Sovereign Sensors

An investigation of factors pertaining to the governance of informational resources in the context of decentralization

Masters of Science Dissertation
Spatial Data Science and Visualisation
UCL Bartlett Centre for Advanced Spatial Analysis
2018 - 2019 Academic Year

I, xxxxxxxxxxxxxxxxxxxxxxxxxxx, confirm that the work presented in this dissertation is my own, inspired and informed by many others. Where information has been derived from other sources, I confirm that this has been indicated.

XXXXXXXXXXXXXXXXX

IPFS hash: QmZyCNozTzYdVk2Ngn3gzpYoSVSchdUP1CnJG6T2K9Fic5

Abstract

As connected sensors capture and process ever greater amounts of information about their local vicinities, growing networks of these devices create the potential for improving situational awareness and, therefore, the efficiency and virtue of resource management. However, due to the technical, commercial and legal realities of these systems in the present day, much of the informational value these networks represent to humanity remains unaccessed. Concurrently, advancements in information technologies - especially cryptography - are enabling the emergence of a decentralized web paradigm designed to preserve individual autonomy and privacy. Most notably, blockchains are emerging to provide the public with transparent, open and neutral information computing and storage resources that operate independently of any central authority, and are highly secure and antifragile. This dissertation seeks to identify and define concepts key to connected sensor networks and the decentralized web. The fundamentals of enabling technologies including computing, cryptography and digital communication are reviewed, and relevant decentralized technologies such as smart contracts, decentralized autonomous organizations, content-based addressing, self-sovereign identities and Ricardian contracts are defined. An agent-based model simulating the operation of connected sensors connecting to a public blockchain network was developed; quantitative results of parameter sweeps of three input parameters are reported and analyzed. Considerations concerning the connection of edge sensors with smart contracts are presented. The technical and political feasibility and ethics of high impact applications of such systems are discussed. Finally, a possible limitation in our current paradigm is contemplated.

Figures

Tables

Statement of Ethics

I engage in this work with an ethical mindset, and hope that this writing conveys that I strive to consider deeply not only the technical, but also the moral, social, economic and political implications of these subjects. Generally, I am oriented toward improving human dignity and supporting us in our role as stewards of the planet.

But unless we recognize the overall failures of our current systems we most probably don’t stand a chance.

Greta Thunberg
2019

… to create a world that better preserves the autonomy of the individual …

Vitalik Buterin
2016

1 Introduction

Objects exist in space. These entities interact over time. We know this because they come into our awareness, moment to moment. We[1] are objects that perceive the world, observing it through these moments while we are alive.

This dissertation is a reporting back of the current state of my understanding of these topics. My overall interests include the patterns and relationships of our lived experience, including the natures of both matter and meaning, and the relationship between the two. My angle is the simplest instance of which I am aware of informational entities communicating: connected sensors - computing nodes interpreting, transmitting and receiving binary data. My goal is to investigate the intricacies of these systems of connected sensors in the current moment - 2019.

The relevance is difficult for me to overstate - I am drawn to pursue this research because I feel it might help us to understand the enormous risks and opportunities inherent in the technologies we are just beginning to invent and discover. My hope is that if we can understand these risks we might be able to mitigate them, just as if we do the same for the opportunities they might be maximized toward the promotion of the values we share. I believe this is a pivotal moment[2] for us, here on Earth today. It seems we have the capacity - and perhaps the capability - to finally achieve what so many just people have spent their lives pursuing: the eradication of violence and the peaceful coordinated thriving of life[3].

The risks of this endeavor cannot be overstated. What I intend to learn will imbue me with great power, the ability to help create the services that people and machines may use to live in the 21st century - an ability that confers moral responsibility. I accept this, but cannot say I understand it - I am only beginning to grasp the ethical implications of this emerging reality. For this reason the work will

[1] As in, us humans.
[2] Perhaps every present moment is a pivotal one? This one seems especially so.
[3] I am not a technological solutionist but I do believe that, as a tool, technology can help us achieve our intentions.

describe a number of ethical principles, adopted or adapted or arrived at in my studies.

Specifically, this research explores the interaction of networks of connected sensors with blockchain networks. In order to adequately define terms and frame the research, the necessary technologies will be reviewed and clear[4] explanations provided. A middle-range agent-based model simulating key features of edge sensors, quasi-Turing-complete smart contract platforms and their interaction was built, as described in the Methodology section. Model runs produced simulation output of parameter sweeps of three independent variables; three dependent variables were measured and analyzed, presented as Results.

In the Discussion I briefly review the model findings, and propose the outlines of a few system configurations that utilize the potential of these systems in a just and inclusive way.

The Conclusion will raise some of the questions I encountered that were not satisfactorily answered in the current scientific paradigm, as I understand it.

A distinction needs to be made, between people that are[5] alive, and everything else; it seems important - paramount - that place higher value in the former. If each person is equal[6], this humanist ideal must be ingrained in us and in everything we create.

[4] Hopefully ...
[5] And will be.
[6] In value, yes. In qualities? Never.

1.1 Research Questions

General

What is information? What is data? How are they related? How are they produced, stored, transported and used? How are they governed?

Specific

What opportunities, risks, constraints and limitations are inherent in the use of edge sensors as oracles for smart contracts?

Agent-based model

What are the effects of network size and data capture and transfer dynamics on public blockchain validation patterns and costs?

2 Literature Review

2.1 Computers

Humans have relied on physical objects - tools - to help perform logical operations since the origin of writing and mathematics (Huang 2017). A common feature of these tools (tally sticks, abacuses, and so on) is their ability to store symbolic representations of numerical values so the operator did not have to remember the values. With the aid of these tools, human understanding of the patterns of mathematics developed.

Leibniz and Babbage[7] paved the way for Alan Turing to create his (now known as) Turing complete “a-machine” (Turing 1937), a mathematical model of computation capable of “simulat[ing] any computer algorithm (Mullins 2012).

On this basis, modern computing arose. Computers apply binary operations to sequences of bits[8], represented in contemporary computers as electrical voltages stored in capacitors. A bit sequence is computed by the processing unit according to its data type[9]; an identical sequence carries different information for the processor can be interpreted and processed differently[10] based on this metainformation (Malan 2014).

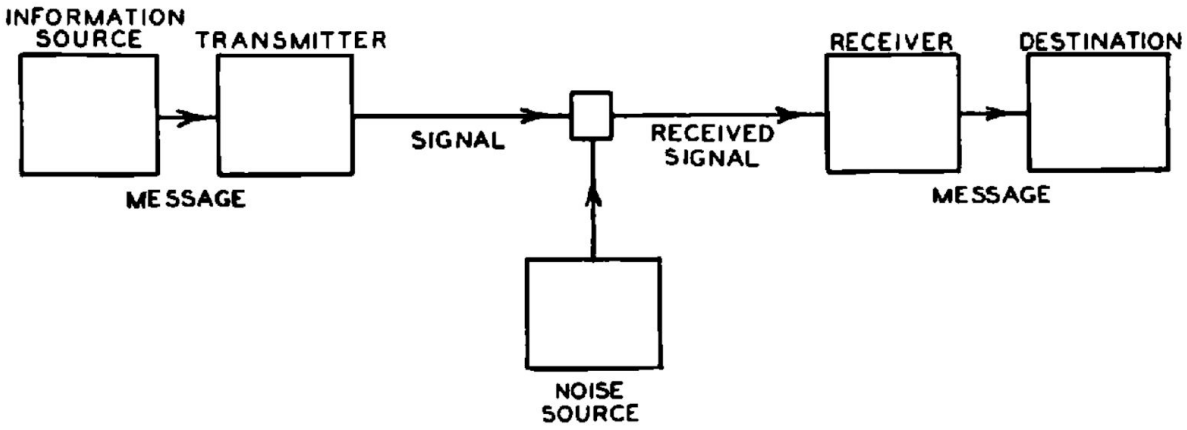
2.2 Communication

Building on knowledge of how information is transmitted on an electrical or broadcast signal[11], as well as the community’s emerging understanding of the behavior of electrical circuits and its implications for data processing and the

[7] See Appendix 5.1 for more on the history of computing.
[8] One byte is 8 bits, one kilobyte is 1000 bytes, and so on, according
[9] I.e. Real, integer, boolean (Wikipedia 2019a).
[10] i.e. different binary operations applied.
[11] See Appendix 5.2 for more on the history of communication.

encoding of information on a serial channel[12], Claude Shannon published his master's thesis[13], *A symbolic analysis of relay and switching circuits* (1938). A decade later, in *A mathematical theory of communication*, Shannon (1948) defined the fundamental principles of the transfer of a message from source[14] to destination[15]: "the theory that lies behind any phenomenon involving data encoding and transmission" (Floridi 2010)[16].

Figure 1: Schematic diagram of a general communication system



Reprinted from Shannon (1948 pp 381)

Shannon's paper described how a binary message is encoded on a signal[17] by the information source, or informer. The signal is transmitted on a (perhaps noisy) channel[18], then detected by the receiver and interpreted by the information destination, the informee.

[12] A one dimensional, serial channel, as opposed to a two dimensional, parallel channel - see Appendix 7.
[13] Just one of many factors contributing to this author's sense of intellectual inadequacy …
[14] Also informer; origin.
[15] Also informee, recipient.
[16] For clarity, this theory pertains to the measurement and communication of units of information: by Shannon's definition, a message consists of a sequence of symbols - Shannon information (Floridi 2010 pp 38). "If base 2 is used, the resulting units may be called binary digits, or more briefly bits, a word suggested by J. W. Tukey" (Shannon 1948 pp 380). In the Conclusion and Appendix 7 the limits of this theory of digital communication will be explored, contingent upon on a definition of "message" broader than Shannon's. This is not to discredit the brilliance and profound impact Shannon's work has had on the world by applying a rigorous conceptual framework to the mathematical treatment of binary information.
[17] A carrier wave of energy propagated through a medium.
[18] Either cable or broadcast.

Digital information is encoded into a digital signal by changing some quality of the carrier wave - modulating the amplitude, frequency or phase of fixed-duration pulses of electrical current (Nair 2002 pp 289) or increments of light (Dume 2012, Economist 2017). The information receiver, able to detect the regular modulations of the signal, can decode the binary sequence from the signal upon reception.

Figure 2: Binary information encoded on a signal

By Mcanet - Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=6024833

Shannon's work formalized the necessary framework for the encoding and transmission of binary information on a digital signal; computers executing instructions encoded in binary were developing in sophistication and usage. From these advancements, networked computing arose.

2.2.1 Cryptography

In the context of digital computing, cryptography applies binary operations and arithmetic to modify these binary sequences - messages intended to carry meaning to an informee - in predictable and sometimes reversible ways. "A message is a discrete unit of communication intended by the source for consumption by some recipient or group of recipients" (Wikipedia 2019b).

To maintain privacy[19] on an open network, message senders need to be assured that:

- only intended recipients can access the contained information - they are "authorized";
- intermediary relayers cannot access the contained information - the message is "confidential"; and
- the message received is identical to the message sent - it has "integrity".

To provide these three assurances - confidentiality, authorization, and integrity - cryptographic algorithms and protocols have been developed, enabling the establishment of secure communication connections on open[20] channels. A basic familiarity with their purpose and function is important to understanding subsequent concepts.

Table 1: Cryptography Key Terms

cipher	"a secret or disguised way of writing; a code" (Oxford 2019a)
plaintext	an unencrypted message, in which the information contained is visible[21]
ciphertext	an encrypted plaintext, in which the information contained is obscured
algorithm	"a process or set of rules to be followed in calculations or other problem-solving operations" (Oxford 2019b)
key	"a sequence of bits" (Techopedia 2019), sometimes with certain mathematical properties
protocol	"a set of agreed-upon conventions" (Cerf 1974)

[19] "Privacy is a human right. It is also the fountainhead for all of the other human rights. If you don't have privacy, you don't have freedom of expression, you don't have freedom of association, you don't have freedom of assembly ... without financial privacy you don't have political rights." (Antonopoulos 2019 57:34). Significant further discussion of the ethics of privacy and the extent of individuals' rights to it is left for future work.
[20] i.e. public.
[21] To anyone familiar with the encoding scheme.

2.2.1a Integrity

Hashing Algorithms

Cryptographic hashing algorithms accept a message of arbitrary length. Based on this unique bit sequence, the algorithm calculates a cryptographic hash (or digest), which is a fixed-length sequence computed from the input value. Most often this hash is substantially shorter than the input data[22] - for example, the SHA-256[23] algorithm returns a 256-bit (32-byte) hash for an input of effectively[24] any length (Veness 2019).

To be useful, cryptographic hashing algorithms must have a few crucial features. They must be deterministic: the same input sequence will always yield the same output hash. Extending this, changing even a single bit in the input message must result in a different output hash[25]. Also, they must be one-way, or trapdoor, functions: it must be extremely difficult to calculate the input data from the message digest[26].

Cryptographic hashes serve as a kind of message "fingerprint": they are effectively[27] unique to the message. This means that, if a message is transmitted along with its hash, the recipient can confirm that the message's integrity is uncompromised by hashing the message on their computer and comparing the output hash with the one included in the message[28]. Additionally, producing and sharing a hash can prove an entity has a specific datum[29] without sharing the datum[30] (Ker 2014).

[22] i.e. it is a "compression function".
[23] "Secure Hashing Algorithm", returning a 256-bit hash. Other common hashing algorithms or algorithm families include MD5, SHA1 and Keccak.
[24] Up to 2,091,752 terabytes (Stackoverflow 2016)
[25] And, ideally (for security purposes), a dramatically different hash - "diffusion".
[26] This is termed "pre-image resistance".
[27] Though not exactly: "note that the compression property implies that the function must be many-to-one, because the domain is infinite and the codomain finite, so infinitely many collisions exist" (Ker 2014 pp 48). In collision resistant hashing algorithms "it should be computationally infeasible to find any" collisions, where different inputs yield the same output.
[28] Often hashes are used as checksums to ensure that a file has downloaded in full and uncompromised (Wikipedia 2019d).
[29] Such as a file.
[30] As long as the recipient knows the hash or has the datum themselves, and can compute the hash.

Figure 3: SHA-256 hashing algorithm Invocation[31]

```
>>> sha256("ucl")
"43c34ee9af7bfaccca6b3bd5d2af0d96bab09732aa5a3dc63a5eaa7015f2a8ce"[32]
```

2.2.1b Confidentiality and Authentication

Symmetric Key Ciphers

Symmetric key ciphers allow communicators to maintain confidentiality even if a message is viewed by an unauthorized entity. The message plaintext is encrypted with a secret key by the informer in a private, secure computing environment prior to transmission on the open network. In this process, an arbitrary-length binary plaintext and a secret key are passed into an algorithm, which returns the encrypted message - the ciphertext. To decrypt, the ciphertext and the key are passed into an algorithm, which returns the plaintext, thereby unobscuring[33] the information represented[34] (Ker 2014).

Asymmetric Key Cryptosystems

While symmetric key ciphers provide communicators confidence in message confidentiality, they are constrained by the need for both parties to have the same secret key. It is infeasible for every pair of communicators establishing a secure connection on a public channel to meet and exchange keys. Transmitting a plaintext key on an open network leaves risk of unauthorized intermediaries detecting it;

---

[31] If the `sha256()` function has declared with the SHA-256 algorithm.
[32] Any online SHA256 hash calculator computing the three letters "ucl", all lowercase will yield the output shown, if the encoding is the same (utf-8).
[33] This phenomenon provides one of the best opportunities to understand the relationship between matter and meaning, in this author's view. This will be cursorily explored in the Conclusion.
[34] While explanation of the mathematical processes specific to various key cipher algorithms is beyond both the scope of this paper and the intellectual capacity of this author, it should be noted that these algorithms rely heavily on the application of the XOR ("exclusive or") binary operator, which returns 1 for differing input operands ($1 \oplus 0 = 1$ and $0 \oplus 1 = 1$) and 0 if input operands are the same ($0 \oplus 0 = 0$ and $1 \oplus 1 = 0$). Because of this operation's properties - it is "commutative, associative and self-inverse" (phant0m 2013) - the application of the operation to the ciphertext with the same key outputs the plaintext. To improve security, more sophisticated symmetric key ciphers have been developed, such as the Advanced Encryption Standard (Daeman 1999, National Institute of Standards and Technology 2001).

attackers could, if aware of the algorithm used[35], decrypt every intercepted ciphertext encrypted with that key.

To resolve this, asymmetric key ciphers and key exchange protocols were developed, based on the work of Ellis, Cocks, Diffie and Hellman, Rivest, Shamir and Adleman, and Merkle (Ker 2014). The algorithms and protocols developed and defined enabled two parties "to communicate confidentially after transmitting a key which is not confidential" (Ker 2014 pp 71).

