

I think that companies developing IT products and other parts of the chain of supply and demand for such products have made the mistake of not paying sufficient attention to making them transparently trustworthy. We should not have to have absolute trust in those supplying us in order to trust what we buy from them. IT should not be a black art, safe

only when practiced by one's own witch. It should be the job of those selling IT products to supply evidence of quality and trustworthiness that others can follow and verify, and it simply should not make sense to accuse the products of being malevolent.

There will be higher expectations of suppliers to the core that makes civilisation tick and keeps it secure, than suppliers of more peripheral IT. This is just as true in medical devices, where there are far higher reliability expectations of life-critical devices (e.g. pacemakers that make hearts tick) and implants than for massage machines and dental floss.

Two other potential arguments for trust are both shown to be imperfect by Boeing's 737 Max failings: regulation, or at least regulation as now practiced, and the argument that a huge company has too much to lose to deliver an imperfect product. The imagined possibility of state pressure on a company, large or small, further negates this "too big to be dishonest" argument.

A piece of hardware might be untrustworthy because it has been designed or built by malevolent or untrustworthy people, or because despite people involved having no malicious intent, the design, procurement or manufacturing process is faulty, or because it is used wrongly.

Fundamentally, I believe that hardware and software has the potential to be completely trustworthy, in a way that a human cannot. One can know exactly how hardware or software will behave, in a way that is impossible with a person. Thus one can trust IT equipment to make unbiassed choices, if it is suitably designed and built, in a way that is not really possible with humans, even with elaborate conflict of interest rules etc. Even the most trustworthy person may be under duress.

We are thus in the paradoxical situation that while it is possible to build technology that can be completely trusted, circumstances have conspired to make trust in technology a major international issue.

It is not our business to stop a US or Chinese president telling their nation not to buy products from the other because (he says) that country is reprehensible in some way, or indeed because they do not want to see their rival attain a monopolistic position in a core technology. Our business is to move the way some sorts of technology are created so that it would ridiculous for these people to say "don't buy X because it might have vulnerabilities that would open us to losing confidential information, or sabotage". In other words, we want politicians to be honest about their motives, rather than blame us, the technologists!

It is accepted economic fact that free trade is a *good thing* and aids prosperity and development. The creation of trade barriers, whether via punitive tariffs, nationalist procurement policies, or by banning perfectly good products for spurious reasons, must therefore be a bad thing, economically. We can now see that the UK is forecasting a significant delay in its adoption of 5G technology through its recent ban on Huawei.

At the same time governments certainly do engage in espionage, cyber warfare and sabotage using whatever means they think they can get away with (or at least tough out if

they get caught, a speciality of Russia right now, it seems). As I write this, Russian espionage against Covid 19 research and attacks on western elections is in the headlines: https://www.thetimes.co.uk/article/russian-actors-tried-to-influence-2019-general-election-says-dominic-raab-z57j6s825?shareToken=0e8dc83a3e7859f0b601fe71e8cd166d And the UK government is being criticised by its own oversight committee for not paying enough attention to this problem: https://www.nytimes.com/2020/07/21/world/europe/uk-russia-report-brexit-interference.html

The Russian state's modes of action have allegedly included social media and "hack and release". Certainly the current US administration could be accused of talking down the Russian threat because it has allegedly been a beneficiary, as possibly have the Brexiteers running the UK government.

It is observed fact that any nation state engaging in espionage or other covert activity will never admit it except possibly (e.g. French sabotage on Rainbow Warrior in New Zealand) after the passage of much time and invariably when it has either long been blatantly obvious or has tuned into a historical good-news story like Bletchley Park. *It is a corollary to this that it is hard to take denials of covert activity seriously, even though they may be sometimes be true.*

The Trump administration is very likely overdoing the perceived threat from Huawei for economic and political reasons and is dragging its allies along. Thus in both this and the Russian electoral interference case it is reasonable to state that the possibility of interference by a foreign state has become a political tool in the hands of real or imagined victims.

