Lovett, Matthew ORCID: 0000-0003-3599-7886 and Thomas, Lee (2021) A fork in the road: Perspectives on sustainability and decentralised governance in digital institutions. First Monday, 26 (11). doi:10.5210/fm.v26i11.12357

Official URL: http://dx.doi.org/10.5210/fm.v26i11.12357
DOI: http://dx.doi.org/10.5210/fm.v26i11.12357
EPrint URI: https://eprints.glos.ac.uk/id/eprint/10378

A fork in the road: Perspectives on sustainability and decentralised governance in digital institutions

Abstract

A digital institution is a set of computer-based rules that perform intermediating roles upon which one or more person's well-being depends. This article argues that governance, the processes and customs by which rules are agreed, is critical to the sustainability of the digital institution and therefore of society more broadly. The objective of this work was to interrogate whether emerging decentralised architectures (blockchain) can offer new perspectives on digital sustainability in the form of decentralised governance. Firstly, the literature on decentralised modes of governance was synthesised. Then, existing digital institutions were reviewed, categorised and mapped onto a multi-domain layered conceptual framework that draws out three distinct modes for enactment of changes to digital institution rules; direct, integrated, and fork-based. We concluded that the coupling of decentralised governance approaches with fork-based or integrated enactment stands to enhance digital sustainability through increased perception of trustworthiness afforded through independently verifiable and cryptographically secure audit trails.

Introduction

In 2015, the United Nations' introduced seventeen Sustainable Development Goals (SDGs). The sixteenth SDG promotes 'peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels' (United Nations Department of Global Affairs, 2015). This article seeks to explore the potential confluences between this framing of public institutions as enablers of global sustainability, and the concept of digital institutions; asking what emergent forms of decentralised governance facilitated by blockchain technology might offer to current debates surrounding governance.

Blockchain technology, in its many forms, continues to evolve and offer increasingly varied approaches to implementing governance structures within decentralised networks. We use the term blockchain or blockchain technology to encompass distributed transaction-oriented computer systems where there is a need to achieve consensus on the state of a shared digital object. Well known examples include cryptocurrencies and smart contract platforms. In this article, our focus has been to interrogate whether blockchain can offer new perspectives on digital sustainability in the form of decentralised governance, and if so, what these perspectives might be.

1. Digital Institutions and Sustainability

According to the SDG matrix, the sustainability of a society is influenced by the effective governance and accountability – and therefore trustworthiness – of its institutions. The Open Data Institute (ODI) describes data institutions as 'organisations whose purpose involves stewarding data on behalf of others, often towards public, educational or charitable aims' (ODI, 2020). Expanding on this, we define a digital institution as a set of computer-based rules that perform intermediating roles upon which one or more person's well-being depends. For example, the intermediating rules might constitute a market, a voting system or a payment system. The intermediating rules can be viewed as transactive, i.e., as a series of discrete adjustments (transactions) to a shared digital object. Typically, in blockchain systems, the rules are instantiated across a network of computers with different owners. In the context of our overarching enquiry, for a digital institution to be sustainable, we understood that the operation of its rules must be perceived as *trustworthy* by its users. The question of how users are able to verify the correct operation of intermediating rules is critical to this trust. Further, we ask how the perception of trustworthiness might be influenced by the digital institution's mode of governance.

In 'Blockchains, trust and action nets: extending the pathologies of financial globalization', Campbell-Verduyn and Goguen frame trust 'as the confidence that humans build among themselves through collective efforts' (Campbell-Verduyn and Goguen, 2019), suggesting that trust persists as a core component of 'effective, accountable and inclusive institutions', be they off- or on-chain. In other words, trust is a feature of traditional institutions alluded to in the sixteenth SDG, along with the digital institutions that are the subject of this article. Within our current definition of digital institutions, it is therefore notable that Campbell-Verduyn and Goguen propose that 'applications of blockchains reorganize the socio-technical underpinnings of trust in manners that perpetuate the dominance of both novel and old centralized third-party actors' (Campbell-Verduyn and Goguen, 2019). This further emphasises the centrality of trust for successful governance within the blockchain paradigm, albeit in an evolving form. So whilst trust may well remain a key enabler of a sustainable society according to UN precepts, as discussed below, Werbach, along with De Filippi et al., expand on how conceptions of trust in the context of blockchain networks, are malleable and open to debate. For example, in *The blockchain and the new architecture of trust*, Kevin Werbach identifies different ontologies of trust:

1) Peer-to-peer (P2P) trust. A trust that 'is based on relationships and ethical norms: I trust you because I trust *you*'.

2) Leviathan trust. Based on Hobbes' concept of the state as Leviathan, the state 'can enforce private contracts and property rights […] and largely operates in the background to prevent others from imposing their will through force or trickery'.

3) Trust in intermediaries. 'In this arrangement, the local rules and the reputation of the intermediaries take the place of social norms and government-issued laws to structure transactions' (Werbach, 2018: 27).

Building on this foundation, Werbach suggests that blockchain technology has ushered in a new form of 'trustless trust': 'on a blockchain network, nothing is assumed to be trustworthy

… except the output of the network itself. This distinctive arrangement defines the landscape for the blockchain's interactions with law, regulation and governance' (Werbach, 2018: 29). In 'Blockchain as a confidence machine: the problem of trust & challenges of governance', De Filippi et al. extend this analysis of trust, and propose that, rather than seeing 'trust' as the central component of the blockchain governance paradigm, we should instead think in terms of 'confidence'. As with Werbach's affirmation that the only thing that is trustworthy about a blockchain network is its output, De Filippi et al. propose that 'in comparison to trust, it is arguably easier to build confidence, because it does not require any communication or mutual commitment by at least two actors, but rather emerges through the cognitive process of one single agent' (De Filippi et al., 2020).