The distinguishing feature of asymmetric key ciphers is the creation of two keys: a public and a private key. These numbers are calculated using a key generation algorithm, and are linked based on their mathematical relationship.[36]

Two pairs of algorithms characterize asymmetric key cryptosystems, based on these public-private keypairs: encrypt / decrypt and sign / verify.

Encrypt / Decrypt

Asymmetric key ciphers include algorithms by which messages can be encrypted and decrypted, as in symmetric key ciphers. However, a critical difference exists: encryption is performed by passing the plaintext and the message recipient's[37] public key into the encryption algorithm. This is a trapdoor function, one "easy to perform and difficult to invert" (Ker 2014). The resultant ciphertext can be decrypted by passing it, along with the recipient's private key, into the decryption algorithm, which returns the plaintext.

By disaggregating these two functions, the creators of public key cryptosystems created a way to establish a secure connection on a public channel: communicators could simply transmit their public keys to each other in plaintext form, while

[35] which is always assumed, according to Kerckhoffs's Principle, concisely stated by Shannon as "the enemy knows the system" (Kerckhoffs 1883, Shannon 1949)

[36] For example, key generation in the RSA cryptosystem, directly from Ker (2014 pp 74):

(1) Choose two prime numbers each b/2 bits in size, call them p and q. Compute their b-bit product n = pq and the so-called totient $\varphi = (p - 1)(q - 1)$.
(2) Choose an integer e which is coprime to φ.
(3) Publish the public key, which is the pair pk B = ⟨n, e⟩.
(4) Find an integer d such that $ed \equiv 1 \pmod{\varphi}$.
(5) The private key is the pair skB = ⟨n,d⟩2.

[37] NOT message sender's.

maintaining the secrecy of their private key. The message sender could then encrypt the message with the recipient's public key, confident that it could only be decrypted with the corresponding private key - which only the intended informee had[38].

Sign - Verify

Due to the mathematical properties of these public-private keypairs, a second pair of algorithms was possible in the cryptosystem: sign and verify. This pair of algorithms provides message recipients confidence that a signed message was sent by the holder of the private key. If private keys are properly managed[39], this allows informational entities to prove their identity (the basis of authentication) by proving they possess the private key, without ever sharing that key.

The signing algorithm accepts a signer's private key and a message to be signed; the function returns a signed message, or digital signature.

```
>>> signed_data = sign( data, private_key )
```

The verify algorithm, when provided the signed message and the signer's public key, enables the message recipient to confirm that the message was signed with the signer's private key (Ker 2014).

```
>>> verify( signed_data, public_key )
returns True if signed_data was signed with private_key,
        False if not[40]
```

This sign-verify functionality of public key cryptosystems is the enabling technology behind the decentralized web and, in this author's view, represents a profound opportunity to improve the systemic justice of the Internet.

[38] This is, of course, a significant simplification of the complexity of these cryptosystems, but hopefully captures the relevant features of these algorithms for the purposes of this paper.
[39] A big "if", and a primary, and valid, critique of decentralized web technologies. If a private key is lost, in these systems it is effectively impossible to retrieve the assets encrypted or controlled by it.
[40] This is an abstraction.

2.2.1c Cryptographic Protocols

By combining the functionalities of cryptographic hashing algorithms and symmetric and asymmetric key cryptosystems, standardized procedures (i.e. protocols) have been established enabling communicators to verify sender authenticity, ensure message confidentiality[41], and confirm message integrity[42] using technical mechanisms (Ker 2014).

2.2.1d Homomorphic encryption

Although a thorough treatment is well beyond this paper's scope and author's comprehension, "homomorphic encryption is a form of encryption that allows computation on ciphertexts, generating an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on the plaintext" (Wikipedia 2019c, Gentry 2010). Crucially, the plaintext information encoded in the encrypted message is never revealed to the computer processing it, meaning these computations can be executed on untrusted devices while preserving the privacy of its subject.

2.2.1e Proxy re-encryption

In these schemes a message sender can direct a third party proxy node to re-encrypt a ciphertext that has been encrypted for one recipient to be decrypted by another without revealing plaintext to the node. This is significant especially in the context of resource-constrained edge devices; re-encryption for multiple recipients can be outsourced to better-resourced computing environments[43] (Nuñez 2018).

[41] via Message Authentication Codes, MACs.
[42] via digital signatures and keyed hashes.
[43] like a cloud server.

2.2.2 Networked computers

As general purpose analytical machines, Turing-complete computers processing binary information began to be adopted by the military, industry and academia (CrashCourse 2017b) to efficiently and accurately compute and catalogue (relatively) large volumes of data. System developers began to see that the transfer of binary information from one computer to another would enable informational resources to be shared amongst them, thereby enabling the execution of more complex applications, as well as the ability to access data stored on other computers[44].

Innovation in techniques for connecting distant computers and transmitting data culminated in the publication of *A Protocol for Packet Network Intercommunication* (Cerf and Kahn 1974). The paper established the Transport Communication Protocol, which "provides reliable, ordered, and error-checked delivery of a stream of octets (bytes) between applications running on hosts communicating via an [Internet Protocol] network" (Wikipedia 2019d).

These protocols established a standard format for data packets[45] to be transmitted to a destination address, included in the message. Critically, in his design of the protocols Kahn adhered to the principle that "There would be no global control at the operations level" (Leiner 1997). Rather than relying on a centralized authority to organize traffic - which Kahn likely understood to be unscalable - the authority would reside in the rules of the protocol, to which system developers would adhere - an early instance of decentralization in computing technology.

2.2.2a The Web (1.0)

With the establishment of the Transport Control Protocol and the Internet protocol suite, local area networks using heterogeneous operating systems had a standardized way to structure data to transmit to external networks. While this represented an enormous leap forward in information accessibility, it was only with the creation of the World Wide Web and web browsers - alongside the increase in

[44] For a more thorough history of internetworking, see Appendix 5.3
[45] binary messages

public access to computers - that these advancements were made more broadly accessible.

The Web 1.0 was characterized by static web pages connected to resources stored elsewhere on the network by hyperlinks: “a single user-interface to large classes of information” “to link and access information of various kinds as a web of nodes in which the user can browse at will” (Berners-Lee 1990).

2.2.2b The Web 2.0

Concurrent improvements in web communications technology, advancements in usability and design, and increases in web-enabled computer access resulted in the emergence of the “Web 2.0” (DiNucci 1999), characterized by interactivity and a social aspect. Services came to provide more utility, entertainment value and customizability for users (Choudhury 2014). Due to the high capital and labor costs required to provide reliable data storage and access services used by the public, economies of scale, and the network effects such platforms created, and competitor acquisitions, large corporations came to dominate different swathes of the web.

In order to sustain operations and deliver a return to shareholders, their assets needed to be monetized. Charging users was rarely viable - users seem to feel that information and information services should be free. Instead, web-connected devices became channels to serve advertising to users. And, due to the information users were sharing with such platforms, advertising could be targeted to a degree unimaginable prior to the digital era. A positive feedback mechanism has emerged by which more data yields better services[46], which yields more data (Ibarra 2017).

In 2019, calls to break up “Big Tech” abound (Waters 2019); public criticism of the ethical and security practices of these technology providers is rising (Williams 2018, Lee 2018).

[46] Due to “the role [they] play in enabling [machine learning]” (Ibarra 2017 pp 4).

The Internet of Things

Increasing connectivity and decreasing equipment costs has led to the inclusion of computing and data transfer capabilities on a widening array of devices. Embedding sensors and connected computing nodes on vehicles, infrastructure, appliances, surveillance equipment (on Earth and in orbit) and so on enables device controllers to access data stored on and captured by the device and, possibly, the ability to actuate device operation (Xu 2014). Devices are increasingly aware of their surroundings and location in space, and able to interpret their environment and adapt (Zanella 2014). Battery-operated, wirelessly connected computing devices capable of transmitting data representing measurements taken from sensors or other on-board data are being deployed on huge scales; it is projected that by 2025, "there will be 41.6 billion connected IoT devices … generating 79.4 zettabytes (ZB) of data" (Shirer 2019).

2.2.2c The decentralized web

As the records underpinning society's operation have been digitized, the responsibility for maintaining and updating those databases remained where it had historically resided: with governments, banks, medical institutions and other centralized authorities. Providing this service incurred substantial cost in terms of capital equipment and operating and labor costs[47], and these authorities continued to be trusted to securely and accurately maintain the records, and to act in the interests of the users, the subjects of the data in their custody.

This presented an agency problem[48] for many of these trusted data custodians: what was in their (often financial) interest sometimes diverged from the interests of their users. Additionally, their centralized data repositories represented a high value target for attackers seeking to gain access to large quantities of valuable data. "Corporations can argue that data are trickier to manage than oil … The hacker only has to be right once to penetrate a system. Defenders have to parry every jab, all

[47] i.e. Servers and computing hardware; energy, water, real estate; legal and software development talent.
[48] "A conflict of interest inherent in any relationship where one party is expected to act in another's best interests" (Chen 2019).

the time; one misstep and they lose" (The Economist 2019). This value of these data lakes rose as Web 2.0 companies began to aggregate more and more data about their users.

A peer-to-peer electronic cash system

In October 2008 - shortly after Federal Reserve Chairman Alan Greenspan's perspicacious observation in "shocked disbelief" of a "once in a century credit tsunami" (Quinn 2008) - a pseudonymous entity called Satoshi Nakamoto published *Bitcoin: A Peer-to-Peer Electronic Cash System* (Nakamoto 2008). The paper described key features of a peer-to-peer network of computers maintaining a ledger representing electronic cash, configured to prevent double-spending. The system would not rely on trust between participants.

On 8 January 2009 Nakamoto released a C++ program implementing the functionality described in the white paper, inviting any member of the public[49] to run a node by executing the program, thereby participating in the consensus network (Nakamoto 2009). By creating this ledger (blockchain) and the system required to maintain it, Nakamoto presented created an alternative money system fulfilling the three functions of money[50] not reliant on a trusted authority.

While a description of the complexities of the operation of the Bitcoin consensus network are beyond the scope of this paper[51], broadly, it is composed of a network of computers competing to win the right to validate blocks of transactions indicating the transfer of funds from holders to recipients. The winner updates the database state and the competition starts again, based on the new state.

From the money spender's perspective, their funds are held at "wallet address" that they created[52]. This address is derived from a public key, created with a matching

[49] Really, the readers of metzdowd.com, an obscure cryptography message board.
[50] 1. A store of value. 2. A medium of exchange. 3. A unit of account. (Antonopoulos 2017b)
[51] "Bitcoin can be thought of as a state transition system, where there is a 'state' consisting of the ownership status of all existing bitcoins and a 'state transition function' that takes a state and a transaction and outputs a new state" (Buterin 2013)
[52] The bitcoin enters their wallet either by being sent there, or as a result of that wallet owner mining a new block, at which point they add a number of bitcoin to their wallet as a reward per the protocol.

private key by a key generation algorithm executed on their computer. To transmit funds to another wallet address, the user must generate a valid transaction, which includes a digital signature created with their private key. This transaction is then transmitted to the Bitcoin network.

Table 2: Example Bitcoin Private Key, Public Key and Wallet Address[53]

Private key	Hexadecimal	43c34ee9af7bfaccca6b3bd5d2af0d96bab09732aa5a3dc63a5eaa7015f2a8ce
	Binary	1000011110000110100111011101001101011110111101111111010110011001100101001101011001110111010101110100101010111100001101100101101011101010110000100101110011001010100101110100011110111000110001110100101111010101001000111100000001010111110010101010000110011100000001010111110010101010000110011101000000101011111001010101000011001110
Public key	Hexadecimal	04aca6b60b848e3bb6da4fee5b8e8be30a7acef0ed82ef82e63fe3ba5d56525729fddc63723b5a269a5facd9a5316b47da24191757d3c54e9044f29249e65f3fc4
	Binary	1001010110010100110101101100000101110000100100011100011101110110110110100100111111101110010110111000111010001011110001100001010011110101100111011110000111011100000101110111110000010111001100011111111100011101110100101110101010110010100100101011100101001111111011101110001100011011100100011101101011010001001101001101010011110111001010100111010010000010001001111001010010010010010011110011001011111100111111111000100
Wallet address[54]	Base58Check	19ZvdpcrQdS8fPQQjw7UHCvnTGBoAixL4E

[53] See *Mastering Bitcoin* Chapter 4 and Chapter 5 (Antonopoulos 2017) for a thorough accounting of Bitcoin key and wallet generation. These keys were generated using the bitcoin python library; the private key is the SHA256 hash of the string "ucl" in utf-8 encoding, 01110101 01100011 01101100 (Figure 3).

[54] A bitcoin address is "derived from the public key through the use of one-way cryptographic hashing"; it is Base58Check encoded for readability and error protection (Antonopoulos 2017 Ch 4).

Miners gather together batches of transactions into a block[55] and assembles a "block header" which, crucially, contains a hash of the previous block[56], as well as a hash representing the transactions included[57]. The miner then "repeatedly hash[es] the header of the block and a random number [nonce] with the SHA256 cryptographic algorithm" (Antonopoulos 2017). When an output hash less than a certain number[58] is found, the miner broadcasts the solution and block to the network, updating the state[59].

The system hinges on a small feature of the protocol: when a miner finds the solution nonce, entitling it to update the state, with the state update *it puts new bitcoins into a wallet that it controls*. This is how a bitcoin is minted: as a reward for participating in the competition to win the right to update the blockchain to its new state. Miners are incentivized to maintain the computing infrastructure required to sustain such a system; users are in full control of their funds. By combining this insight into human nature with the technical implementation to leverage it, Nakamoto's was "arguably one of the highest-leverage actions in human history" (Ehrsam 2017).

---

[55] Required to be smaller than 1 megabyte, the "block size".

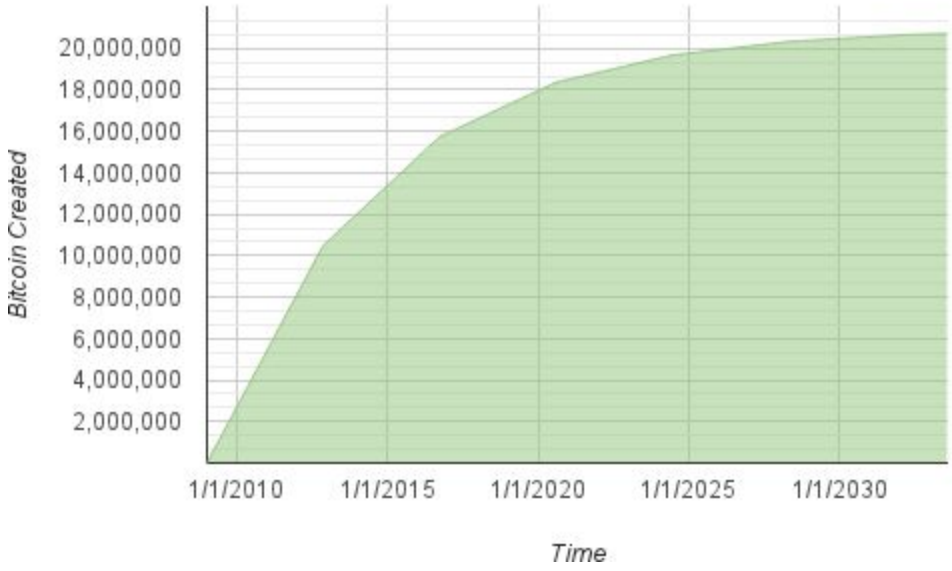
[56] This inclusion of the prior block hash, due to the properties of cryptographic hashing algorithms described, means that once a block is mined changing it would entail re-mining all subsequent blocks, including the sequence of new block hashes. This is the basis of the term "blockchain", which is, in the strictest sense, simply a database in which the current state is cryptographically linked to the prior state, providing a guarantee of immutability. The term has come to be understood, however, as the combination of "the state transition system with a consensus system in order to ensure that everyone agrees on the order of transactions" (Buterin 2013).

[57] The inclusion of this Merkle root hash further strengthens the security of the system. If a dishonest node changes a single bit (say, giving itself more bitcoin), the deviation is detected by the rest of the network, and rejected.

[58] ""Proof-of-work involves scanning for a value that when hashed, such as with SHA-256, the hash begins with a number of zero bits" (Nakamoto 2008); i.e. the output hash has a number of leading 0s - formalized as the "block difficulty". By adjusting this block difficulty the network can raise or reduce the difficulty of finding a nonce yielding an output hash, thereby self-adjusting to the amount of computing power operating on the network. The hash rate of the Bitcoin network is a measure of its security (Long 2019) meaning, sadly, that the security of the network - a strength - is proportionate to its energy consumption - a substantial drawback (Vincent 2019).