Computer scientists have been arguing for the past 40 or 50 years that hardware and software should be created to be correct by design. Yet this objective has, with rather few exceptions, been lost to commercial expediency and the technical difficulty of achieving it. It must now be in the commercial interest of those selling technology, and especially critical technology to be able to prove what it does, and that it cannot be perverted. Equally importantly, society must understand (in both a technical and non-technical sense) and accept such proofs. It accepts assays on drugs, vaccines, chemicals etc: extending this to computer hardware requires a technical and seemingly also a cultural shift.

Digital civilisation seeks to build an environment that all users can trust. This trust can come from the nature of the algorithms such as blockchains that are used, but must also come from the hardware and low level software this is run on top of. In some cases, such as protecting private information, there is inevitably vulnerability to a single point of failure. And though it is possible to trust a blockchain with some corrupt participants because of assumptions about overall trustworthiness, this becomes so much harder in the almost certain scenario that the members of the chain are using only a very small range of types of hardware and software. I recall that when we visited a well known Chinese company to look at their blockchain ideas, **they** were producing the software, to run on **their** hardware.

I am quite sure that the same competitive forces are having the same effects on large tech companies everywhere, given the advertising one sees for things like the "IBM Blockchain".

We conclude that digital civilisation is intimately connected with, let us call it, **assayable** or **auditable** hardware and software. By *assayable* I mean something different from verifiable, though the two are closely linked, and verification might be a large part of the assay/audit process. We need to shift the culture of IT to that in many other areas of commerce and human discourse, where checking of the honesty of the product is routinely expected by both producer and consumer, and the ultimate consumer is confident that the overall "ecosystem" does indeed carry out such checks.

Let us study the medicinal example.

Comparison with pharmaceutical products

The pharmaceutical industry is in some ways very similar to the IT industry: it operates at the forefront of science and technology; it is similarly a mixture of multi-nationals and innovative start-ups; it is extremely competitive; IP is crucial; and as we currently see with Covid, national pride and economics are at stake. Ethics has always been deeply embedded there in a way that has only emerged relatively recently in IT. And yet the way trust is managed is different, perhaps more mature.

If I buy a medicine (particularly in the role of wholesaler or pharmacist), or prescribe it as a medical professional, I want to know

1. Exactly what it is: chemical structure in the case of most drugs.
2. That the product I think I am buying will do no harm, except perhaps for a list of side effects and "contra-indications". This will require clinical trials and monitoring.
3. That it will do good in a specific set of circumstances, usually for treating a disease. Both this and 2 require clinical trials, for which there are very definite ethical standards.
4. That the product I am accessing is what I think it is. In unregulated markets for drugs, whether legal or not, often either the product is fake or it is mixed with something nasty. This sometimes happens in regulated markets too: https://nltimes.nl/2020/07/10/carcinogenic-found-paracetamol-possibly-available-dutch-market-report .)

The word "assay" refers mainly to the last of these, and is measured against the first together with analysis of what else is there. The second and third speak to the utility of the product as specified.

I note that pharmaceutical products can be sold in multiple forms: licences to the design, licences to design plus manufacturing process, raw product and packaged product. Of course the same also applies to chemical and biological precursors.

Somewhere in this framework fit warranties and product liability, and we would expect the way in which the above operate to be closely regulated, as part of which suppliers and intermediaries will often require licences.[i]

Let's consider the above in the context of IT.

1. This should be a complete specification of functionality or a complete circuit or source code, and preferably both. The circuit/source is the better analogy for chemical structure. They can of course by protected by patent, NDA, copyright, etc. In the case of hardware there should also be the standard of manufacture.
2. Analogues here would be: security threat models that the product is proof against: for example should I trust it if the attacker has control of its physical, electrical, wifi environments, are there limits on the criticality or secrecy of their use? What is the nature of the threat?[ii] For example, what if the enemy can read the inputs and outputs of a chip, or see its memory; what if the enemy can monitor its power or memory usage, or detailed timing in pursuit of side channels. How easy is chip to disrupt? Can it be used in space? What are its power consumption and use of other resources?
3. That it will do good should be a corollary to the functional specification, or perhaps this is the place for functional verification.
4. That I am buying what I think I am should be straightforward in the case of software that can be inspected in source form and compiled, particularly if I inspect it and install it myself rather than letting you install it on thousands of pieces of equipment. The difficulty really comes with electronic equipment. We need an analogue of assay that can exclude malicious intent. Note that in medicine one is not required to examine every pill, but in the case of critical hardware this might be necessary. The reward (to the bad player) and damage models are different.