This leads De Filippi et al. to suggest that blockchain networks are not 'trustless' or 'trust-free' systems. They function via user confidence, rather than trust *as such*. Their contention is that 'confidence essentially operates as a platform for trust', since, on their reading, confidence operates within the domain of one individual's relationship to a given situation, whereas trust is a property of a relationship between two or more parties. Thus, for De Filippi et al., blockchain systems operate 'by maximizing the degree of *confidence* in the system as a means to indirectly reduce the need for *trust*' (De Filippi et al., 2020). While these may – to an extent – be semantic positions, this analysis does suggest that a new conception of trust has co-emerged with blockchain. Whether we wish to see it as 'confidence-trust' or 'trustless-trust', the point is that digital institutions, in the form of blockchain networks, operate by foregrounding users' trust in a system based on a particular technological paradigm, and de-emphasise a range of traditional institutional behaviours that would otherwise be seen as contributing to successful and sustainable governance.

The 2018 World Economic and Social Survey recognises the role that decentralised networks, blockchain and cryptocurrencies have to play in upholding 'ethical standards and effective and accountable institutions', specifically in relation to the capacity of decentralised technology systems to log digital transactions (UN, 2018: 25). It also recognises the challenge that decentralised software systems bring to regulatory systems and traditional models of the 'formal institution' (UN, 2018: 37), particularly in the way that familiar notions of trust in such systems and institutions are being transformed. In this regard, the survey notes the importance of fairness, transparency and accountability; proposing that transparency in digital systems is central to sustainable and ethical development (UN, 2018: 64).

In defining sustainability, the ODI addresses an institution's capacity to 'continue [or sustain its] operations so that it can deliver on its purpose', specifying financial sustainability as 'the ability to maintain financial capacity over time' (ODI, 2020). Although the ODI does recognise that 'the sustainability of a data institution is linked to the ecosystem that surrounds it and the community that supports it' (ODI, 2020), its definition essentially frames sustainability as a matter of the institution's own survival, rather than stewardship or broader issues of governance. Indeed, the ODI suggests that, in terms of measuring the sustainability of data institutions, it is partially a matter of implementing 'sufficient governance to ensure that the aims of the organisation can be delivered in the long term' (ODI, 2020). As we shall see in the discussion below, there are also differences of opinion over the meaning of governance within the blockchain community. In general, this divergence reflects two views: whether governance refers to a network's ability to maintain self-identity and coherence (defined here by the ODI as 'sustainability'), or if it denotes the behaviours of a more distributed network of stakeholders, both network participants, along with a variety of non-participants who are nonetheless affected by the network's activity.

We see that a key quality of decentralised governance is the capacity of new forms of trust to integrate traditionally disparate and informal relations within a self-sustaining and coherent network. What emerges is a holistic view of institutional governance that encompasses the tightly delineated rulesets of decentralised transacting, along with established models of what the economist Elinor Ostrom referred to as 'polycentric' co-governance (Toonen, 2010). Here, governance becomes a matter of achieving system coherence given the challenges of both recognising and meaningfully integrating the dynamic relations between a network's users, its internal rulesets, its wider set of stakeholders and the range of contextual forces that constitute its external environment.

2. Polycentric Institutions and Sustainability

What follows is a brief literature review centring on polycentric governance. This lays the groundwork for our later discussion of contemporary debates relating to governance in the context of decentralised technology.

In her work on the management of Common Pool Resources (e.g. water, fisheries), Elinor Ostrom used the concept of polycentric institutions to show how they could be managed by a combination of governmental and non-governmental actors [3]. Her research examined matters of governance in what she referred to as the public service industry (PSI), which, according to Ostrom, was a 'market-like network of users, providers, and producers of public goods and services'. In his analysis of Ostrom's work, Toonen describes this as an 'interdependent' alliance between government and non-government actors such as 'citizens, neighbourhoods [and] societal organisations', collaborating to achieve what he calls 'co-governance' (Toonen, 2010: 196). In 'Resilience in Public Administration: The Work of Elinor and Vincent Ostrom from a Public Administration Perspective', subtitled 'A Conversation with Elinor and Vincent Ostrom', Elinor Ostrom informs Toonen that polycentric institutions (PIs) facilitate 'trustworthiness, [and] levels of cooperation of participants', which lead to 'more effective, equitable, and sustainable outcomes at multiple scales' (E. Ostrom in Toonen, 2010: 664). Furthermore, Vincent Ostrom (E. Ostrom's co-researcher and husband), explains that PIs are characterised by 'multiplicity, diversity, interdependency, checks and balances [and] complexity' (V. Ostrom, in Toonen, 2010: 194).

It is also worth noting that the Ostroms' work prefigures some of the issues that we shall consider later in the article in terms of decentralised software architectures. Elinor Ostrom reasoned that polycentric, but not-yet decentralised, institutions 'generally outperformed' systems with only one or two central departments, and she argued against the need for continued central coordination within institutions (E. Ostrom, in Toonen, 2010: 197). For our current purposes, what is interesting about the foundations for Elinor Ostrom's thinking in this area is that she understood that governance of CPRs transected accepted governance frameworks relating to both the production and exchange of private goods - which would be seen as the domain of the market - along with the regulation of non-private goods, as the domain of a government (Ostrom, 2010: 642).

In the context of digital institutions, Ostrom's contention that effective institutions are the amalgamation of an heterogeneous set of actors have an important part to play in establishing a framework for digital sustainability. For example, De Filippi et al. provide the following polycentric mapping of a blockchain-based network's key stakeholders:

- a few economic players, such as the largest mining pools, mining farms, popular online exchanges and blockchain explorers

- core developers and open source contributors
- cryptocurrency, token holders, as well as more general users
- regulators (De Filippi et al., 2020)

Whilst they do not define this as a governance framework as such, De Filippi et al. do describe blockchain-based systems as ‘socio-technical systems’, understanding that they are governed by ‘a set of rules and procedures, [which are] defined by the underlying blockchain protocol, and are enforced by a distributed network of miners and validators maintaining the system’ (De Filippi et al., 2020).

Ostrom also proposed the development of ‘autonomous institutions within communities’ (Choe and Yun, 2017: 122), which would be sympathetically matched to the CPRs they were designed to manage. Trust was critical to the development of these institutional frameworks and rule systems, and for Ostrom, the successful building of trust - particularly in regard to the challenge of finding consensual ‘ways of monitoring one another’ - was repeatedly manifest in polycentric institutions that had achieved successful CPR governance (Ostrom, 2010: 664). Whilst Ostrom’s work predates our contemporary digitally networked environment, public-private utilities such as bandwidth, energy resources, and even networks themselves can be thought of as CPRs within a polycentric governance context.