[59] This so-called Proof-of-Work consensus mechanism serves to effectively randomize which validator node wins the right to update the database state, and puts in place a substantial computational cost to change prior states. Solution nonces for all subsequent blocks would need to be computed, as changing the prior data would most likely result in the hash of subsequent block headers not fulfilling the block difficulty requirement to be accepted as valid, according to the protocol.

Figure 4: Bitcoin money supply



Every 210,000 blocks the block reward is reduced by 50%[60]; "in approximately[61] 2140, almost 2,099,999,997,690,000 satoshis, or almost 21 million bitcoin, will be issued. Thereafter, blocks will contain no new bitcoin, and miners will be rewarded solely through the transaction fees." (reprinted from Antonopoulos 2017, Ch 10)[62]

A quasi-Turing complete world computer

In 2013 Vitalik Buterin recognized that while the Bitcoin protocol provided some capability to execute scripts upon transaction validation - "a weak version of a concept of 'smart contracts'" - "the scripting language as implemented in Bitcoin [had] several important limitations", including a "lack of Turing-completeness".

Buterin's key insight was that while performing the state update required to validate transactions, arbitrary computations could be performed by the validating computer, and arbitrary data written to the ledger. He proposed Ethereum, "a blockchain with a built-in Turing-complete programming language, allowing anyone to write smart contracts and decentralized applications where they can create their

[60] According to the protocol. Block solutions broadcast to the network that violate the protocol will be rejected by the rest of the network.
[61] The exact date is uncertain due to stochasticity in block mining times.
[62] Bitcoin is inherently deflationary.

own arbitrary rules for ownership, transaction formats and state transition functions" (Buterin 2013).

With this adaptation Buterin created a system that, in this author's view, is a technological innovation representing one of the greatest opportunities to improve the dignity of the human condition in history[63]. The decentralized web - Web 3.0, or Web3 - is arising from the realization "that entrusting our information to arbitrary entities on the internet [is] fraught with danger" (Wood 2014). This movement toward decentralization has been occurring, in some ways, since the earliest days of networked computing, and even before (Wikipedia 2019e)[64].

The Ethereum Protocol

Based on a Proof-of-Work consensus mechanism similar to Bitcoin[65], the Ethereum protocol extended the functionality of the blockchain network to enable users to deploy and store executable computer programs - termed "smart contracts"[66] (Szabo 1994) - on the blockchain.

Transactions specifying the transfer of funds from one account (the sender) to the wallet address[67] of another (the recipient) were of course possible, as in Bitcoin, but compiled program bytecode could also be deployed and made available for execution in the Ethereum Virtual Machine (EVM)[68]. Upon deployment[69], the program's "contract address"[70] returned. Transactions subsequently sent to a

[63] Hyperbolic? Only time will tell ...
[64] For example, the establishment of a (necessarily) open protocol - TCP/IP - by Cerf and Kahn (1974) decentralized the ability to develop software compliant with that protocol. This sort of precompetitive coordination is of deep interest to this author.
[65] Initially - though an alternative consensus mechanism, Proof of Stake, is being developed for Ethereum 2.0 (Kim 2019).
[66] Smart contracts - "computerized transaction protocol[s] that execute the terms of a contract" - were first conceptualized by Nick Szabo (1994), who envisioned sophisticated systems of ownership and custodianship subject to conditions implementable in code. Buterin has since expressed "regret [at] adopting the term 'smart contracts'. I should have called them something more boring and technical, perhaps something like 'persistent scripts'." (@VitalikButerin 2018).
[67] Externally-owned account, sometimes EOA.
[68] A "transactional singleton machine with shared-state" (Wood 2019).
[69] Deploying a contract requires a special "contract creation transaction" to be sent to the zero address (0x0), with compiled bytecode included (Antonopoulos 2018 Ch 7).
[70] A second type of address in the Ethereum system, contract addresses reference the contracts deployed there. They do not have a private key - it "in fact does not exist - we can say that smart contract accounts own themselves" (Antonopoulos 2018 Ch 7)

contract address will result in the EVM attempting "to execute the contract, … [trying] to call the function named in the data payload of your transaction" (Antonopoulos 2018). These smart contracts were capable of computing data, as well as writing data to the blockchain[71], providing the necessary functionality for decentralized applications (dApps).

The costs of being public

As a public, permissionless service, the Ethereum system needed to disincentivize wasteful or malicious behavior. This is achieved by requiring a fee to be paid by any users seeking to transfer funds, or invoke smart contract functions. Each computational or write operation has a corresponding "gas" cost, representing the work the network would need to do to perform the operation (Wood 2019 Appendix G).

As smart contracts run, the gas costs accumulate; once program execution completes, or a limit is reached[72], the total cost of transaction validation[73] is calculated and deducted from the transaction originator's account[74]. This fee is paid in ether, the native[75] currency maintained by the system[76]. Each block has a total limit to the amount of gas that can be consumed, "to keep block propagation and processing time low, thereby allowing for a sufficiently decentralized network" (jnnk 2015, Ethereum Foundation 2019). The gas mechanism protects the Ethereum

[71]A note fundamental to the premise of this dissertation: the system is agnostic to the data being passed into these functions or written to the chain; so long as the data type is compliant, it will be accepted. It makes no judgments about the veracity of the information the data represents (which is a quality). Within the blockchain vocabulary, external entities that provide data to a blockchain are called "oracles"; these entities represent a potential weakness in the system (Fecke 2018).
[72] As specified by the transaction originator, or system block gas limit.
[73] To enable the market to determine the value of performing operations on the Ethereum Virtual Machine, and to act "as a buffer between the (volatile) price of Ethereum [sic] and the reward to miners for the work they do", account holder submitting transactions specifies the price (in a subunit of ether - "wei") they are willing to pay per unit of gas (Antonopoulos 2018 Ch 13). This allows miners, who receive transaction fees of validated blocks, to select which transactions will provide them the highest rewards.
[74] Externally-owned accounts must have some ether in them for any transactions submitted to be validated successfully.
[75] Withing smart contracts decentralized application developers can create their own cryptocurrencies, which have many of the properties of the ones described, namely, that provision of the right digital signature authorizes transfer.
[76] And the currency of the block reward for participating in the validation process.

system by making users “pay proportionately for the computational, bandwidth, and storage resources that they consume”; “thereby disincentivizing attackers” (Antonopoulos 2018 Ch 13).

Three forms of decentralization

In this research three forms of decentralization that characterize the decentralized web have been identified . These systems realize the benefits of openness and mitigate the risks through game theoretic incentivization of system participants. They seek to eradicate any “single point of failure”, instead exhibiting antifragility (Taleb 2013), the quality of being self-healing: they tend to strengthen, on balance, in response to stress, shock and volatility.

1 - Open Source

The first form of decentralization is in the authority to propose and make changes to the code and protocols. This - the open-source model of system and software development - opens access to source code to public scrutiny. In doing so, the pool of potential contributors to a project is expanded, and the likelihood of discovering unintentional errors or intentional abuses included rises[77].

Tools such as git (Torvalds 2005) - a “distributed version control system” (git-scm.com 2019) have radically improved the ability of groups of loosely-coordinated developers to work on the same codebase. In theory, an open-source project can be a pure meritocracy. Ideas - code update proposals - are judged by the community on their quality; if worthy, they are accepted and incorporated. Without a central authority directing efforts, participants are free to pursue solutions they think appropriate[78]. Openness improves security[79], diversity and adaptability.

---

[77] Linus’s Law: ”Given enough eyeballs, all bugs are shallow.” (Raymond 1999)

[78] An example demonstrating the benefit of this diversity is the development of multiple implementations of a blockchain protocol - in the Ethereum ecosystem, geth, parity, pyeth, cpp-ethereum, etc. (Antonopoulos 2018 Ch 3) - which reduces the impact of a common mode failure (Buterin 2017) in any one of those implementations, enabling a network to respond to such a failure without going offline.

[79] Of course, open source code can be examined and exploited more easily by malicious entities. It seems that on balance open source code is more secure, but counter-examples surely exist; some systems are more secure if source code remains private.

2 - Distributed Ledgers

Second is the decentralization of responsibility to maintain and update informational assets: the so-called "distributed ledger". This form of decentralization is enabled by the creation of a strict protocol enabling nodes to confirm block validity by recomputing each transaction: "each node verifies the results of each transaction" (Ryan 2017).

3 - Private Key Custodianship

Third, and most importantly, is the decentralization of the responsibility to hold the private keys needed to interact with the systems. This aspect of the decentralized web is enabled by the asymmetric key ciphers described earlier, namely, that a message recipient (in this case, a blockchain miner verifying the validity of a transaction), given a message and a public key, can mathematically confirm that the message was signed by the corresponding private key. This means that all interaction with these systems can be done by a user without ever transmitting the private key over the Internet. The only way to access the funds in a Bitcoin or Ethereum wallet, or to confirm identity as necessary in the invocation of some Ethereum smart contracts, is to present a valid transaction, including a digital signature. This signature is impossible to generate with the private key (Ker 2014, Antonopoulos 2018, Buterin 2013).

This represents a significant break from the centralized database technologies of the prior web, in which the secrets used to authenticate users (passwords) are stored on a centralized server, and must be transmitted[80] to that server for storage. In a simplified model, when a user tries to log in, the server-side software confirms that the password matches the one that it has in its records[81]; if so, the user is considered authenticated, and granted appropriate rights.

Based on Wood's (2014) observation on trusting arbitrary entities on the Internet, this model carries risks: the authorities managing a user's data might mistakenly delete or alter it[82], reveal it to some malicious attacker[83], or may choose to deny

[80] Albeit on an encrypted channel like HTTPS - sometimes (Vyas 2016)
[81] Hopefully stored securely … though sometimes, even in the case of highly skilled and resourced firms, not (Krebs 2019).
[82] as in the cases of DreamHost, several government agencies …
[83] … JPMorgan Chase, Equifax …

access or revoke service provision[84] (McCoy 2015, Haselton 2017, Hopkins 2017, Fernandez 2019, Galperin 2015, Prince 2019)[85].

Technologies enabled

Decentralized Autonomous Organizations

Smart contracts enable a broad range of decentralized applications to be developed and deployed on a blockchain network. Of particular interest is that of the DAO: "decentralized autonomous organization", an organizational structure enabled by smart contract technologies. As with many of these concepts, the collective understanding of DAOs is nascent and few are operating successfully at the time of writing, but generally these entities operate according to rules and data encoded on a blockchain. When deployed on public blockchains, total transparency as to the governance (Merkle 2016)[86], data, membership and activity of the DAO is available for public review[87].

Due to the versatility of smart contracts these rules can enforce any policy, but DAOs can be conceived of as providing members an interface through which they can interact with the informational and financial assets necessary to coordinate organizational projects[88]. At this early stage DAOs appear to hold enormous promise to improve organizational efficiency, thereby enabling the provision of services to unserved market segments[89]. They also hold the potential to enable as of yet unforeseen organizational structures (Aragon 2019, DAOstack 2019, Rea 2019).

---

[84] … Facebook, and Cloudflare, to name a few.
[85] While a topic of intense fascination, an in-depth discussion of the ethics of the denial of access to informational services is beyond the scope of this dissertation. Web3 purists seem to err toward a complete rejection of the morality of authorities deciding to prevent behavior they deem harmful. This author takes a more nuanced view, but overall would rather see those (dis)incentivization mechanisms (such as censorship, sanctions and so on) built atop a system that is fundamentally oriented toward individual liberties and rights.
[86] i.e. smart contract source code.
[87] Of course, some organizations require privacy for ethical, legal or business reasons. Due to its constraints, this paper cannot thoroughly explore the mechanisms and ethics of organizational privacy, nor mechanisms for preserving privacy on public blockchains.
[88] A common use case would be the collection of funds and collective decision-making about their disbursement, based on some voting mechanism (DAOstack 2019).
[89] For example, IBISA is a project building a decentralized autonomous organization offering crop insurance to smallholder farmers in developing countries. Their customers are unserved by traditional insurers due to the high costs of offering products relative to the low revenue opportunity the customers represent (Bitvalley 2018).

Private Blockchains

Public, permissionless blockchains incentivize members of the public to utilize or contribute the computing resources necessary to sustain the system; to prevent abuse high fees are charged for use of compute and data storage resources. The community develops the protocol and software and decides on which features to implement. Furthermore, all data - including accounts and balances - on a public blockchain can be accessed by anyone with an Internet connection[90].

For use cases where system developers are not willing to sacrifice control and privacy, blockchain protocols and clients designed to be deployed on private, permissioned networks exist, such as Hyperledger (The Linux Foundation 2019), Corda (2019) and Quorum (2019).

If configured properly - say, with validating nodes hosted by competing firms, or loosely coordinated firms situated across a value or supply chain, where each would benefit by the existence of some shared database - private blockchains can provide many of the benefits conferred by decentralization[91] without requiring the participants to rely on public networks. Furthermore, operating costs can be substantially reduced if network participants trust each other to some degree, or have legal mechanisms in place to disincentivize malicious behavior.

Ricardian contracts

Blockchains and the cryptocurrencies and smart contract systems they enable have emerged independent of the direction of any sovereign government, regulatory agency or monetary authority. While this feature attracts some, others identify an opportunity to utilize smart contracts within the existing regulatory and legal framework.

First proposed by Grigg (2004), Ricardian contracts contain both human- and machine-readable elements, a legal (prose) and smart (code) contract. This allows them to document the intent of the agreeing parties, be interpreted in a court of

[90] It should be noted privacy preserving techniques on public blockchains are in development.

[91] Persistence, immutability, security, among others. Web3 purists may disagree with this point.

law, and delegate certain functions to smart contract platforms (Braendgaard 2016, Clack 2016).

Zero-knowledge proofs

Zero-knowledge proofs provide a verifier an assurance that prover knows a piece of information without revealing anything more. It is a privacy-preserving method with applications in authentication systems, identity applications, nuclear disarmament (Glaser 2014), and has been proposed as a privacy-preserving mechanism on public blockchains (Orcutt 2017, Wu 2014, Wikipedia Contributors 2019f).

Content Addressing

In location-addressed systems, resources are referenced by their address, often a URL referencing a location in a web server's directory structure, or containing parameters for a database query enabling the server to respond with the requested data.

This, however, represents a single point of failure[92]: if a file is moved from its location, the URL does not resolve and the web server returns a "404 Not Found" error; this resource is inaccessible. In the serverless Web3 paradigm[93], this problem is solved by addressing data by its content. The technology is built largely on hashing algorithms and their ability to prove the "sameness" of two equal-length binary sequences. Files are identified by a unique and deterministic hash; a user requesting a specific file could receive segments from various peers, re-assembling it on their local machine[94].

Juan Benet (2014) proposed the InterPlanetary File System (IPFS), "a peer-to-peer distributed file system" based on content addressing. The use of content addresses enable short hashes[95] to reference larger files[96]. Due to the cost of writing to a blockchain, and integrity guarantees, IPFS and other content addressing systems

[92] Commonly experienced by this author.
[93] In the spirit of removing single points of failure, and improving performance.
[94] BitTorrent (Cohen 2003) forms some of the basis of Benet's solution.
[95] In IPFS, v0 Content Identifiers are 46 bytes in length (Protocol Labs 2019)
[96] And, relatedly, "*Proof-of-Replication* was developed to solve the specific problems of a verifiable, decentralized storage network that could incentivize and reward file storage", thereby improving on the fragility of the location-addressed client-server model (Protocol Labs 2019b).

and decentralized file storage systems are technologies crucial to the decentralized web.

Self-Sovereign Identities

> "ed: The vulnerability that is being exploited in all systems is identity."
>
> @santisiri, 2019

In the emerging vision of the decentralized web, digital identities are controlled and owned by the individuals and organizations they represent: "users [are] the rulers of their own identity". While the community's understanding of the idea is nascent and developing, Allen defined 10 principles of self-sovereign identity (2016), oriented toward individual autonomy and privacy (Appendix 3). These developments are enabled by the advancements in networked computing and cryptography described - self-sovereign identities rely on users being able to create and manage private keys securely.

Sovereign Sensors

Humans interact with the Internet through computers, which detect inputs and establish connections with other computing nodes. In the most literal sense, every personal computer is a sensor: they "sense" keyboard or touchscreen inputs, as well as incoming transmissions from other computers, either wirelessly via an antenna or through some wired input. As such, for the purposes of this research the definition of "connected sensors" includes any digital computing node with the ability to receive and transmit data.

However, a distinction should be made by computers that are controlled by direct human interaction[97] and ones that are autonomously controlled by software agents installed by human developers[98]. This research was inspired by a peculiar question: what if these autonomous agents had self-sovereign identities of their own? What if

[97] Desktops, laptops, tablets and smartphones, generally.
[98] To this author's knowledge no computer exists that is controlled entirely by software that was developed and deployed by another computer, though it is not unfathomable (Boyd-Rice 2018). It seems obvious, however, that all future software and hardware technology ultimately originated in a human action.

computers had private keys that only they had access to? How might such sensors serve society, and how might they threaten its functioning?