One does not usually suspect a pharma manufacturer of malicious intent, at least by direct action of its products.

The regulation and relative transparency of the pharmaceutical market are consequences of its close connection to the universal basic principles of medical ethics and the fact that IP is there typically protected by patent rather than obscurity. I certainly would not, however, hold up the pharmaceutical market as an ideal model of regulation, given how restrictive markets lead to ludicrous price variations between countries for the same drugs.

It is of course true that both software and hardware can be supplied in multiple forms. We will review some of these below, and show how the combined integrity and privacy preserving calculation of digital civilisation can contribute enormously to generating trust in these products without completely transforming the way that IP and disclosure work in our industry.

Auditable software

Assay, as described above, when done relative to a guarantee or bill of sale, is a form of *audit.* I think that that term fits best with software and hardware.

For both hardware and software we want to demonstrate that an artefact was created to meet the need it being sold for and that if it has additional capabilities these are explained. Software is usually complex and built from various libraries. In terms of explainability, it is clear that source code is much better than object code, and code that has been generated, preferably in automatic fashion, from a clear specification such as an object oriented description.

Then such models can be subjected to vulnerability analysis.

Almost certainly we should expect some sort of coverage analysis of software to avoid coded-in backdoors.

The libraries used by the software should be given similar care.

All of this may sound a little ambitious; but why? If we are giving software a crucial role in running some infrastructure, it is reasonable that we are given all the evidence we need to decide if we trust it to perform its function and to be secure. It would be sensible for regulators to define exactly what is required. Why should safety critical software or hardware, or national infrastructure on which the safety and security of a country depends, be subject to less scrutiny than a pharmaceutical product? Why should less ethics apply? I have no idea: pharma may simply be a more mature and better regulated market.

Historically, software vendors are extremely reluctant to let others see their source code, in order to preserve their IP and proprietary know how. And yet we think that source level inspection is a vital part of software assay. Potentially a blockchain mediated trust environment can provide a way of storing sources, testing and verification data in such a way that supplied products can be demonstrated trustworthy without providing an opportunity for IP theft.

Auditable hardware

In the modern world, hardware is almost always software with an infrastructure to run it on. The structure of the software may be directly reflected in the physical structure of the hardware, with direct mappings between logical variables and operations, and physical registers and calculations implemented as gates, or the software may be realised as microcoded instructions running on a specialised processor.

Insofar as hardware is thus “frozen software”[1] it can substantially be audited as the software that generated it, together with appropriate description of how the freezing is

[1] This analogy is inspired by Hegel, who once said that architecture is frozen music: that is a poetic truth, ours is close to literal truth.

implemented and the net-list or other description that it generates in such a way that the tie-in with the software is clear and reproducible.

Parts of the hardware that lie outside such process, and non-standard components included in it, may well require separate justification. Of course we can use the same auditing philosophy as above.

This sets out the overall approach we might follow, but the details and standards will need establishing, depending on the sort of hardware. If there are aspects of the hardware that are analogue or otherwise non-digital such as sensors, antennae, motors and lasers these may well require quite separate quality and inspection regimes that may operate by delivering certificates to the digital processes.

Unlike software, which can relatively easily be confirmed to be as designed, hardware presents us with the considerable challenge that what is delivered actually corresponds to *the design, the whole design, and nothing but the design*, paraphrasing a well-known oath.

If a receive a shipping container full of bananas, how do I know there is no poisonous spider hiding in it?

I have no magic solution to this problem, other than to observe that designs and manufacturing processes for IT hardware can, like pharmaceutical analogues, be licensed. Unpackaged chips are clearly more scannable than packaged ones.