Hansen et al. and Ngeta and Waage expand on the polycentric governance approach in a way that is applicable to collective control of digital intermediaries. Hansen et al. propose that governance ‘refers to ordered rules and collective action in society, where a system of rules around decision-making is implemented by social actors in a coordinated way’ (Hansen et al., 2014: 44). Similarly, Waage et al. suggest that governance ‘refers to the institutions, mechanisms, or processes backed by political power and/or authority that allow an activity or set of activities to be controlled, influenced, or directed in the collective interest’ (Waage et al., 2015: 84). In this account, the defining characteristics of governance include users’ accountability and responsibility to the collective interest, listing ‘hard and soft laws, regulations, agreements, institutions (national, local, and regional government; international bodies; secretariats; civil society; and the private sector), shared norms of behaviour, and the balance of power therein’ as common polycentric institutional features (Waage et al., 2015: 84). They also claim that if sustainable development is to be achieved, then ‘responsibility, transparency, accountability, capacity, and legitimacy' must be implemented at all levels of regional, national and international governance (Waage et al. 2015: 80).

Ngeta's examination of the history of governance during the postwar period, and the emergence of its contemporary usage, shows how governments and non-governmental agencies have partnered to shape and deliver an 'agenda of economic liberalisation in developing countries' (Ngeta, 2014: 26), and reflects this transition to distributed, polycentric governance. He suggests 'governance must be structured around a system of rules that allow decision-making and implementation by social actors in a coordinated way' (Ngeta, 2014: 26). This foregrounding of rules and decision-making enables us to shift our focus towards matters of digital governance, particularly in the context of decentralised information technology.

3. Decentralised Technology and Governance

Whilst Ostrom's work has much to offer in terms of providing a foundation for the analysis of decentralised governance, problems emerge when we begin to engage more specifically with how blockchain networks and systems function. De Filippi et al. demonstrate that, ultimately, blockchain governance is not a matter of polycentric, collective consensus, instead, they see that 'a few influential actors ultimately have the power to affect the operations of the overall network—often in order to further their own interests' (De Filippi et al., 2020). Similarly, Campbell-Verduyn and Goguen also argue that mining pools operate in opposition to the collective-orientation of polycentric governance, stating that 'a handful of companies [pool] their resources to reach economies of scale in producing cryptocurrencies' (Campbell-Verduyn and Goguen, 2019). These perspectives illustrate the complexities involved in formulating an effective model for decentralised governance, and while there is much that we can take from Ostrom's work, there are a number of other factors to consider as well. For example, Zwitter and Hazenberg create a continuum of development that mirrors Ngeta's historical perspective in order to address governance in the decentralised digital domain. Although this continuum does not engage with the tacit structural imbalances highlighted by De Filippi et al. and Campbell-Verduyn and Goguen, Zwitter and Hazenberg put forward three 'modes' of governance as a means to articulate governance as an evolving phenomenon. Mode one operates within a 'hierarchical', 'command and control' paradigm, and ‘relies on authoritative institutions to make policies through the enforcement of hard law’ (Zwitter and Hazenberg, 2020: 4). Mode two consists of ‘horizontal’, public and private partnerships, where ‘authority is not necessarily acquired by identity but rather through performance, knowledge, and expertise’ (Zwitter and Hazenberg, 2020: 5). Given that both modes one and two fundamentally rely on trust, and indeed reflect Werbach’s three-tiered taxonomy that was discussed above, Zwitter and Hazenberg propose that trustless transacting by its very nature subverts these accepted forms of governance. As such, since they see blockchain technology facilitating the existence of ‘different roles and power relationships often residing in a single, anonymous, actor’ (Zwitter and Hazenberg, 2020: 6), their third mode of governance operates transversally; neither hierarchically nor horizontally, but as a property of the networked environment.

Such analysis emphasises how the evolution of trust in the context of anonymous decentralised transacting has impacted how governance is understood and implemented across networks. As has often been observed, the fact that Bitcoin and blockchain emerged in the wake of the 2008 global financial crash was ‘no coincidence’ (Musiani et al., 2018: 133). Its capacity to facilitate trustless transactions was intended to put Bitcoin beyond the control of ‘any central authority’ (Musiani et al., 2018: 133), demonstrating a fundamental distrust of conventional financial or governmental institutions. For Hütten and Thiemann, ‘the history of money is a history of trust, not between two individuals, but between an individual and a community as a whole’ (Hütten and Thiemann, 2018: 26). Moreover, money is ‘a social convention [with] social and political origins’, although its physical presence often obscures its more socially-embedded history (Hütten and Thiemann, 2018: 27). In this sense, Hütten and Thiemann considered Bitcoin, to be ‘a political experiment’ designed to bring about radically different approaches to trading and transacting, which gave rise to a new form of trust that evaded and superseded recognised forms of ‘sovereign power’ and ‘institutional underpinning’ (Hütten and Thiemann, 2018: 30).

Turning again to matters of blockchain governance, we see that this emphasis on trust has had significant impact on how governance is perceived, not only in familiar stakeholder contexts, but also now ‘on chain’, within decentralised networks. For example, ‘The Wood-Zamfir Governance Debates’, is an extensive analysis of Ethereum co-founder Gavin Wood, and Ethereum Foundation and CasperLabs researcher Vlad Zamfir’s debate concerning the nature and the implementation of governance within blockchain-based architectures (CleanApp, 2018). The article draws on a 2018 episode of the Zero Knowledge Podcast, where Zamfir and Wood offer differing theories of governance, based on their views about what constitutes a stakeholder within a blockchain ecosystem (Rose et al., 2018). For Wood, the matter is simple. Only token-holding users within a blockchain system should be counted as stakeholders. Zamfir's view, in contrast, is that 'anyone who is impacted by governance decisions is a stakeholder, [although] a lot of people who are affected by decisions [may be unrecognisable as stakeholders, and] don't know what their interests are', giving the US government as an example (Zamfir, 2018). For Wood, the issue of maintaining system cohesion is central to his vision of blockchain governance, a view that resonates with the ODI's perspectives on sustainability. Indeed, he asserts that 'the job of governance […] is to keep the whole *whole*; avoiding the pieces falling apart […] keeping things together is good for the thing and good for the stakeholders' (Wood, 2018). Whilst Wood therefore sees that 'governance is what changes a multiparty system into a moral person [and that it] is the glue that binds [people] into a single decision-making economic actor' (Wood, 2018), Zamfir's approach seems to follow Ostrom, understanding governance as the process of 'coordinating of shared resources' (Zamfir, 2018).