Such a capability has been an area of active research for some years, primarily led by the Trusted Computing Group, including "AMD, Hewlett-Packard, IBM, Intel and Microsoft" (Merritt 2003). Detailed discussion of the technical aspects of trusted computing is well beyond the scope of this paper[99], but the fundamental premise is of a chip that contains in non-volatile memory a "public and private key pair, … created randomly on the chip at manufacture time [that] cannot be changed" (Safford 2003) and cannot be accessed by any external entity[100]. With this technology, it seems that computers have a right to privacy[101].

With this technology, ethical concerns abound (Anderson 2003), but if configured properly and governed transparently, such a system could offer myriad ways to improve the security in the emerging Internet[102], and the ability for computer networks to monitor environmental, physical and information security conditions.

The notion of computers in sole and secure possession of private keys attains a new level of significance in the context of the decentralized web[103]. As sole controllers of wallet addresses, edge devices can be used as oracles in smart contracts without any point of human interference, and a computer can hold funds accessible only to

[99] As well as the cognitive capacity of this author.
[100] If the device securely maintains custody of the private key (i.e. it has the only instance of that private key in existence), all digital signatures with that key must have been performed on the device.
[101] Should they?
[102] Security and privacy are often viewed as a trade-off (Malik 2018): increased surveillance erodes privacy but increases the likelihood of authorities detecting threats to public safety. Acknowledging the constraints entailed by the computational intensity of homomorphic cryptography, this author is still wondering: could homomorphic algorithms analyze data about some entity to detect a threat, and only reveal the contents of the data to authorities if a threat is detected? Could software agents analyzing data about human activity be configured as a sort of benevolent panopticon - an artificial intelligence oriented toward minimizing harm and damage to life and property?
[103] In the Ethereum Yellow Paper, Wood recognized the potential of machines providing inputs to smart contracts: "A transaction (formally, T) is a single cryptographically-signed instruction constructed by an actor externally to the scope of Ethereum. While it is assumed that the ultimate external actor will be human in nature, software tools will be used in its construction and dissemination. [Footnote 1] Notably, such 'tools' could ultimately become so causally removed from their human-based initiation—or humans may become so causally-neutral—that there could be a point at which they rightly be considered autonomous agents" (Wood 2019 pp 4).

it. A vehicle could detect necessary on-board maintenance - and release funds to perform it. A weather buoy could submit data to a smart contract - data that could only have originated on the buoy. Access to the imagery captured by a satellite with a synthetic aperture radar or a surveillance camera installed in an urban environment or a balloon (Harris 2019)[104] could be managed by a DAO governed by an open and transparent set of rules, subject to identity-based, spatial and temporal conditions.

2.3 On resource governance

"Governance refers to all processes of social organization and social coordination" (Bevir 2012). In this context - of computers - governance is closely associated with the concept of access control, the processes by which access to and use of the informational resources on a computer system is managed (Rouse 2019). It also refers to decision-making about the direction of blockchain software and protocol development (Ehrsam 2017, Zamfir 2017)

A thorough exploration of the theory and practice of governance is beyond the scope of this work. Broadly, however, if governance is conceived of as "processes of rule" (Bevir 2012), the principles underlying governance design, as well as the structures manifesting such processes[105] hold great bearing on the justice of the system: its efficacy in equitably governing the subject community.

[104] "'We do not think that American cities should be subject to wide-area surveillance in which every vehicle could be tracked wherever they go,' said Jay Stanley, a senior policy analyst at the American Civil Liberties Union." Leaving aside the apparent inevitability of this, this author believes that such a capability is not inherent unethical. If such information is stored, analyzed and accessed by an open source system designed to respect the individual's privacy, with appropriate anonymization, it is believed that such a system has the potential to fundamentally improve physical security in the service of the public interest. The key phrase: "designed to respect the individual's privacy"; likely incorporating zero-knowledge and range proofs, as well as homomorphic cryptographic techniques, etc. If "the surveillance state is inevitable" (Weigert 2015), then the nature of the state must be adapted so human rights are respected.

[105] Government, as in "the system or group of people governing an organized community" (Wikipedia 2019g). Note that this conception of government is much broader than the common state-based understanding of the term; it is "a means by which organizational policies are enforced, as well as a mechanism for determining policy" (Wikipedia 2019g).

The governance of the physical commons is fairly well understood, most clearly conveyed in the eight "design principles" of common resource governance identified by Ostrom (1990). However, these principles seem to primarily apply to the governance of physical resources; no clear consensus on the nature and intricacies of informational resource governance appears to exist. On one hand, information is "a good - … an object of economic transactions", "typically non-rival and sometimes nonexcludable"[106] (Varian 1998). On the other, "numerous English authorities have affirmed that information or data is not property" (Bilbow 2019)[107].

It seems that informational objects exist in a space with different laws to physical space. Given the commitment to investigating the factors pertaining to the governance of informational resources, the differences between these two spaces in which reality manifests is of acute importance to the research agenda. This will be explored further In the Conclusion.

---

[106] This author is critical of the general economic definition of "non-rivalrous", which forms the basis of the distinction between public goods and common goods (Vu 2015). The difference is appreciated: the use of a rivalrous good decreases the value of the asset, thereby making that value unavailable to future users (i.e. when a tree is cut down and lumber removed no future person can use that specific tree for lumber), whereas the utility of non-rivalrous goods is not diminished with use (i.e. a waterway can be used to transport a shipment without reducing the waterway's utility for future users). The point missed within the definitions reviewed is the temporal rivalry of many public goods: *while I am using this waterway*, availability to others is reduced. Informational resources, especially in the decentralized, content-addressed web, may (almost) be the first non-excludable resources in history (as even high-bandwidth servers have a limit to the number of users who can access an informational resource in a given time). True non-excludability may be impossible - except perhaps in the case of an informational asset in the custody of its user.

[107] "can data be owned?" (@santisiri 2019b)

3 Methodology

3.1 Overview

3.1.1 Purpose

An agent-based approach was employed to investigate the dynamics of sensor networks recording data on a public blockchain. As the scale, resolution and connectivity of the Internet of Things grows, so does its potential to improve situational awareness (Shirer 2019). If leveraged properly, enormous gains in the efficiency and prevention of harmful or malicious activity might be realized in the networks monitored by providing system operators a more accurate understanding of systemic patterns and dynamics - with substantial social, environmental, military and commercial implications.

As outlined, public blockchains have unique limitations when compared with traditional database management systems, cloud servers and private distributed ledger implementations. This modeling effort was conducted to explore emergent dynamics of connected sensor networks reporting measurements from the edge to a public blockchain. By understanding these dynamics and the trade-offs inherent to using public and permissionless blockchain architectures, these findings might contribute to helping system designers to take a more informed approach to architecting such systems, thereby balancing the costs and benefits of decentralization.

This model is focused on investigating factors related to network costs and transaction processing times. These scaling questions are important in the context of connected sensor networks, which are often characterized by high data capture and transmission volumes, and provide the most value to system operators if the data captured at the edge is made available for analysis rapidly and is available long term (Gutierrez 2015).

3.1.2 State variables and scales

Using Python 3's (van Rossum 1995) Mesa framework (Core Mesa Team 2019), two subclasses inheriting Mesa's `Agent` class were defined: `Sensor` and `Blockchain`. A `SensorBlockchainNetwork` subclass of Mesa's `Model` class served to instantiate an individual simulation, configure model schedulers and collect data from model reporters. Parameter sweeps were performed using Mesa's `BatchRunner` tool in an iPython notebook (Pérez 2007); collected data was analyzed and visualized, also with iPython, using the `numpy` (Oliphant 2007), `pandas` (McKinney 2010), `statsmodels` (Seabold 2010) and `matplotlib` (Hunter 2007) packages.

3.1.2a Sensors

Sensor agents represented edge devices capable of recording empirical data about their surrounding environment[108]. Logic designed to simulate on-device data computation and transmission operations was included within the class definition as methods.

[108] The type of recording is abstracted; image sound, image, temperature, air quality, rotation, acceleration, etc.

Table 3: Sensor agent elementary properties

Name	Value	Initialized Parameter
`battery_life`	Energy level	`10000`
`mortal`	Boolean, whether sensor should be removed from scheduler upon battery depletion	`False`
`dead`	NaN, replaced with block number on battery depletion (if mortal)	`NaN`
`record_cost`	Energy consumed with each recording	`1`
`gas_price`	Price (gwei[109]) per unit of gas, required in transaction (Ryan 2017)	`20`
`stochasticity`	Arbitrarily introduced variation to mimic reality	`0.05`
`gwei_spent`	Gwei spent that tick	`0`
`last_sync`	Block number of previous transaction submission	`0`
`nonce`	Iterator for valid transaction generation	`0`
`db`	Block-indexed array with bytes recorded each tick	Empty array

3.1.2b Blockchains

`Blockchain` agents were designed to simulate key features of the quasi-Turing complete public blockchain networks described, with the Ethereum platform serving as the primary inspiration. The `Blockchain` class defined for the model here is based on key features of the Ethereum platform[110] (Antonopoulos 2018).

[109] Gwei stands for gigawei; 1 wei = 10^{-18} ether, 1 gwei = 10^{-9} ether = 1 nanoether (gwei.io 2019). Think pence to pounds or cents to dollars. It must be noted that the term "gwei" is simply borrowed from the Ethereum lexicon; here it represents a theoretical cost, and should not be thought of in terms of value in fiat currency. Rather than attempt to estimate true costs, this dimension was modeled in an effort to understand trends in total transaction costs per sensor across the parameters swept.

[110] Of note: the `Blockchain` class does not have a `step()` method, nor was it added to the Mesa scheduler. This was intentional, to emulate the inherent reactivity of smart contracts executed on blockchain networks. All smart contract computations can be traced back to an origin externally-generated transaction sent to a contract address. Tasks cannot be scheduled on Ethereum (Antonopolous 2018 Ch 13).

In the `Blockchain` class (which inherits `Agent)`, attributes including the gas limit per block in gas and gas cost per byte written in gwei, were defined upon initiation. These integer values were fixed both through a single model run[111] and across model runs.

Class methods defined allowed for the blockchain network to perform necessary actions, including adding submitted transactions to its mempool (`Blockchain.add_to_mempool(tx)`) and mining a block (`Blockchain.mine_block()`).

The latter method included logic to select the highest value and most recent transactions from the mempool[112], validate them - including invocation of that appropriate sensor's `Sensor.confirm_tx(tx)` method - and update `Blockchain.chain` dataframe's state.[113]

[111] Note that this simplifies this dynamic when compared with Ethereum, in which block gas limits are adjustable (Antonopoulos 2018 Ch 13). While the dynamic nature of these parameters are critical to the adaptability and scalability of blockchain systems, these models focused on variation in the size and behavior of the networks connecting to the blockchain in each model run.

[112] Thereby simulating an economically rational miner's selection of transactions from the mempool, designed to maximize block rewards.

[113] The `Sensor.confirm_tx(tx)` method was designed to emulate the response sent from the blockchain to the client upon transaction confirmation. This aspect of smart contract behavior is the reason for the asynchronous functionality of the web3.js library, in which JavaScript's async await syntax is employed to pause execution of client-side code reliant on transaction confirmation and, possibly, values returned from smart contract computations (web3.js 2019). If edge devices were running NodeJS programs (Node.js Foundation 2018) to execute on-device computations and transmit data, the web3.js library would likely be used to interact with the Ethereum blockchain. This would be done by establishing a connection to the network (local, test or main net) and invoking methods of the contract instance, referenced by the contract address. This would be done in a local, private environment as one would need to sign the transaction with the private key of the wallet containing the ether necessary to pay for the gas costs.

Table 4: Blockchain agent elementary properties

Name	Value	Parameter / Range
`block_gas_limit`	Block limit of computation or data write volumes in gas	`8000000`[114]
`gas_per_byte`	The gas cost to write one byte of data to the chain	`625`[115]
`chain`	A dataframe of boolean values representing whether sensor recordings have been validated on-chain	Empty dataframe
`tx_ct`	An iterator so each transaction submitted gets a unique id	`0`
`mempool`	A dataframe containing transactions, including 'mined' boolean column	Empty dataframe

In the middle-range model developed (Gilbert 2008 pp 42) blockchain complexity is limited to write operations: no on-chain compute operations were simulated. This notable exclusion of a key feature of smart contract platforms was deemed necessary to reduce model complexity[116].

3.1.2c Higher-level entities

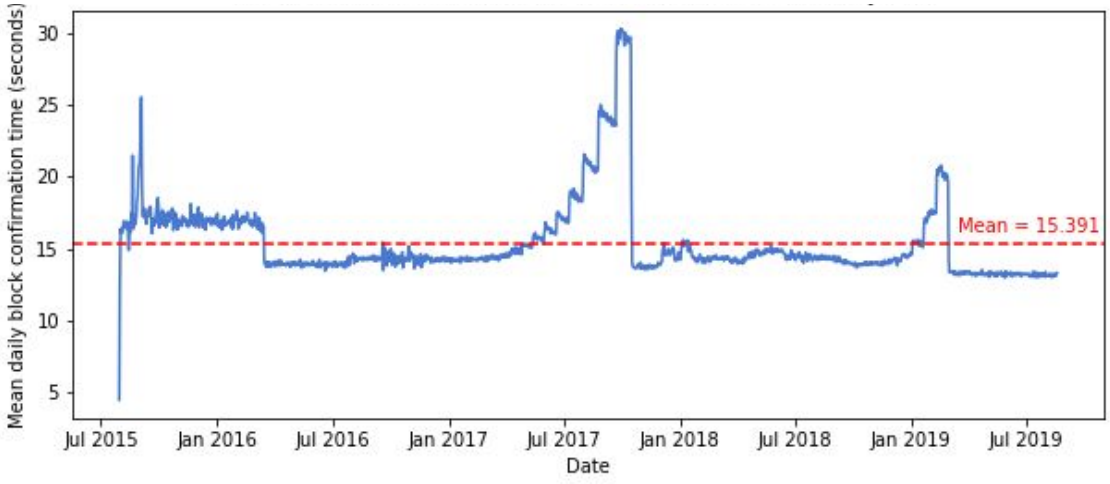
No higher-level groupings of agents exist, though much opportunity to model heterogeneous networks exists within the modeling framework developed.

3.1.2d Scales

The physical location of model agents is not specified. A single tick in the model represents a block confirmation. For context, Figure 5 depicts block times on the Ethereum mainnet since network launch: 15.391 seconds, on average.

[114] (Moriya 2018).
[115] (Ryan 2017).
[116] Methods were included in the `` `Blockchain` `` subclass to simulate on-chain compute operations, namely, gas costs.

Figure 5: Ethereum mainnet block confirmation times



(Etherscan.io 2019)

Model runs were executed for 300 ticks[117]. Scales were selected based on a series of initial model runs. Visual inspection of the results of these model runs confirmed the intuition that network block confirmation times and informational currency measures would stabilize around the optimal values of 1 block and 1.0 respectively. Increasing network loads through adjustment of record volume, record frequency and network size parameters resulted in these measurements degrading rapidly at a tipping point, representing the network reaching some operational performance threshold. Fixed parameters, and the ranges of swept parameters, were chosen to enable investigation into network dynamics around this threshold.

However, development of a facsimile model exactly simulating each of these aspects was well beyond the scope possible given time, computational and analytical constraints. While such a model would be difficult to build, it could provide researchers the opportunity to simulate unanticipated stresses on blockchain networks, as well as the emergent complexities of on-chain interactions such as DAO behavior, voting and governance mechanism performance high-demand periods such as token listings (Peaster 2018), dApp launches (BBC 2017), asset ownership transfers (Gupta 2019), and so on.

[117] This substantially but necessarily abstracts a blockchain network's complexity: the value of such a platform is largely found in its persistent operation.

3.1.3 Process overview and scheduling

To model the operation of a network of sensors connected to a blockchain, key steps in the process were defined for simulation, while others were abandoned in abstraction. Judging the relevance of features of a complex system is a primary point where bias can be injected into such an enquiry; good faith efforts were made to retain important aspects.

Actions can be conceived of as taking place in two computing environments: on board the edge sensor and on the computing nodes forming the blockchain's consensus network, where transactions are validated and the state is updated, including execution of any smart contract code. The data follows a serial process (Table 5).