It is worth remarking that vulnerabilities which stop a system from working properly can be reduced by replication, and that this is closely related to the principles underpinning blockchain. Such replication is unlikely to help against a faulty component design as the fault itself is likely replicated. However it is likely to help if subterfuge has created a vulnerability in only a few chips in the hope that these will evade sampling-based analysis.

On the other hand, this argument does not work for confidentiality flaws, unless the confidential information is spread via multi-party computation. Depending on the assumed threat model, this may indicate identifying a subset of privacy-critical functions such as signature and other functions involving keys, and distribute them using a secret-sharing scheme amongst multiple TEEs.

Trust through diversity

In a blockchain environment it is usual to insist that over half the power (defined in an appropriate way) follows the rules accurately and properly. This clearly imagines that the bad few operate that way because they choose to try to subvert the chain. There is a generally unwritten assumption that when an agent chooses to behave well it has a perfect program to implement the protocol running on reliable hardware.

Similar assumptions are made in other decentralised algorithms and architectures with Byzantine fault models.

Trust through verification

Verification of software, and of hardware designs, is obviously of great importance in building trustworthy IT. This is true whether the trust we are looking for is that a system will follow a particular protocol accurately, or that the protocol itself is fair in some defined sense, or that it can be trusted not to communicate its keys or unencrypted data, or that it cannot be hijacked or turned off, or that it has some other abstract property such as deadlock freedom or determinism.

This illustrates the range of properties that can be verified. These vary in the level of abstraction of the property proved, and also the completeness. They also usually, explicitly or implicitly, make assumptions about how their environment behaves and the format of interactions. If a verification of security makes such assumptions, those relying on it must reassure themselves that they are adhered to, either because of logical or physical limitations. The latter, at last, will be easier to ensure if hardware is under one's own control rather than someone else's.

I believe that when specifically looking for trust from verification, there is an even stronger reason than usual for using a method which in effect compiles the implementation from a clear and trustworthy specification, with verification of the latter. Of course the code translation/compilation technology itself must be known to be trustworthy.

If one has proved that a piece of software or hardware S satisfies property R provided its environment satisfies assumptions A, and the implementation mechanisms of S (compilation, fabrication etc) are faithful, it will mean we can trust it will do this. However it is vitally important to realise that this does not imply that we can trust S in any other sense.

Testing for Trust

Testing is an important part of any software engineering process, even where there there is a lot of verification. It provides reassurance in that case and the primary demonstration that the product conforms to expectations when there is no verification.

It makes huge sense to do this when we are judging the public personality of software or hardware. If we are worried that there may be covert aspects we are not seeing then there is a strong possibility that testing will not find it. We really need to employ testing coupled with sufficient knowledge of the logic of the design to ensure that it has all been well tested. Presenting evidence of such testing is best done with evidence that the test cases were chosen independently of the vendor. Thus evidence of testing is again something that fits perfectly into the trust model of digital civilisation, but needs to be carefully designed relative to the logic embedded in the product, which must be sufficiently accessible to the testers.

Conclusions

It is really disappointing that the way politics, international relations and our industry works have created an environment where trade in core IT equipment is difficult and controversial. In this paper I have shown how more careful behaviour by manufacturers and use of the concepts of Digital Civilisation may be able to provide an environment where trust comes from devices themselves.

Acknowledgements

We have a large team at TBTL and OxHainan and many of these have contributed to my understanding of what can be achieved via blockchain and TEE. These include Bangdao Chen, Ivan Martinovic, Feihu Song, Srdjan Capkun, Pedro Antonino, Liu Han and Rafi Yahalom. It was our sponsor Yang Chunzhi of NanHai (South Sea) Cloud, Chengmai, Hainan (SSC) who identified the need for a blockchain based digital civilisation project, and generously sponsored it.

The first chapter of this paper was written in Hainan in 2019. The others were written in Europe while unable to go there because of the Covid 19 pandemic during March-August 2020.

I have had interesting conversations with many people who have read drafts, and the paper has benefitted from these.