Clearly, Wood and Zamfir differed in their views about where blockchain governance happens, and as we move through the literature, a range of views emerge about how governance is being addressed in the context of decentralised software systems. In this sense, Hsieh et al. suggest that not only is governance an evolving concept, but it can hold different meanings for different stakeholders (as Zamfir and Wood's disagreement demonstrates):

'Blockchain governance is about determining who has authority (internal and external actors); how these actors are endowed (e.g., ownership rights vs. decision authority), in what form (formal and informal governance forms/structures), and at which level' (Hsieh et al., 2018: 49).

Hsieh et al. also recognise that governance of blockchain-based networks and organisations can happen in various ways; by decentralised communities, by centralized corporations, or jointly by both as hybrids (Hsieh et al., 2018: 50). Further to this, Zwitter and Hazenberg (2020) show how decentralised network governance can take three forms: (1) platform strategy, (2) private strategy and (3) legal strategy (Zwitter and Hazenberg, 2020: 9). In scenario (1), on- and off-chain stakeholders interact, and off-chain parties - with state support

- can influence on-chain activity. In scenario (2), negotiations and consensus activities occur within a network, and in scenario (3), off-chain regulation can be drawn on and implemented to resolve conflicts within the network (Zwitter and Hazenberg, 2020: 9). What is interesting about this modelling is the proposal that, whilst decentralised blockchain systems may not entirely change the underlying nature of governance - that is to say, decentralised governance is not necessarily a brand-new and unprecedented form of governance - these systems do allow for new forms of negotiation and interplay between different types of stakeholders to emerge. Zwitter and Hazenberg's research suggests that 'in a context where traditional governance fails, decentralised network governance opens up a space for contestation in which actors in concert govern each other' (Zwitter and Hazenberg, 2020: 9). As such, it is the specificity of how stakeholders are connected with each other - via delineated contracts within a decentralised network - that offers the means of balancing interactions between weaker and stronger stakeholders, and the three forms of governance articulates how this appears in practice. This suggests that blockchain governance operates by amalgamating different standards and forms of organisation, networking, integration & control. Further to this, Chalmers et al.'s analysis highlights how blockchain architectures embody particular sets of perspectives concerning the design, function and purpose of governance structures; described as the 'interplay between distinct external enablers in a given industry or social context' (Chalmers et al., 2019: 2).

Clearly, the task of defining blockchain governance is fraught with complexities and contradictions. However, perhaps Zwitter and Hazenberg's 'decentralised network governance' model best aligns the radical innovations of decentralised software - and all of the behaviour changes it has brought - with Ostrom's PSI management framework for CPRs, which at its core facilitates the creation of heterogeneous and multilateral institutions. For Zwitter and Hazenberg, their model enables them to articulate both 'changing and multiple roles of actors' within a decentralised network, and at the same time foreground the need 'to identify roles depending on network clusters and policy domains' (Zwitter and Hazenberg, 2020: 11). It would seem therefore that a workable conception of blockchain governance must reflect two trajectories that blockchain systems have catalysed: on the one hand, the mutability of the governance framework depending on context, and on the other an acknowledgement that the participants within the framework also have mutable and many-sided roles. In this sense, blockchain governance becomes a decentralised interplay between multiple stakeholders, not only between on- and off-chain parties whilst a network is live, but at the design stage, as the network is formulated. This leads us to ask how a digital institution might be broken down into component parts, with associated trusted third parties, whose trustworthiness can be considered separately.

4. Taxonomy of digital institutions

Trust in digital institutions

Within digital institutions, a number of separate trusted components can be discerned. Trust in these components spans social and technological spheres. In the social sphere, an individual participant implicitly or explicitly holds trust in other individuals and the set of rules and norms they follow. In the technological sphere, trust is held in hardware and code (software and legal codes). Typically for each component in the technological sphere there is a connected set of trusted parties in the social sphere. Some examples of trusted components are shown in Table 1.

Table 1 – Trusted components within digital institutions

Trusted components in digital institutions	Example trusted parties
Social/legal agreements	Legal professionals
Automatically enforced contracts (smart contracts)	Smart contract companies/ engineers
Consensus protocols (how a network of computers and their owners agree the state of a digital object)	Protocol designers Implementers Miners/stakers
Network protocols	Standards agencies Implementers
Computer software	Software companies Software engineers
Computer hardware (e.g. digital key signature devices)	Manufacturers Firmware developers
Cryptography (e.g. Public Key Cryptography, hash funct ions)	Standards agencies Algorithm developers

Concentric domains

Whilst De Filippi et al. and Campbell-Verduyn and Goguen are right to define blockchain systems as socio-technical networks, our research goes further to propose that a digital institution exists across three concentric domains; the computer domain, the interaction domain, and the sociocultural domain. The innermost, the computer domain, is entirely defined by a hardware and software configuration of computers. The second, the interaction domain, encapsulates the minds that have influence over the operation of the institution; users (the individuals that interact with the institution's information exchange rules) and the governors (those with the capacity to change the intermediating rules, such as the core developers and contributors referred to by De Filippi et al.). Finally, the sociocultural domain encapsulates the influence that wider society has on the governors and users through a third set of individuals labelled outside influencers. The outside influencers do not have direct influence on the functioning of the computer domain. However, they may have social or cultural influence over those who do (the governors). Note that the initial designers and builders of a digital institution are early governors in this model, as they have direct influence over the institution's operation. If the design or instantiation of a digital institution leaves vulnerability to activities such as corruption, hacking, sabotage, power outages and other non-standard interactions, it opens the possibility that a subset of the governors may be clandestine.