Table 5: The path of data in the network simulation

Step	Action
1	The edge device takes an empirical recording from the edge and stores it in on board memory, such as an SD card.
2	When ready to transmit data to the blockchain, the edge device gathers all un-submitted data[118], includes it in a transaction, signs the transaction and transmits it to the blockchain network.
3	Upon receipt of the transaction data by the miners, the transaction is added to the mempool, a local data store where transactions await validation.
4	Each block (tick), the network mines the next block by selecting a set of transactions requiring less gas to execute than the block gas limit. Upon validation, the blockchain updates its state to reflect the inclusion of the data (edge recordings) contained in those transactions.
5	A message is transmitted back to each edge sensor confirming transaction validation. This enables sensors to keep a record of the financial expenditures required to perform write operations[119].

To simulate the lack of coordination in reality amongst edge nodes, `Sensor` agents were activated in random order. The tick was completed upon block confirmation, at which point the next transactions in the mempool move up the queue; for this reason the final action taken in each model step was the invocation of the `Blockchain.mine_block()` method.

Each tick represented one block validation by the blockchain network, during which the blockchain's state was updated to incorporate data included in each transaction mined.

[118] i.e. data recorded since the prior transaction submission.
[119] Note: This is not an accurate representation of reality: in Ethereum the accounts and balances are stored on the blockchain itself, and when an externally-owned account's transaction is validated the transaction costs are deducted from that account's balance. No communication is required with the submitting edge node to complete transaction validation. However, this structure was used for conceptual simplicity, meaning the Sensor objects kept track of expenditures - relevant for Mesa's reporting procedures. This meant that costs were measured by adding costs incurred to the `Sensor.gwei_spent` value, rather than seeding sensor-controlled accounts with currency and subtracting from it each transaction validation, as would happen if such a system were actually developed for the Ethereum network.

3.2 Design concepts

3.2.1 Sensing

As informational entities, the agents as defined and simulated sense their environment in a few ways. Edge nodes convert the qualities of the analogue data contacting their sensory interface into some quantitative representation of that quality, represented as binary information[120]. The moment of this analogue-to-digital (qualitative-to-quantitative) conversion is explored more thoroughly in Appendix 7, as understanding it is critical, in this author's view, to understanding the nature of information and thus the governance of informational resources. However, this is largely abstracted in this model, with only the "volume" of data captured in bytes being explicitly referenced in model execution.

All other agent "sensing" pertains to the detection of signals encoded with binary information, carried on some channel. Specifically, these include the detection of a transaction submitted to the blockchain network[121], and the subsequent detection of the message from `Blockchain` to `Sensor` confirming transaction validation[122].

3.2.2 Interaction

Agents interact by transmitting data to a recipient - `Sensor` agents submitting transactions to `Blockchain` agents, and `Blockchain` agents sending confirmation of transaction validation to `Sensor` agents. Exploration of peer-to-peer data transfer systems - in which interactions occur between `Sensor` agents and between different `Blockchain` agents - holds potential to improve system scalability, but was beyond the scope of this research.

3.2.3 Stochasticity

A degree of stochasticity was included at some points in the model design to simulate variations arising in system functioning. `Sensor` agents used the

[120] A number, array, array of arrays, and so on.
[121] via invocation of `Blockchain.add_to_mempool(tx)` within the `Sensor.transmit()` method body.
[122] `Sensor.confirm_tx(tx)`, called within the `Blockchain.mine_block()` method. Again, the specifics of this process are almost entirely abstracted. This was necessary to conduct this study within its constraints, but the dynamics of information transfer and dissemination through these networks of computing nodes is very far from irrelevant.

stochasticity measure to introduce variability in the number of bytes captured when a new recording was taken. Additionally, if instance variables `Sensor.record_freq` or `Sensor.transmit_freq` were assigned float values between 0 and 1.0 exclusive (rather than unsigned integers), a probabilistic conditional test was applied before the `Sensor.record()` or `Sensor.transmit()` operation was executed. This was meant to inject variability in the frequency of these actions taking place, simulating some on-device logic or intermittent Internet connectivity. The degree of randomness is quite system-dependent; if a more accurate model were being built a more thorough consideration of the process and measure of stochasticity introduced would be necessary.

3.2.4 Collectives

No higher-level groupings of agents were simulated in this version of the agent-based model. However, much about system performance could relate to these collectives and their interaction. Indicating the owners or manufacturers of devices might enable simulation of changes affecting subsets of the network such as a corporate decision-maker - or a hacker exploiting a software vulnerability - removing a number of devices from participation in the network, or changing their behavior somehow. This is of special concern if edge devices are in custody of private keys controlling financial assets, trusted hardware, as outlined. A malicious actor co-opting a botnet of such sovereign devices could carry significant financial consequences (Sabanal 2016).

3.2.5 Observation

Upon completion of all agent activation, on each tick agent-level data was collected for further analysis.

The mean gwei spent per sensor was calculated for each model run to gain insight into the effects of network size on financial costs to system participants. Transaction costs disincentivize malicious or wasteful behavior; the higher marginal costs represent a primary way public blockchains differ from private implementations and traditional data storage and cloud computing resources.

A metric representing the proportion of information captured that has been submitted, validated and represented on-chain was developed, termed "informational currency". [123] The metric based on a rolling window of the most recent 30 blocks mined, $IC = N_{records\ mined} / N_{records}$.[124]

Additionally, summary statistics describing transaction mining times was collected upon termination of the model run and a mean mining time was calculated by averaging the difference between block submitted and block mined for all transactions.

3.3 Details

3.3.1 Initialization

Upon initialization, a `Blockchain` object is instantiated with an empty mempool. A specified number of `Sensor` agents are also instantiated with instance variables specified, including reference to the `Blockchain` instance present. Sensors are added to the model scheduler. Swept variable parameter ranges are shown in Table 6.

Table 6: Initialization variables - swept parameters

Object class	**Variable**	**Parameter value or range**
`SensorBlockchainNetwork` (`Model` subclass)	`num_sensors`	{s: s = n * 50, n ∈ [1, 2, ... 10] }
`Sensor` (`Agent` subclass)	`record_freq`	[0.01, 0.1, 0.3, 0.5, 0.7, 0.9, 1]
	`record_bytes`	{r: r = 1 + n * 20, n ∈ [0, 1, 2, ... 25] }

[123] Calculation of mean informational currency for each iteration excluded the warm-up period of 30 ticks. The metric quantified the proportion of data captured at the edge reflected on chain, having completed the full process of recording the data, transmitting a transaction, and that transaction being validated through mining.
[124] Note that record size was not incorporated; more sophisticated informational currency metrics will be necessary if heterogeneity in edge recording dynamics is simulated.

3.3.2 Input

No environmental inputs were introduced during model execution; all activity was in response to pre-programmed agent behaviors and state variables as initialized.

To move toward a facsimile model of connected sensor networks, the programming of responsiveness of sensors to environmental conditions represents an interesting line of further research[125].

3.3.3 Submodels

See Appendix 1 for details of submodels.

Models, data and iPython notebooks can be found at https://github.com/robisoniv/sovereign-sensors.

---

[125] For example, some sensor monitoring coral reef health could transmit a transaction in the event of reef damage, triggering the release of funds to some local repair crew. This type of parametric insurance is being considered by the insurance industry; Willis Towers Watson recently launched a reef insurance project for sites in central America (Vincent 2018). It is unknown, however, if such parametric policies are currently connected to smart contracts holding funds for disbursement, or how parametric triggers are monitored and policies paid in the event of conditional thresholds being exceeded. A recent Lloyd's of London report suggests it is unlikely that such a system has been connected to smart contracts on a blockchain mainnet holding real funds (Lloyd's 2019).

4 Results

Analysis of data captured during model runs revealed interesting system dynamics, though questions remain regarding whether the patterns observed reflect system operation in reality.

This investigation sought insight into the effects of three independent variables on three dependent variables. Interesting and unexpected results will be reported; all other results are included in Appendix 2.

4.1 Network size

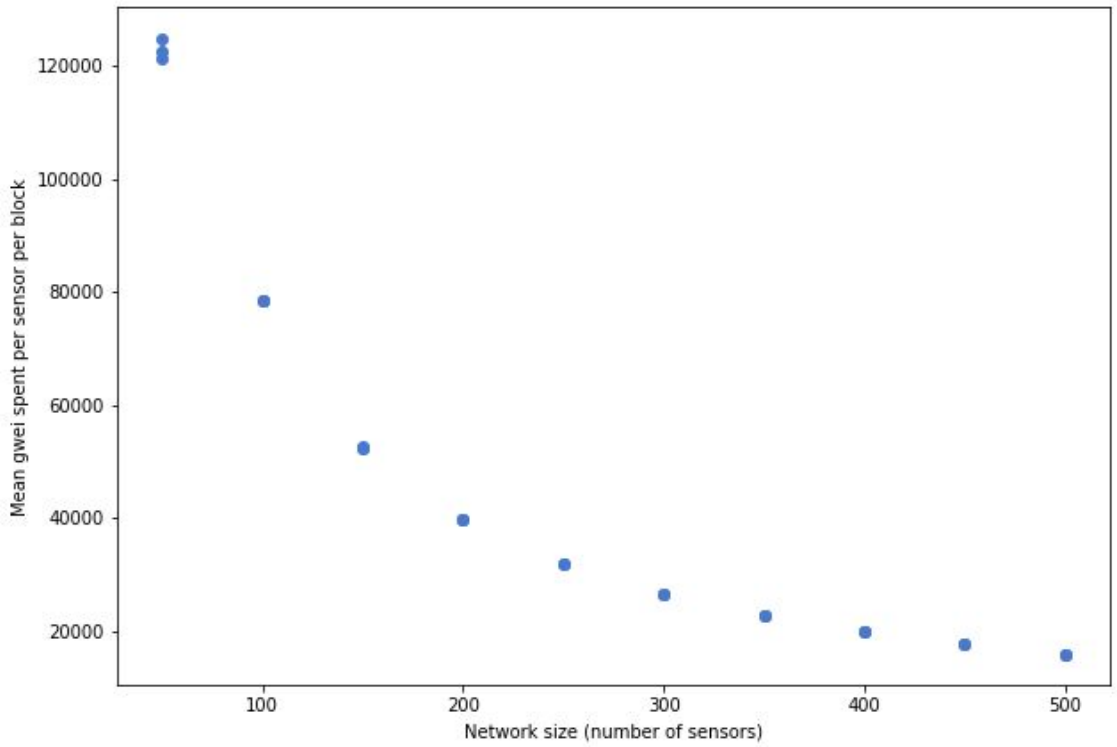
The network size was defined as the number of sensors submitting transactions to a blockchain. Model simulations were executed ranging from 50 to 500 sensors, on intervals of 50.

4.1.1 Gwei spent

Table 8: Summary statistics, Mean total gwei cost per sensor across network size parameter sweep

n iterations		30
Total gwei cost per sensor	Mean	10.44038
	σ	0.64924
	Minimum	9.67701
	Median	10.27899
	Maximum	11.73449

Figure 8: Network size against mean gwei spent per sensor per block



In Figure 8, a concave negative non-linear relationship between network size and mean gwei spent per transaction is visible.

A plot of a log transformation of both variables yielded an apparent negative linear relationship. An OLS simple linear regression analysis produced a model with an R-squared value of 0.9938, indicating that the network size strongly predicts mean gwei spent per sensor. Figure 9 visualizes the scatter plot of the transformation and the calculated line of best fit.

Figure 9: Log of network size against mean gwei spent per sensor per block

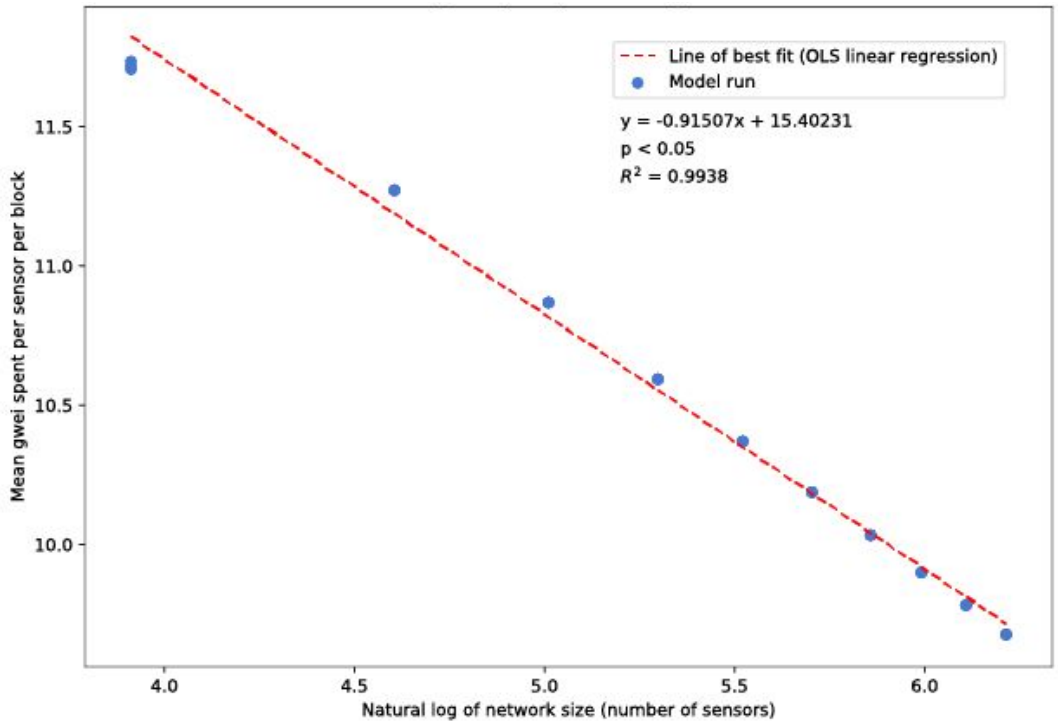
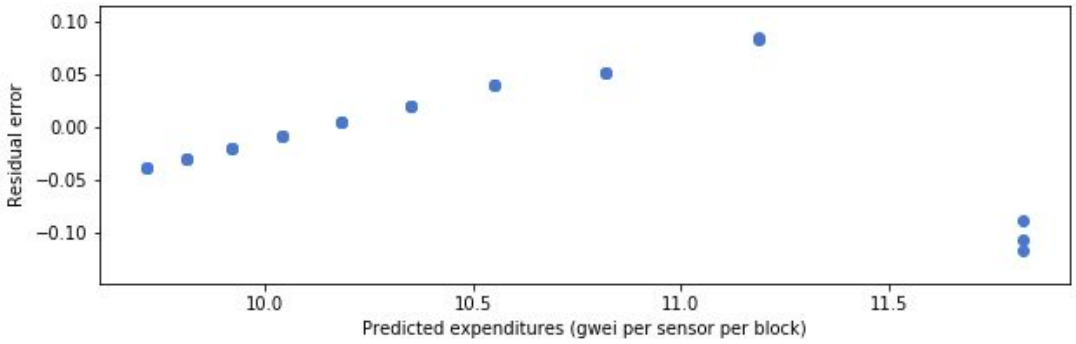


Figure 10: Residual errors versus gwei spent per sensor per block predicted from network size



Visual inspection of the residual errors plotted against predicted values based on the model (Figure 10), however, suggests that the errors are not normally distributed, invalidating the OLS simple linear regression as a method for assessing the significance of the relationship between the variables. Still, the negative relationship between the variables is irrefutable.

As with mining dynamics, the decrease in gwei spent per sensor per block observed with increasing network sizes is caused by the static block gas limit. As the number

of sensors attempting to submit data to be stored on chain increases, the likelihood of each sensor's transactions being validated, contained data being written, and gas costs incurred, decreases. Costs in gwei are only deducted from the sensor's externally owned account upon transaction validation.

Oddly, in model runs with networks larger than 150 sensors, mean gwei costs per sensor were the same across the three iterations executed at each network size simulated, while some, albeit minor, variation was observed in smaller networks (Table 9). Due to the inclusion of the stochasticity metric, as well as a probabilistic record frequency value, this was not expected. Regardless, it almost certainly is a quirk of model design and not an indicator of actual system behavior.