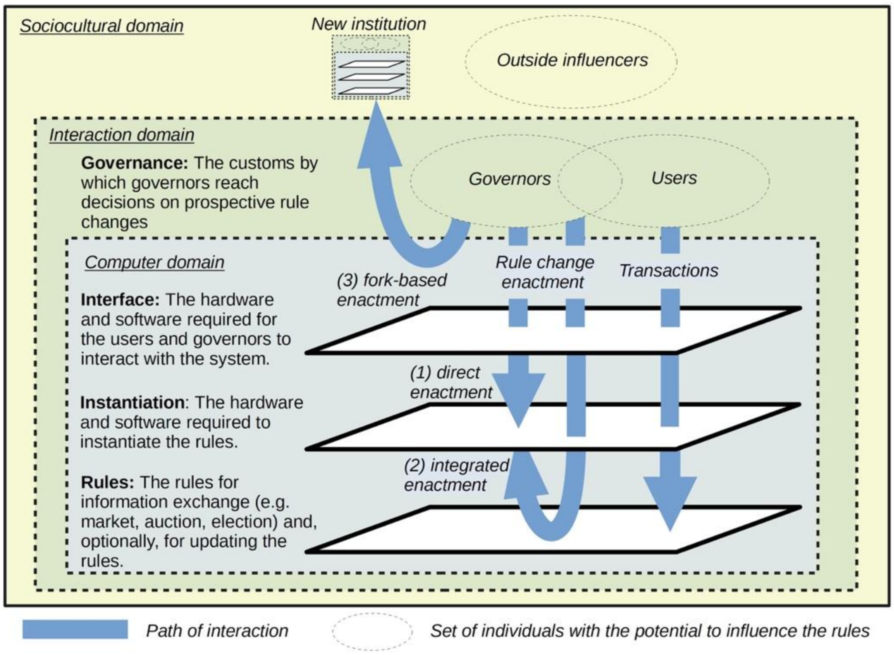

Figure 1: Layered model of a digital institution

Figure 1 presents a layered model of a digital institution. The model was created through iterative conjecture, criticism and alteration following Deutsch (Deutsch, 1997). The model consists of four layers. Three layers (the interface layer, the instantiation layer and the rules layer) fall into the computer domain and one, the governance layer, falls into the interaction domain. The interface layer is the software and hardware required by the users and the governors to interact with, or update, the institution's rules. The instantiation layer encompasses the hardware and software required for the system to operate. The rules layer defines the abstract rules for information exchange amongst participants (e.g. market, auction, election). It also may include rules by which the rules themselves may be updated. Finally, the governance layer encapsulates the customs by which the governors reach decisions on prospective rule changes. This includes traditional hierarchical management-based structures as well as approaches such as polycentric governance explored in the sections above.

When the governors have decided on a rule change, there are three ways to enact the change. The first is in making direct changes to the instantiation layer, labelled (1) in Figure 1. For example, traditional web-server based architectures are oriented around developer access to server computers. This creates the possibility for surreptitious, harmful rule changes to be made without the awareness of the users or other governors. In contrast, emerging architectures (oriented around digital signatures, cryptographic fingerprinting, and consensus algorithms) hold potential to enhance the trustworthiness (the transparency and independent verifiability of transactions) of the institution. The core innovation in these new architectures is that transactional interaction with an institution's rule set can be made to be independently verifiable and transparent. This results in the potential for a public audit trail for an institution and allows us to consider two emergent conceptions of rule change enactment in digital institutions; fork-based enactment and integrated enactment.

Fork-based enactment asserts that any rule changes result in the creation of a new institution (the old institution may persist at the option of its stakeholders) [1]. In this conception, governance is the process by which a decision is made to jump to a new institution (i.e., a new instantiation of a new set of rules derived from the old institution). One version of fork-based governance is that there are no formal governance rules at all. This is a characteristic of decentralised systems such as Bitcoin. In this case, any individual can verify where (in the audit trail) a jump to a new institution is made.

Whether or not a rule change results in what is popularly recognised as a new digital institution, however, depends in large part on those in control of the public brand and marketing for the digital institution in question. Where there is opposition to a rule change that is supported by the brand holders, as in the Ethereum "The DAO" fork, the old institution persists, but under different branding [4]. This may result in situations where the modified public-facing identity is inconsistent with the cryptographic audit trail. Nevertheless, as long as the audit trail is independently verifiable and the brand controllers act transparently, there is arguably no breach of trust in this situation.

The notion of integrated enactment follows Wood's commitment to *sustaining the whole*, as discussed above (Wood, 2018). In this conception, mechanisable governance procedures (e.g., votes) are integrated into the intermediating rules of the institution. The act of coming to a decision (governance) is merged with the enactment of it. This results in a situation where rule changes are undertaken transactionally with a resulting cryptographic audit trail. A potential benefit of this approach is that it improves the independence and quality of the audit trail for digital institution rule changes. It also stands to reduce the risk of a branding inconsistency as all governors and users a-priori consent to the institution's mechanised governance rules. Furthermore, it is conceivable that the governance rules could allow for the spawning of new institutions. Perhaps though, unless perfect rules are designed, desire may arise that the mechanised governance rules themselves should be changed. In this case the fallback is either direct, or fork-based rule change enactment.

Digital Institution Architectures

Our research identified seven types of digital institution architecture. They are characterised below and mapped onto the layered model, see Table 2.

1. Traditional web-server

Example: Airbnb, Amazon, eBay, Uber

This is the commonplace model of a web server controlled by an organisation. This type of digital institution typically consists of a management structure around a set of implementers (e.g., coders) all encapsulated within a legal entity such as a public company. Any changes in rules are decided by the management-implementer network. The decisions may be influenced, with discretion, by user feedback and external factors such as technical standards. In most cases, the exact intermediating rules are hidden from the user. Typically, a set of users unhappy with the rules would seek out a competitor organisation.

Public Blockchain

This type of digital institution is characterised by a transparent ruleset instantiated across a network of computers with different owners. They make use of consensus algorithms to maintain a common view of a shared data structure. Three subcategories were identified based on their rule change enactment mode. Public blockchains with fork-based enactment, with integrated enactment, and hybrid blockchain-webserver systems with significant direct enactment.

2. *Public blockchain, fork-based enactment*

Example: Bitcoin (Nakamoto, 2009), Ethereum (Buterin 2015; Wood 2014)

In this category the intermediating ruleset does not include explicit definition of how the ruleset might be changed. This is left to emerge from the natural interaction between the users, implementers and the operators (e.g., miners). Implementation of new intermediating rules results in the creation of a new parallel institution (that may retain the original branding).

3. *Public blockchain, integrated enactment*

Example: Polkadot (Wood, 2016), Tezos (Goodman, 2014)

In this category, the intermediating ruleset does include explicit definition of how the ruleset is updated. The governance decision making process is integrated with the enactment process. This might, for example, be the requirement for a particular multi-party signature to change the consensus algorithm.

4. *Hybrid blockchain web-server*

Example: Steemit (Kiayias et al., 2019)

This category encompasses blockchain systems that have a significant reliance on direct enactment, e.g., through reliance on a small number of website entry points or individuals, without which the system would become inaccessible to the public, become fragile, or fail.

5. *Coalition blockchain networks with hierarchical governance*

Example: Diem/Libra (Diem Association, 2020)

This category covers blockchain networks created in coalitions of multiple parties, where the system is instantiated using computers controlled by an exclusive group of parties. The system's governance is drawn from a committee comprising representatives from the parties and enactment is direct. It has defined interface standards but may include intermediating rulesets that are not transparent, or that are open to discretionary editing by the coalition.

6. Federated server templates

Example: Fediverse (The Federation 2020)

This category encompasses digital institutions that adhere to a set of server standards and intercommunication protocols such that they can be considered to be "federated" instances of a central standard. Governance is undertaken via changes to the standard, typically a committee of interested people. For example, servers that adhere to one or more of a set of public social media protocols are able to participate in the "Fediverse" network of servers (The Federation 2020).

7. Decentralised file transfer systems

Example: Hypercore (Hyperdivision, 2020), IPFS (Benet, 2014), Bit Torrent (Cohen, 2008), Secure scuttlebutt (Tarr et al., 2019)

This category encompasses digital institutions that are decentralised peer-to-peer file transfer systems oriented around peer-to-peer communication and information handling protocols. Other than the protocol standards themselves, they do not rely on the maintenance of consensus around the state of a shared digital object. The protocol rules are typically defined by sets of developers. These may be loosely connected via conversation platforms, or oriented around a foundation or company.

Mapping of Digital Institutions

The identified digital institution types were mapped onto the layered model, see Table 2. For each digital institution, two comparison metrics were defined and used to estimate the qualities of the architectures:

1. **Independent verification**: Does the architecture allow for independent verification, with no influence from those that define and run the system.

2. **Tamper resistance.** Is the architecture resistant to the possibility of targeted persecution through abuse of transaction outcomes by third parties.

Table 2 – Taxonomy of digital institutions

	1. Traditional web server	2. Public blockchain fork-based	3. Public blockchain integrated	4. Hybrid web-server/blockchain	5. Coalition blockchain	6. Federated server templates	7. Decentralised file transfer systems
Examples	Airbnb, Amazon, eBay, Uber	Bitcoin, Ethereum	Polkadot, Tezos	Steemit	Libra	Fediverse	Hypercore, IPFS, BitTorrent
Governance	Hierarchical management structure	Informal discussion (on public platforms or in private communication).	Integrated, at least in part, into the rules layer.	Project based, hierarchical management structure	Hierarchical management structure Coalition agreed licensing requirements	Informal coalition of software developers	Project based, typically hier archical management structure or a loose coalition of software developers
Rule change enactment	Direct	Fork-based	Integrated	Direct and fork-based	Direct	Direct	Direct
Interface	Private computer access, Email address, Bank account	Private computer access, Private key / transaction creation software	Private computer access, Private key / transaction creation software	Private computer access, Private key / transaction creation software	Private computer access, Email address, Bank account, Social media account	Private computer access	Private computer access
Instantiation	Network of computer systems, controlled by a single party	Network of computer systems, controlled by multiple unknown parties	Network of computer systems, controlled by multiple unknown parties	Network of computer systems, controlled by a single party, with some use of a network of computer systems, controlled by multiple unknown parties	Network of computer systems, controlled by licensed parties	Network of computer systems, controlled by multiple parties	Network of computer systems, controlled by multiple unknown parties
Rules	Application specific	Application specific, include consensus algorithms (ensure strangers agree on transaction outcome)	Application specific, include consensus algorithms and governance mechanisms	Application specific	Application specific	Application specific, applications share common communication and user identity standards.	Varies by project (some have an economic system – Filecoin) Others do not define economic interactions.
Independent verification?	No	Yes	Yes	Varies by project	Yes	No	Protocol dependent
Tamper resistant?	No	Yes	Yes	Varies by project	No	Partial	Varies by project

5. Conclusions

What emerged from our research is that digital sustainability in the context of decentralised governance is a recognition that effective institutions are not necessarily those who require self-sameness over time as a defining feature. Instead, we recognise an emergent form of sustainability as the capacity to allow for transactionally integrated and/or fork-based changes to modes of operation, which are tamper resistant, and captured via independently verifiable public audit trails. Our analysis suggests that *forking* does not break with notions of effective institutional governance; instead, it is a defining feature of the emergent decentralised governance paradigm. By transforming perceptions of trust in this way, we suggest that these new governance practices stand to enhance the sustainability of digital institutions.