Table 9: Standard deviation of mean gwei spent per sensor per block across identical model runs

Network size (sensors)	**σ - mean gwei spent / sensor / block**
50	0.014425
100	0.000423
150	0.000339
200	0.0000
250	0.0000
300	0.0000
350	0.0000
400	0.0000
450	0.0000
500	0.0000

4.1.2 Informational currency

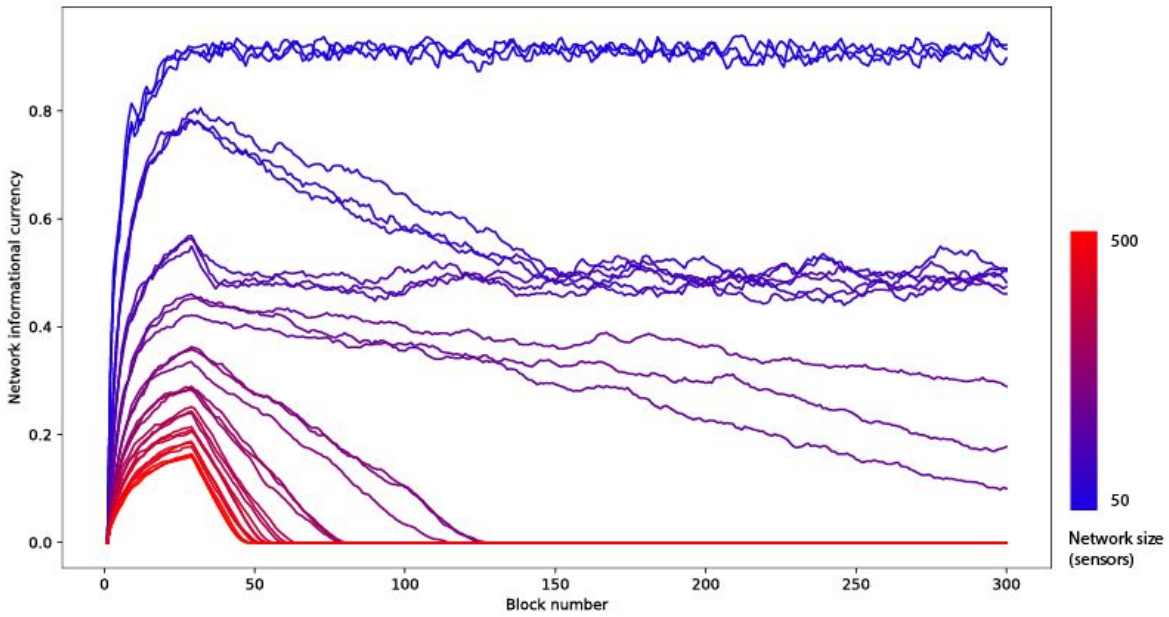
Interesting and unexpected patterns of network informational currency were observed across the parameter sweeps - and across the time series of individual model runs. A linear negative correlation between network size and mean

informational currency values was expected, as was stationarity within model run time series plots. While the pattern expected was observed across model runs, this was not the case within individual iterations.

4.1.2a Time Series

Figure 11 depicts time series plots of mean network informational currency by tick for each of the model iterations recorded, colored by network size. In the smallest networks (50 sensors), informational currency rapidly rose in early ticks and - prior to tick 30, at which point the rolling window meant older blocks were excluded from the measure calculations - reached a maximum value of > 0.9.

Figure 11: Informational currency over time through network sizes



After the warm-up period for these 50-sensor networks, mean network informational currency was a stationary process, as Augmented Dickey-Fuller test results show (Table 10).

Table 10: Augmented Dickey-Fuller test results, mean informational currency over time for 50-sensor networks

Iteration	ADF Statistic	p-value
1	-7.2910	< 0.05

2	-5.3041	< 0.05
2	-6.3199	< 0.05

p-values < 0.05 indicate it is unlikely that these time series do not have a unit root; they are stationary.

Networks larger than 50 sensors exhibited unusual behavior: after achieving a maximum mean informational currency after warm-up, the measures decayed over time until a threshold was reached, at which point the process became stationary. The greater the maximum measured value, the slower the decay rates to the floor threshold. Specifically, ~0.5 appeared to serve as a threshold support level; any networks that exceeded this mean currency receded to this point over time, then achieved stationarity there. Networks of 100 and 150 sensors exhibited this behavior. In networks of 250 or more sensors, which did not reach an informational currency of ≥ 0.5 at any point, after the initial 30-tick warm-up period the measure diminished, ultimately receding to 0. The closer these larger networks got to this threshold level, the longer it took for them to recede to 0 - and the slower the rate of decay.

The eventual establishment of stationary processes around these threshold mean informational currency levels was unexpected; it is again unclear if these behavior patterns are due to a quirk of model design or represent a valid emergent dynamic of these complex systems. The decay in the measurements is likely due to the blockchain block gas limit: if more data is being submitted each time step than can be recorded on the ledger, and because mining prioritizes earlier transactions over more recent ones (given equal gas prices), transaction mining times will increase as the number of unvalidated transactions in the mempool grows. Higher mining times and lower informational currency measures can be attributed to the common cause of block gas limits.

However, the threshold mean informational currency of 0.5 observed in smaller networks remains unexplained. Despite thorough review of model source code, no obvious point causing such behavior was found. Further investigation into this unexpected emergent dynamic would help to identify its cause, and to understand if

it represents a result representative of actual system behavior or simply a quirk of model design.

4.2 Recorded Data Volumes

Empirical sensors operating at the edge record data representing some quality of their environment; the information contained in these data has value to other informational entities[126] seeking insight into conditions in the vicinity of that sensor.

Here the effects of changes in data volumes recorded in each observation on the dependent variables of mean transaction mining times (in blocks), financial costs (in gwei) to each sensor and network informational currency are analyzed.

4.2.1 Mining dynamics

A positive relationship between data volumes recorded per observation and mean transaction mining times is visible in Figure 12: as sensors capture more data per observation, transactions tend to take longer to mine (r_{xy} = 0.74258). Excluding two outliers, for which mining times were exceeded the upper Tukey fence[127], a Pearson's correlation coefficient of r_{xy} = 0.78448 was calculated. Summary statistics for the sample excluding outliers are shown in Table 11.

[126] Humans, and other computing nodes, including blockchain networks.
[127] $Tukey_{upper} = Q_3 + 1.5 * (IQR)$; IQR = 4.8671.

Table 11: Summary statistics, Mean transaction mining times across edge record volumes parameter sweep

n iterations		76
Transaction mining time (blocks)	Mean	7.8928
	σ	3.9433
	Minimum	1.0010
	Quartile 1	5.9778
	Median	7.8253
	Quartile 3	10.712
	Maximum	16.535

Figure 12: Sensor data capture volumes versus transaction mean mining time

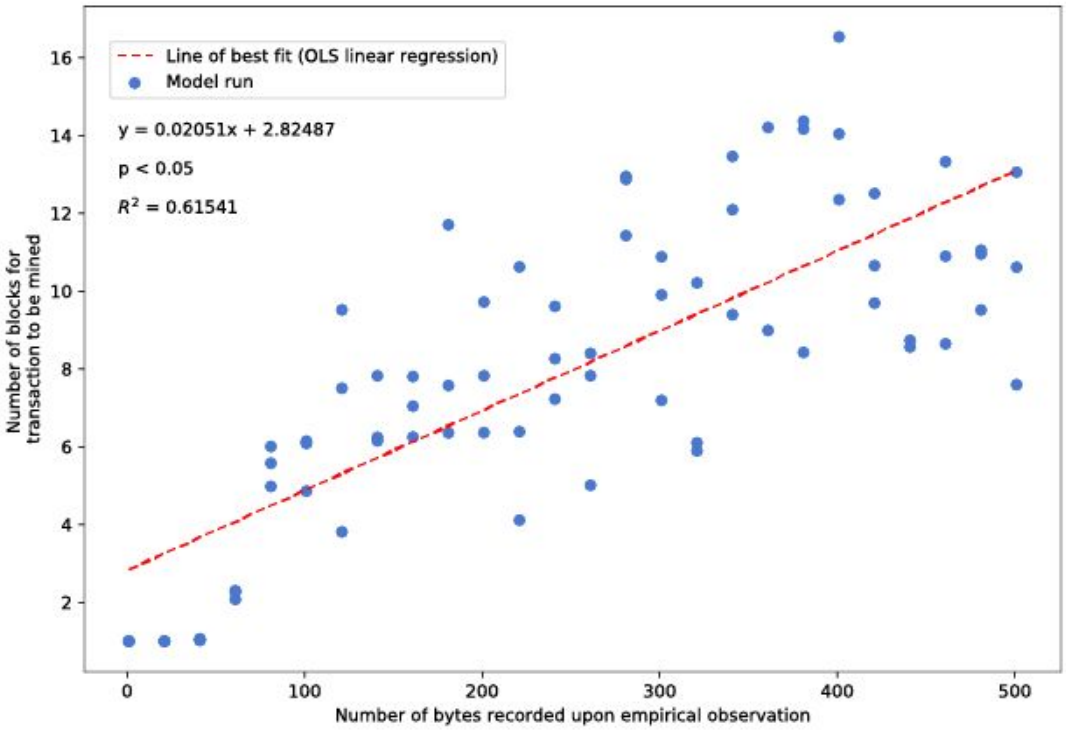


Figure 13: Residual errors versus mean mining time predicted from network size

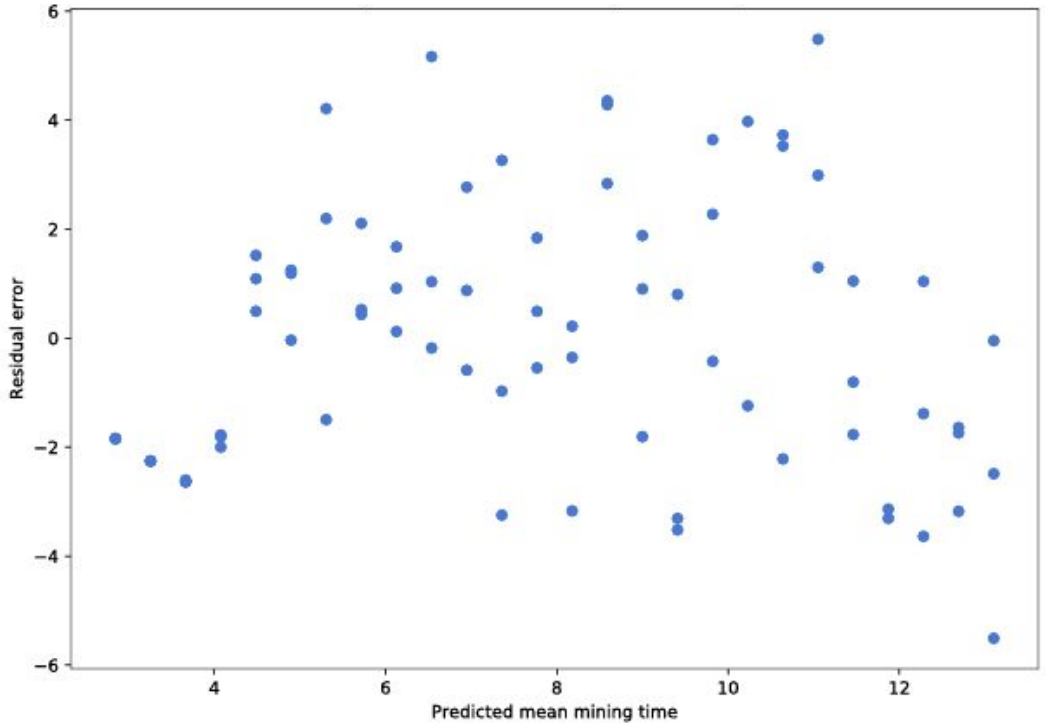
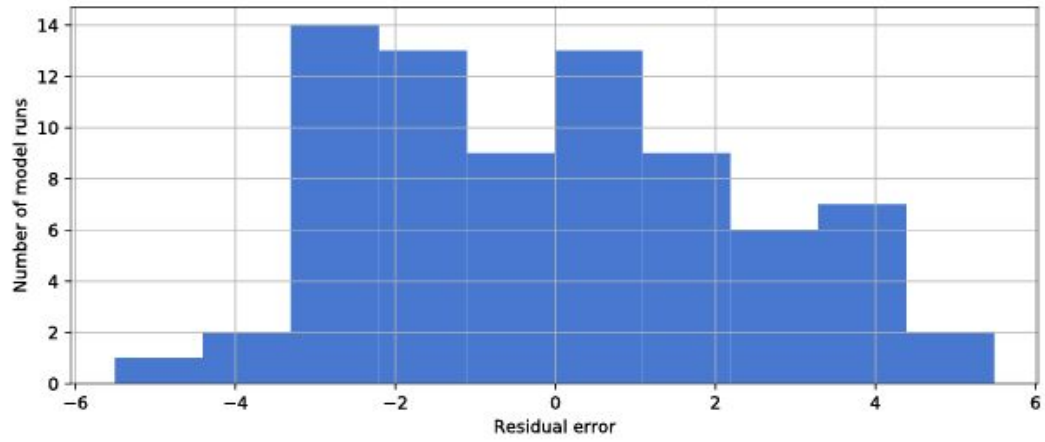


Figure 14: Distribution of residual errors of mean mining times

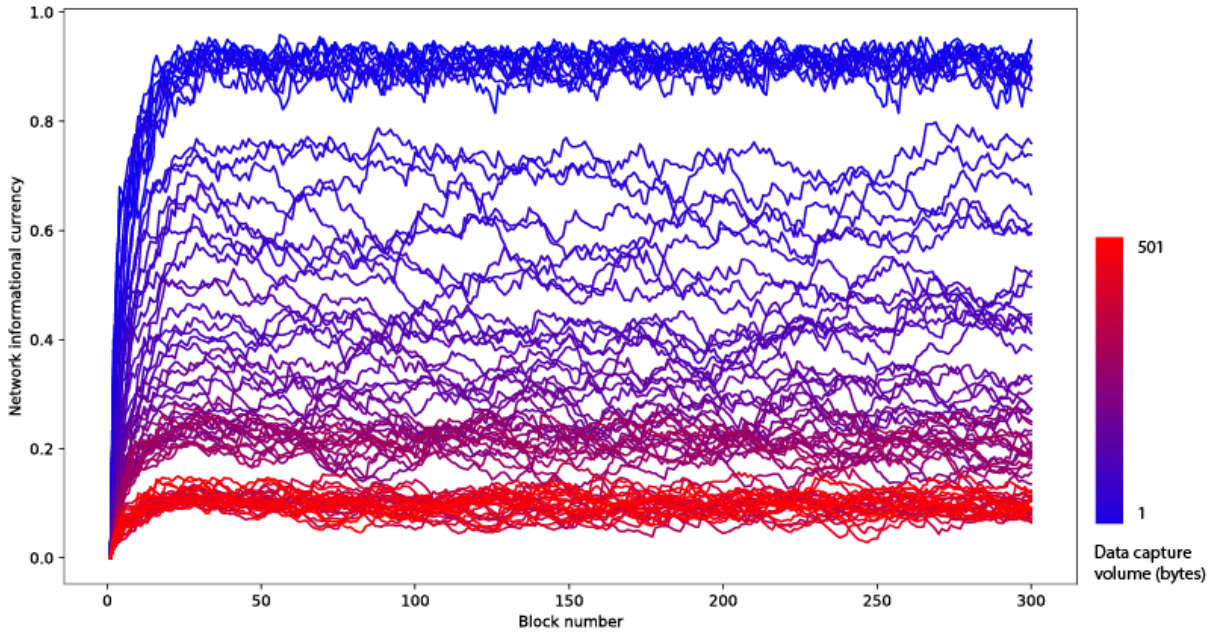


Figures 12 depicts the plot of edge recording volumes against mean transaction mining times with the line of best fit. The calculated R^2 value of 0.61541 indicates that the independent variable accounts for ~61.5% of the variance observed in the dataset. Visual inspection of Figures 13 and 14 suggest that residual error is normally distributed. The mean mining time increased by 0.02051 blocks for each additional byte of data recorded per sensor observation.

4.2.2 Informational currency

As record volumes increased, informational currency was expected to decrease, due to limitations in the amount of data that could be written to the blockchain each tick - the block gas limit. As shown in Figure 15, this was observed: after the warm-up period, in each model iteration a mean level of informational currency was established. Variations around this mean are explained by the stochasticity introduced and the probabilistic conditional transaction transmission frequency.