The layered model encompasses both the Zamfir's and Wood's perspectives. Whereas Zamfir's ideas largely reflect the horizontal, co-governance approach of the Ostrom-polycentric institution model, Wood's views on governance centre on safeguarding the future cohesion of a given digital institution reflect the ODI's self-sustainability model. In this sense, the capacity of decentralised architectures to facilitate implementation of decisions in a publicly auditable way, through integration of governance processes (e.g., votes) and enactment within the system's information exchange ruleset, or through enactment in the creation of a network fork, are both examples of decentralised governance. Furthermore, we also recognise that conceptions of trust have evolved, or at least concretised, as a result of the emergence of decentralised blockchain systems. It is now essential that we differentiate trust in systems, and trust in users, developers and moderators, from trust in the *output* of a system.

The literature synthesis, layered model and taxonomy contributed here provide a framework for contextualising and articulating sustainable decentralised governance of digital institutions. As a result, in terms of identifying the key contribution that blockchain technology can be seen to have made to the question of digital sustainability – as the layered model and taxonomy of digital institutions indicates – we find that it is the capacity to facilitate decentralised governance through integrated and/or fork-based changes to intermediating rulesets, which are tamper resistant, and captured via independently verifiable public audit trails, that marks out a paradigm shift with regard to the sustainability of digital institutions.

About the authors

Dr Matthew Lovett is Academic Lead for the Music and Sound Department at the University of Gloucestershire. Rooted in systems thinking and ecomusicology, his work explores the common ground between music, technology, economics and philosophical theory.

Dr Lee Thomas is an independent researcher with a background in power engineering and computing. He is founder of ooda.space, a transparent and decentralised system that aims to accelerate the revelation of valuable ideas.

Notes

[1] The concept of forking is widely understood within open source software development. Its contribution to sustainability, in the sense that a software application will continue to serve its developers and users, stems from the right (granted by software licence) to duplicate and edit the codebase (Nyman 2013). As software architectures become more decentralised and increasingly performs broader roles, including societal level institutional functions (e.g. payments, votes, contract enforcement), the scope and impact of forking takes on a broader meaning. There is question of how the norms around the distributed governance of open source software, as examined by Schweik (Schweik 2003) for instance, might apply in this broader sense. Further, there is a question of whether new software integrated modes of governance might emerge as a result of the decentralised nature of transaction-oriented systems, such as blockchain technology.

[2] Although the wider implications of Ostrom's work go beyond the current study, it is relevant that amongst her key insights was her distinction between public goods and common pool resources. A public good is non-excludable and non-rival, which is to say that a public good can be used freely, and one person's use of it does not diminish another's (Apesteguia and Maier-Rigaud, 2006: 647). A common pool resource however, shares this non-excludability, but is rivalrous, which is to say that if one person's use of it increases, then another's decreases (Choe and Yun, 2017: 117). Given our current analysis of how blockchain is contributing to debates surrounding the evolving nature of governance and

digital institutions, it is noteworthy that Choe and Yun problematise Ostrom's original conception of a CPR by suggesting that CPRs must be understood as historically located, and that seemingly permanent characteristics of air and water, such as non-excludability and non-rivalry, can change (Choe and Yun, 2017: 117). Their study details how air has become a rivalrous (their equivalent term is 'subtractable') good due to 'population growth and air pollution', and has become excludable because of emission trading schemes. Similarly, varying degrees of water scarcity mean that for some, water too is a rivalrous good (Choe and Yun, 2017: 117).

[3] In his examination of the DAO – 'a failed decentralized autonomous organization' built on Ethereum (DuPont, 2018: 157) – DuPont describes how the system was compromised and then ransacked to the tune of millions of dollars' worth of ETH tokens. Ethereum's solution to create a 'hard fork' version of the software erased the DAO and stemmed the flood of tokens, but simultaneously created a rift between 'moderate' and 'ideological' minors. DuPont relates that 'the "moderates" saw the hard fork as evidence of the flexibility and practicality of Ethereum and its leaders, while the more ideological saw the hard fork as censorship by a powerful cabal, or proof that blockchain technology was unable to live up to its idealistic promises' (DuPont, 2018: 165), where the former continued to use Ethereum in its new – and still current – form, whilst the latter migrated to Ethereum Classic in a move that demonstrated their commitment to the sanctity of code, rather than the necessary flexibility of decision-making and governance frameworks.

References

Benet, J. 2014. 'IPFS - Content Addressed, Versioned, P2P File System'. *ArXiv* abs/1407.3561.

Buterin, V. 2015. 'A Next-Generation Smart Contract and Decentralized Application Platform'. https://ethereum.org/en/whitepaper/ (accessed December 2020).

Chalmers et al. 2019. 'Blockchain as an external enabler of new venture ideas: Digital entrepreneurs and the disintermediation of the global music industry'. *Journal of Business Research*. Available at https://doi.org/10.1016/j.jbusres.2019.09.002 (accessed September 2020).

Choe and Yun. (2017) "Revisiting the Concept of Common Pool Resources: Beyond Ostrom," *Development and Society*, 46(1), pp. 113–129.

CleanApp. 2018. 'The Wood-Zamfir Governance Debates: Two of the sharpest minds in crypto disagree on what it means to agree. Or do they?'. Available at https://medium.com/cryptolawreview/the-wood-zamfir-governance-debates-80e92436a457 (accessed November 2020).

Cohen, Bram. 2008. 'The BitTorrent Protocol Specification'. 10 January 2008. https://www.bittorrent.org/beps/bep_0003.html. (accessed December 2020).

Deloitte, 2020. Deloitte's 2020 Global Blockchain Survey: From Promise to Reality. Available at https://www2.deloitte.com/content/dam/insights/us/articles/6608_2020-global-blockchain-survey/DI_CIR%202020%20global%20blockchain%20survey.pdf (accessed November 2020).

Deutsch, D. 1997. 'The Fabric of Reality'.

Diem Association, 2020. ‘White Paper v2.0’. Available at https://www.diem.com/en-us/white-paper/. (accessed December 2020).

DuPont, Q. 2018. ‘Experiments in algorithmic governance: A history and ethnography of “The DAO,” a failed decentralised autonomous organization’. In Campbell-Verduyn, M. (ed.) (2018) *Bitcoin and beyond: cryptocurrencies, blockchains, and global governance*. London: Routledge, Taylor & Francis Group.