Figure 15: Informational currency over time across a range of data capture volumes

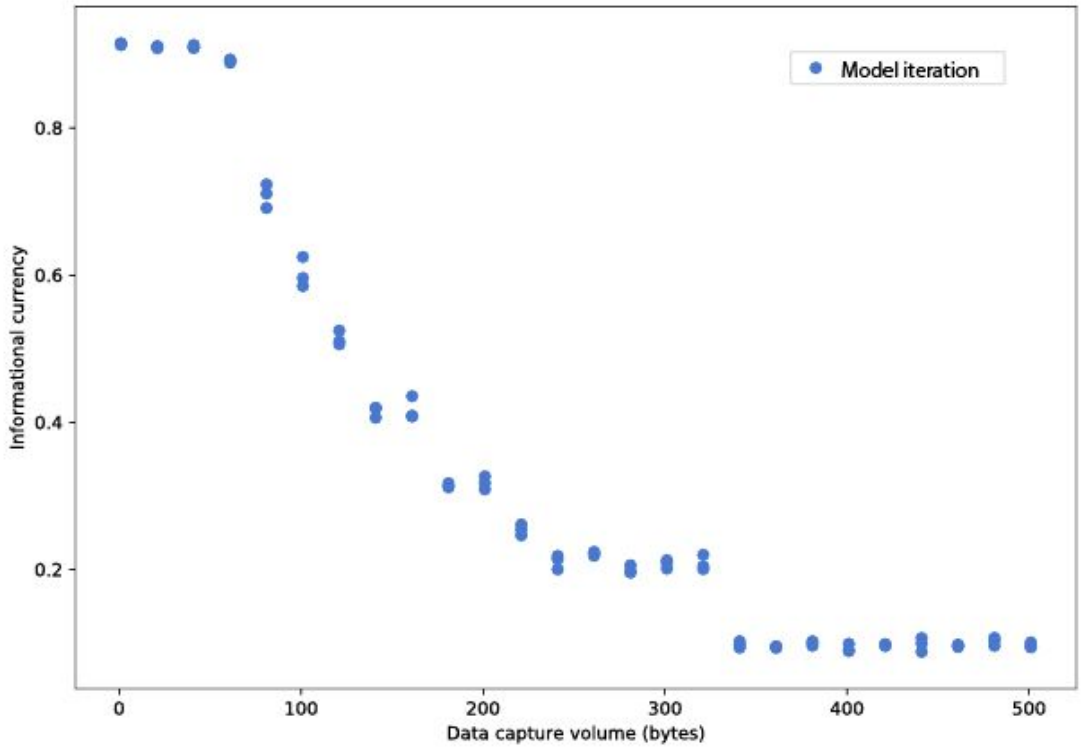


Calculating mean informational currency values for each model run (excluding the warm-up period) enabled the visualization of the effects of data capture volumes on the stable informational currency levels depicted in Figure 16.

A non-linear negative relationship is visible - perhaps a sigmoid curve. Interestingly, as the curve approaches its lower horizontal asymptote at IC ≈ 0.2, a point is reached between 321 and 341 bytes per record where the dependent variable drops

to a new stable level of ~ 0.09. The cause of the discontinuity at this threshold is unclear.

Figure 16: Mean informational currency across data capture volume parameter sweep



4.3 Sensor observation frequency

Sweeping the frequency with which edge sensors recorded empirical observations about their environment was intended to yield further insight into the effects of network activity on blockchain behavior.

Initial examination of results obtained from sweeping record frequencies using fixed parameters as in other independent variable sweeps, it was evident that the model behaved predictably up until near the limit of the sweep. Notably, with fixed network sizes of 20 sensors, the maximum gwei spent per sensor per block of 390000 meant that the block gas limit of 8000000 gwei was never reached.

Since these models were intended to simulate the challenges blockchains might encounter while scaling, and block gas limits are one of the primary constraints to network scalability, the parameter sweep was performed after doubling the number of sensors simulated in each iteration. In these sweeps, block gas limits were reached in the median swept values, enabling the analysis of model behavior as the network transitioned from underutilized to oversubscribed. This is noteworthy because for the analysis below, fixed parameters were identical to investigations into the independent variables' effects conducted above, with the exception of the number of sensors in each network iteration.

4.3.1 Mining dynamics

Initial examination of the scatter plot of mean mining times across record frequencies indicated that for record frequencies below 0.2, transactions tended to be mined immediately after submission. For simulations in which sensors took more frequent recordings, an exponential positive relationship was visible; this segment of the sample was analyzed. An OLS linear regression of record frequencies and the square root of mean mining times yielded a line of best fit shown in Figure 17, with an R^2 value of 0.99243.

Figure 17: Record frequency versus square root of mean transaction mining time

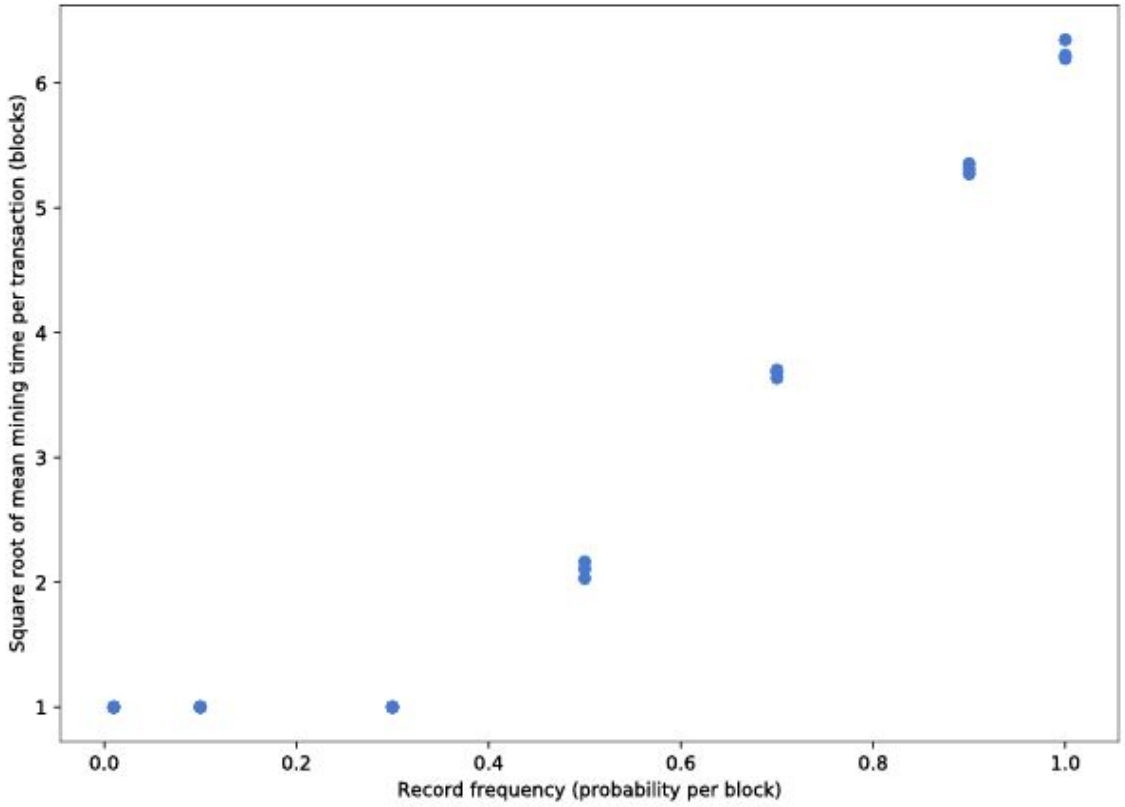
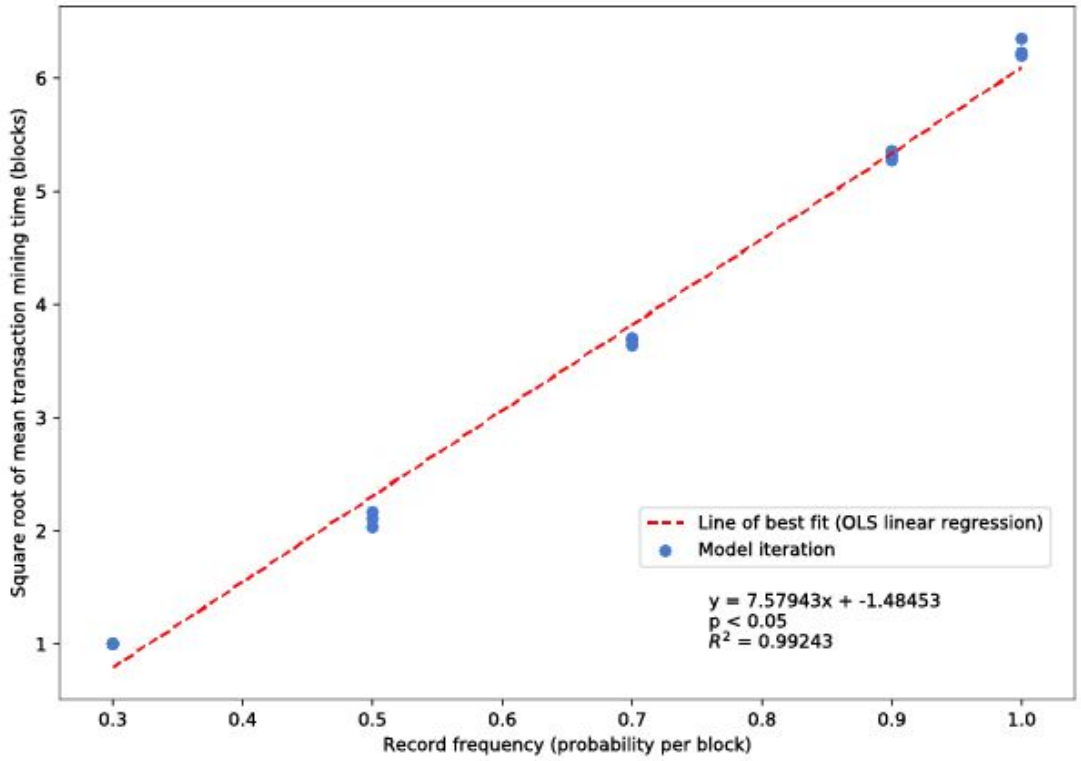
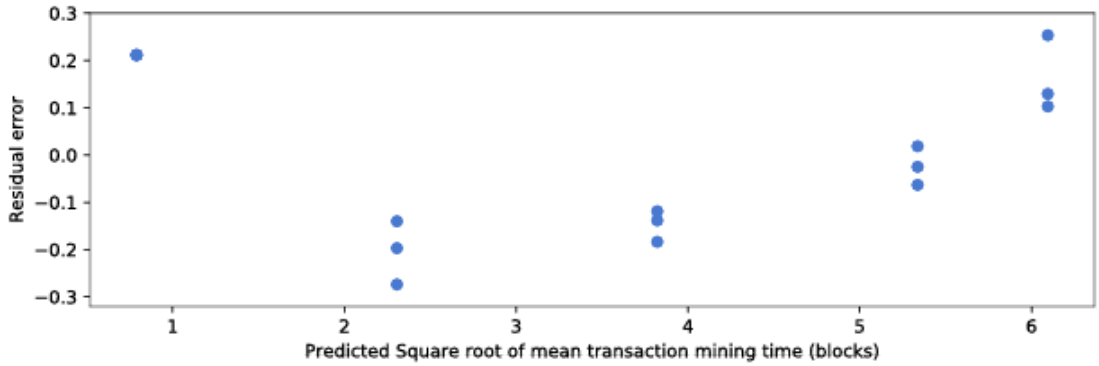


Figure 18: Record frequency (probability per block > 0.2) versus square root of mean transaction mining time (blocks)



Analysis of this plot, and of the residual errors (Figure 19) reveals that the square root transformation did not fully straighten the positive nonlinear relationship observed in the untransformed data.

Figure 19: Residuals versus predicted square root of mean transaction mining time



Performing an additional square root transformation yielded a distribution of residual errors nearer to normal, though uncertainty remains as to the true nature of the nonlinear relationship observed in the untransformed dataset (Figures 20 and 21). Deeper investigation of these nonlinearities is beyond the scope of this analysis.

Figure 20: Record frequency (probability per block, > 0.2) versus double square root of mean transaction mining time

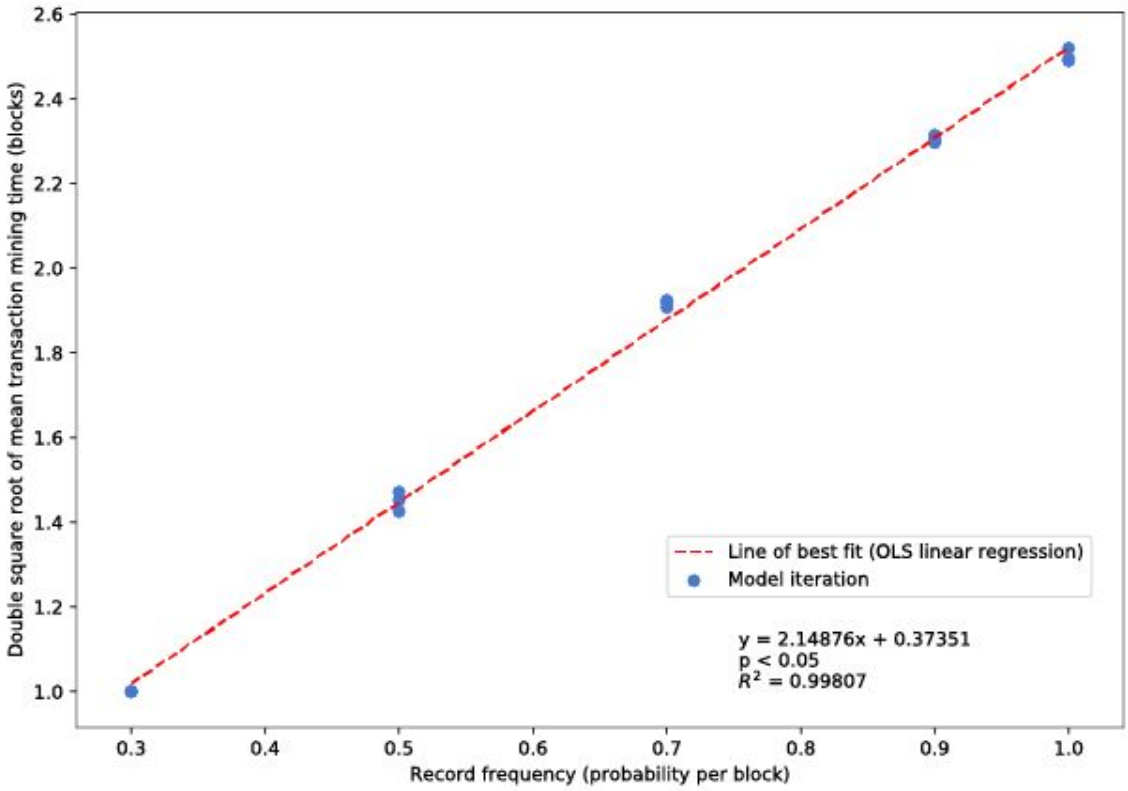
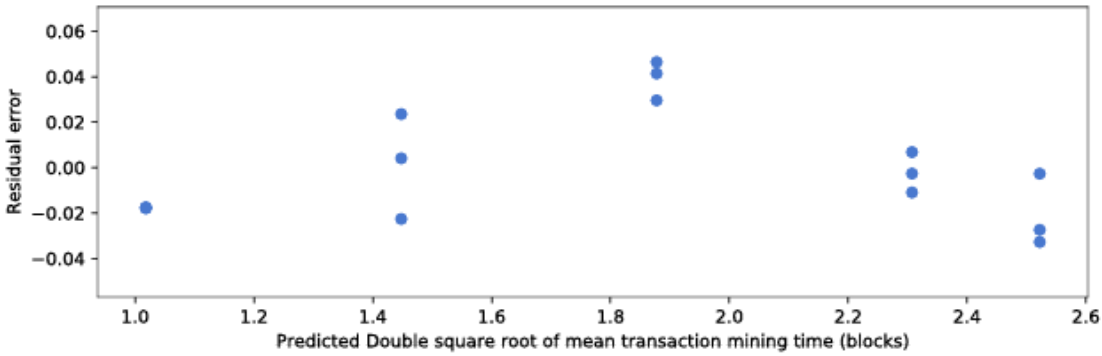


Figure 21: Residuals versus predicted double square root of mean transaction mining time



It does seem clear, however, that once a certain threshold level of transaction activity per block is reached, mining times per transaction rise in a nonlinear fashion. This result is important, and deserves further investigation, as it indicates that blockchain update performance would diminish more and more rapidly as demand increases. Substantial research into mechanisms for managing these scaling challenges due to the inevitable fluctuations in network demand is ongoing [128]; a solution to seems necessary if blockchains are to achieve their potential to form the critical informational infrastructure of the web.

4.3.2 Informational currency

Measures of informational currency collected from model runs across the parameter sweep revealed that a recording frequency of ~0.5 yielded resulted in the highest values. Less frequent updates meant the blockchain was not updated often enough to result in high measures of the metric; as record frequency (and therefore transaction volumes) increased above this optimal level increasing mining times adversely affected network informational currency.

[128] For example, sharding (Ethereum 2019b).