The Federation. 2020. ‘The Federation Webpage’. 2020. Available at https://the-federation.info/#protocols (accessed November 2020).

Gallersdörfer, U., Klaaßen, L. and Stoll, C. (2020) “Energy Consumption of Cryptocurrencies Beyond Bitcoin,” 4(9), pp. 1843–1846. doi: 10.1016/j.joule.2020.07.013.

Goodman, LM. 2014. ‘Tezos—a Self-Amending Crypto-Ledger White Paper’. Available at https://tezos.com/static/white_paper-2dc8c02267a8fb86bd67a108199441bf.pdf (accessed December 2020).

Hansen, M et al. 2014. ‘Localizing global environmental governance norms: Implications for justice’. In Sowman, M. (Ed.), Wynberg, R. (Ed.). *Governance for Justice and Environmental Sustainability*. London: Routledge.

Hsieh, Y-Y et al. 2018. ‘The internal and external governance of blockchain-based organisations: evidence from cryptocurrencies’. In Campbell-Verduyn, M. (ed.) (2018) *Bitcoin and beyond : cryptocurrencies, blockchains, and global governance*. London: Routledge, Taylor & Francis Group.

Hütten, M and Thiemann, M. 2018. ‘Moneys at the margins: From political experiment to cashless societies’. In Campbell-Verduyn, M. (ed.) (2018) *Bitcoin and beyond : cryptocurrencies, blockchains, and global governance*. London: Routledge, Taylor & Francis Group.

Hyperdivision, 2020. ‘Hypercore Protocol’. Available at https://hypercore-protocol.org/ (accessed December 2020).

Kiayias, A., B. Livshits, Andrés Monteoliva Mosteiro, and O. Litos. 2019. ‘A Puff of Steem: Security Analysis of Decentralized Content Curation’. In *Tokenomics*.

Krause, M. J. and Tolaymat, T. (2018) “Quantification of Energy and Carbon Costs for Mining Cryptocurrencies,” *Nature Sustainability*, 1(11), pp. 711–718. doi: 10.1038/s41893-018-0152-7.

Li, J. *et al.* (2019). “Energy Consumption of Cryptocurrency Mining: A Study of Electricity Consumption in Mining Cryptocurrencies,” *Energy*, 168, pp. 160–168. doi: 10.1016/j.energy.2018.11.046.

Musiani, F et al. 2018. ‘Governing what wasn’t meant to be governed: A controversy-based approach to the study of Bitcoin governance’. In Campbell-Verduyn, M. (ed.) (2018) *Bitcoin and beyond: cryptocurrencies, blockchains, and global governance*. London: Routledge, Taylor & Francis Group.

Nakamoto, Satoshi. 2009. ‘Bitcoin: A Peer-to-Peer Electronic Cash System’. Available at https://bitcoin.org/bitcoin.pdf (accessed December 2020).

Ngeta, K. 2014. ‘Theoretical deliberations on understanding natural resource governance’. In Sowman, M. (Ed.), Wynberg, R. (Ed.). *Governance for Justice and Environmental Sustainability*. London: Routledge.

Nyman, Linus, and Juho Lindman. 2013. ‘Code Forking, Governance, and Sustainability in Open Source Software’. *Technology Innovation Management Review*, 6.
Open Data Institute. 2020. ‘Designing sustainable data institutions. Available at http://theodi.org/wp-content/uploads/2020/04/OPEN_Designing-sustainable-data-institutions_ODI_2020.pdf (accessed November 2020).

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, E. (2010) ‘Beyond Markets and States: Polycentric Governance of Complex Economic Systems’, *American Economic Review*, 100(3), pp. 641–672. doi: 10.1257/aer.100.3.641

Rose, Anna, Fredrik Harrysson, Gavin Wood, and Vlad Zamfir. 2018. 'Vlad and Gav Go Head-to-Head on Blockchain Governance'. Available at https://www.zeroknowledge.fm/52 (accessed November 2020).

Schweik, Charles M., and Robert English. 2007. 'Tragedy of the FOSS Commons? Investigating the Institutional Designs of Free/Libre and Open Source Software Projects'. *First Monday*, May. https://doi.org/10.5210/fm.v12i2.1619.

Tarr, Dominic, Erick Lavoie, A. Meyer, and Christian Tschudin. 2019. 'Secure Scuttlebutt: An Identity-Centric Protocol for Subjective and Decentralized Applications'. *Proceedings of the 6th ACM Conference on Information-Centric Networking*.

Toonen, T. (2010) "Resilience in Public Administration: The Work of Elinor and Vincent Ostrom from a Public Administration Perspective," *Public Administration Review*, 70(2), pp. 193–202.

United Nations Department of Global Affairs. 2015. 'The 17 Goals. Available at https://sdgs.un.org/goals (accessed September 2020).

UN. 2018. "World economic and social survey 2018: frontier technologies for sustainable development", United Nations. Available at https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/WESS2018_full_web.pdf (accessed November 2020).

Waage, J et al. 2015. 'Governing Sustainable Development Goals: interactions, infrastructures, and institutions. In: Waage, J and Yap, C. (eds.) *Thinking Beyond Sectors for Sustainable Development*. Pp. 79–88. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/bao.i

Werbach, K., 2018. *The blockchain and the new architecture of trust*. Cambridge, MA and London, England: The MIT Press.

Wood, Gavin. 2014. 'Ethereum: A Secure Decentralised Generalised Transaction Ledger'. 2014. https://ethereum.github.io/yellowpaper/paper.pdf (accessed December 2020).

Wood, Gavin. 2016. 'Polkadot: Vision for a Heterogeneous Multi-Chain Framework'. White Paper. 2016. https://polkadot.network/PolkaDotPaper.pdf (accessed December 2020).

Zwitter, A and Hazenberg, J. 2020. 'Decentralized Network Governance: Blockchain Technology and the Future of Regulation'. *Frontiers in Blockchain* 3:12. doi: 10.3389/fbloc.2020.00012 Available at https://www.frontiersin.org/articles/10.3389/fbloc.2020.00012/full (accessed November 2020).