Figure 22: Informational currency over time across recording frequencies

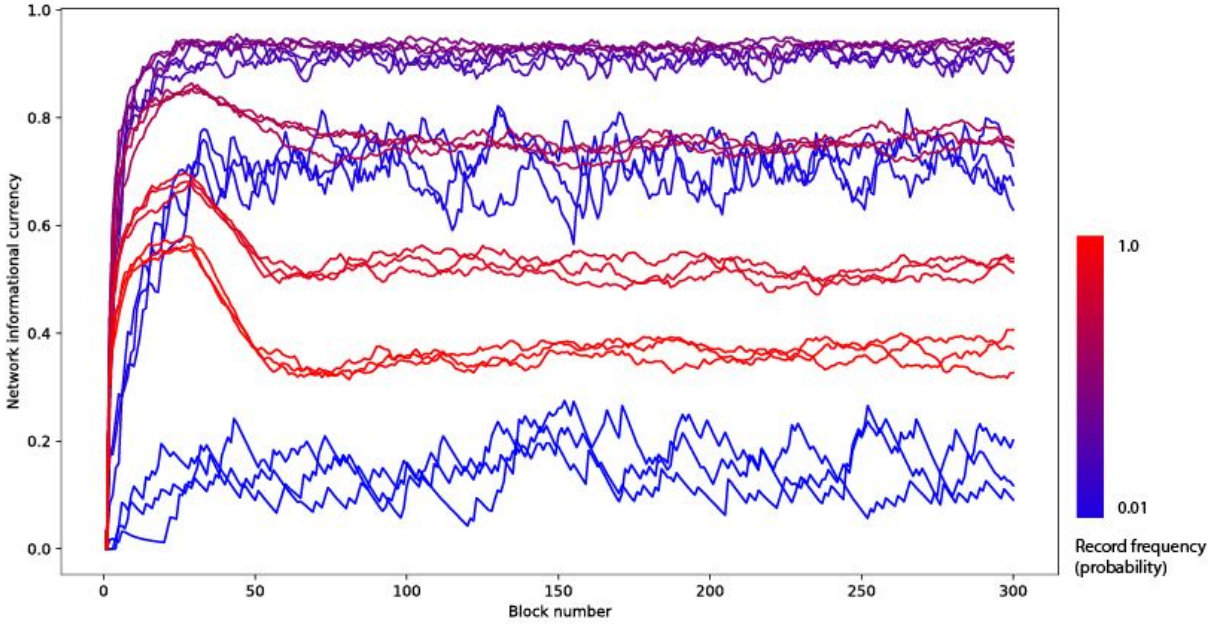
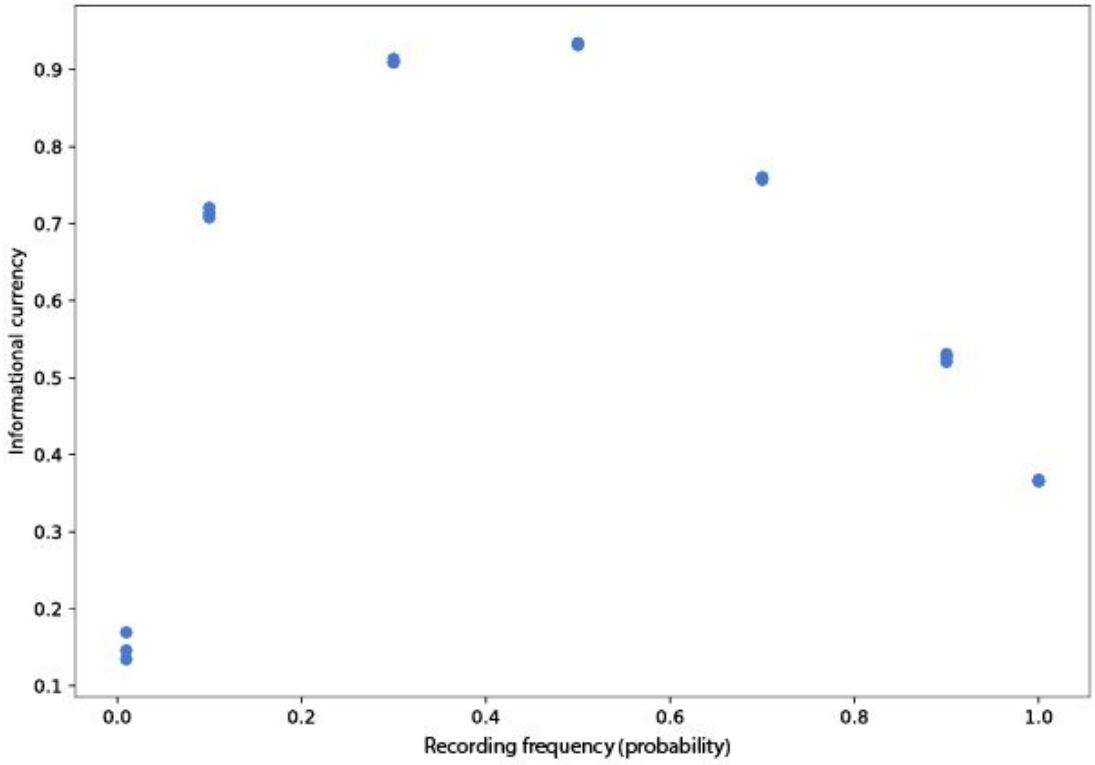


Figure 23: Record frequency vs mean model run informational currencies[129]



[129] Excluding a 30-tick warm-up period.

Interestingly, the model results depicted in Figure 22 show that at higher record frequencies, from a time series perspective within model runs informational currency measures tended to achieve their maximum toward the end of the warm-up period (equal to the window used for measuring the metric), then decayed to a stable level, at which point a stationary process was established. This pattern only occurred with record frequencies greater than 0.5.

It seems that stable states may exist at different transaction volumes. As with other results, distinguishing between authentic system behavior and unintended behavior resulting from model design is difficult. Nonetheless, the result is interesting, and warrants further research.

---

5 Discussion

This research has explored the technologies constituent to a system of trusted sensors acting as oracles to smart contracts. A middle range agent-based model of such a system was developed; initial results were analyzed. The insights gleaned from these simulations and analyses, along with a broader conversation of the implications and opportunities of such systems, are presented here.

5.1 IoT + Blockchain

Public blockchains can only function if the costs associated with utilization of the public resource deter excessive, lazy or malicious use: they represent “an expensive but decentralized and very high-assurance store of data and computation” (Buterin and Weyl 2018). Sensors embedded on devices generate large quantities of data. At first glance, blockchains do not seem a sensible informational architecture to store or compute IoT data. This intuition was validated by the observation of diminishing informational currency measures and increasing costs associated with increasing network loads.

However, techniques such as edge computing and content addressing can dramatically reduce data volumes transmitted from edge sensors without necessarily reducing the user’s access to informational value. Due to the benefits of decentralization described, the computing services blockchains offer should be considered by IoT system designers, especially when the connected sensor network might monitor the public space[130], or external actors might benefit from accessing the information it captures.

While the ideological values of public blockchain systems generally align with the author’s, private blockchain implementations should not be overlooked, as they represent a way for system designers reluctant to relinquish full control to an

[130] This research has raised many questions about the current and future status of data ethics regulation. This author believes strongly that the value captured in a sensor recording should benefit not only the device owner, but also the entities represented in (i.e. subjects of) the recording. The technical barriers to this are significant, but a range of context-dependent solutions can be conceived of.

independent organization[131] to still capture many of the benefits of decentralization, at substantially lower cost. System configuration and governance will largely determine the extent to which such private instances realize these benefits.

Designers of IoT systems connecting to blockchain networks will need to balance the trade-offs between edge resource consumption, financial costs, informational availability and control over data and computing resources. In the near future, focus should be placed the highest-value use cases for connecting edge devices to smart contracts, where the cost could be justified. Parametric insurance products, which trigger a payout if some threshold condition[132] is reached, seem an optimal initial test case (Lloyd's 2019). While this technology is still in its infancy experiments should be conducted such that the harm caused by failure or unexpected contract behavior would be inconsequential.

5.2 Proposals and Recommendations

5.2.1 Ricardian treaties

Trusted sensors connected to blockchain networks might provide a transparent, secure and trusted way for loosely coordinated entities[133] to maintain situational awareness. This author is especially interested in applications of the technologies for sovereign nations holding one another to account. He imagines an agreement between sovereign nations - a valid, legal treaty[134] - that expressly delegates some authority to a smart contract. Contract code would be open to inspection by all. Parties could be required to submit digitally-signed evidence of treaty compliance; violation of terms would automatically trigger agreed-upon enforcement mechanisms. Such technical mechanisms[135] incentivizing proper behavior could

[131] i.e. a blockchain developer community, mining network and so on.
[132] Flood waters exceed a certain height in a building, say, or a flight is delayed by more than a number of minutes.
[133] Especially ones with no trusted central authority.
[134] "A treaty is a formal written agreement entered into by actors in international law, namely sovereign states and international organizations. A treaty may also be known as an international agreement, protocol, covenant, convention, pact, or exchange of letters, among other terms. Regardless of terminology, all these instruments may be considered treaties subject to the same rules under international law." (Wikipedia 2019k).
[135] Such as the forfeiture of staked funds, or the automated enforcement of sanctions on participating states that fail to adhere to agreement terms.

supplement legal ones, which can be slow and difficult to implement. The infinite programmability of such Ricardian treaties could enable the application of a sophisticated system of microincentives to states (Hoopes 2019), perhaps integrated with central bank digital currencies such as the Synthetic Hegemonic Currency proposed by Bank of England governor Mark Carney at the Jackson Hole Symposium in August 2019. It could be an opportunity to "sort out the deep flaws in the international monetary and financial system" - to "change the game" (Carney 2019).

Arms control represents a particularly high impact application of this system of oversight.

As an example, if secured in containers embedded with trusted sensors (Becha 2019), a protocol requiring intermittent data transmissions including location and local environmental characteristics could provide governments confidence that their adversaries are behaving in compliance with mutually-agreed rules.

Of course, privacy and security concerns about the entities authorized to access this sensitive information are valid; it seems that a private, permissioned blockchain implementation[136] would be more appropriate and palatable to sovereign nations considering such an arrangement[137]. Many applications beyond arms control are conceivable, including in citizenship, immigration, international finance, military intelligence sharing, environmental monitoring, international trade, disaster response, conflict mitigation, and so on. Such Ricardian treaties could bring many of the benefits of decentralization to international affairs, and improve sovereign nations' ability to leverage the improving capacity to maintain situational awareness around the world. After all, it could be argued that in the international arena no clear central authority exists, involved parties "have conflicting incentives", and "there is a need for a shared common database" - a context for which blockchains are well-suited (Paul 2018).

[136] Likely with nodes hosted by adversarial governments, as well as entities with a range of incentives - a consortium bound by legal agreements and incentives.
[137] This may be another situation in which homomorphic cryptographic techniques and zero-knowledge and range proofs could provide assurance to adversarial states without revealing additional strategic information. For example, Glaser's (2014) work on "A zero-knowledge protocol for nuclear warhead verification" and other zero-knowledge proof protocols.

5.2.2 A DAO for access control

This work was inspired in part by the author's recognition of an enormous opportunity being missed. Legions of sensors are being deployed and connected to the Internet - in theory[138], able to provide information feeds to people who could derive value from that information. However, due to the current configuration, based on the client-server web paradigm and an approach that is often required that private interests are placed over public ones[139], much of the value to humanity being created is lost as the relevance of the data fades as it is left siloed on its device or in a proprietary relational database.

This is, in this author's view, a tragedy of the informational commons[140]: the failure to share information with someone who would benefit from it. There are certainly valid instances where sharing such information would provide a competitive advantage to the informee, but in many cases this is not so; squandering this value is tragic.

As a solution to this, a peer-to-peer data access protocol is imagined[141]. Edge nodes would need to be visible to a central administrator - in this conception, a DAO hosted on a blockchain. Data consumers could - for a fee, or under certain conditions, or possibly if assessed as deserving by some[142] review mechanism - connect to edge networks and retrieve data for analysis. This could provide sensor owners an additional source of revenue - and they could price the data according to their perception of its value - and it would provide the public an opportunity to extract value from the unprecedented capability for situational awareness emerging.

A thorough exploration and feasibility study of this system is beyond this paper's scope, but it seems that work related to decentralized public key infrastructures (Allen 2015), content addressed storage (Benet 2014), and proxy re-encryption

[138] As in, technically (very often).
[139] Many sensors would not be installed if their owners could not monetize them.
[140] One of three identified. See Appendix 4 for the three tragedies described.
[141] Based on this author's limited understanding, likely employing IPFS content addressing technology, or as recently proposed by Chen, Ramsundar and Robbins (2019).
[142] Ideally decentralized.

techniques (Nuñez 2018) might resolve many of the most difficult challenges to its implementation.

5.2.3 The voluntary transition to self-sovereign identities

Centralized authorities that have entrenched themselves as society's informational infrastructure through the Web 2.0 period now hold much of the data about us[143]. It has been compellingly argued that "forces of intelligent persuasion" "undermine the integrity of the human will" (Williams 2018), and that data should be treated as labor rather than capital (Ibarra 2017).

As the legal custodians of the personal information of billions of humans, these centralized data custodians should voluntarily lead the transition of data custody to their self-sovereign users. Acknowledging the "irrationality"[144] and possible legal challenges in doing so, it is, in this author's view, clearly the right and sustainable thing to do.

[143] This dissertation is being composed on a Google Doc. Does that mean it belongs to Google? Should Google be allowed to read it as it is being written?
[144] Economic irrationality. Rationality exists on many spectrums and varies with the scope of perspective.

6 Conclusion

This research was motivated by my curiosity about the world and guided by a dawning understanding of the magnitude of the opportunity and risk we face in the coming decades. Information is a fundamental component of our reality[145], and yet we have only begun to grasp its nature. In this moment - an apparent inflection point - I am compelled to leverage my privilege, skills and knowledge to promote the ideals I hold.

Computers are in many ways the crowning technological achievement of the modern world, providing an intensely useful tool for us humans. As a result, our globalizing society is weaving them into the fabric of our lives, meaning the physical and informational infrastructure we rely on to sustain civilization increasingly depends on these devices and the insights they yield.

As with any tool, computers are ambivalent to their application[146]: it is the users who decide if they will be used morally or not. And like every innovation, their intrinsic unfamiliarity has left us uncertain of how to properly govern them, and of the effects they have on human dignity and planetary well-being when adopted at scale.

These three qualities - the usefulness, ambivalence and unfamiliarity of computing technologies - explains my interest in pursuing this research. To invoke Greta once again: “We are failing but we have not yet failed” (Thunberg 2019 pp 44). Mistakes are to be expected, missteps and misjudgments forgiven. What is no longer acceptable, however, is to move fast, break things, then fail to learn, adapt, and evolve our systems and laws and norms and cultures and behavior in response to our discernment of harm and inequity - even if it is difficult to do so.

Much as cooling water reaches a point at which its molecules become ordered and aligned, a similar phenomenon might occur as we approach saturation of

[145] Both individual and shared - perhaps another interesting aspect of information and other conceptual objects is that they are all that comprises each of our individual realities. Physical reality is shared - it is through interpretation we find incompatibility.
[146] Although - as informational entities - do computers have the potential to care?

information transfer between sentient entities. As understanding improves, entities might adapt to be less likely to act in a way that would disrupt another's intentions. Deconfliction of behavior may be possible[147].

Toward a theory of conceptual reality

Reflecting on the breadth of technical, scientific, ethical, political and economic reading and learning I have done in this research effort[148], I am struck by a missing link, one that I am coming to realize is fundamental to a complete and coherent worldview inclusive of the informational domain.

Communication is the transfer of meaning between informational entities. Shannon (1948) formalized a mathematical theory of communication, based on digital[149] data. However, not all communication is mathematical in nature[150]. While Shannon's definition of a "message" is useful in his context, it is obviously only partial. Further, it seems to be a profound simplification to reduce the model of communication to a one dimensional series with only two radices - the two binary digits[151].

My recognition of this missing link is pointing to an interesting finding: information does not objectively exist. Information is subjectively perceived, and only exists within the awarenesses of sentient[152] agents. Furthermore, data seems to *only* objectively exist - it is physical. Data can be stored on a physical object as matter[153] or be transmitted as physical energy through space. Data is governed by the laws of physics[154]. For the information contained within data to be discerned, it must be sensed and interpreted by another informational entity[155].

---

[147] These ideas on another scale entirely: could entangled (Simonsen 2018) blockchain networks offer a solution to the dark forest theory of deterrence (Liu 2008)?
[148] Incomplete though it is ...
[149] In his case, binary.
[150] Perhaps none of it is? Communication is the transfer of meaning *through space and time*. Mathematics exists in the conceptual space - in which neither time nor space exists.
[151] See Appendix 7.
[152] If "sentient" means able "to perceive one's environment" (State of Victoria 2017).
[153] And thereby persist through time.
[154] Albeit quantum physics, which is less well understood than classical mechanics.
[155] i.e. in the awareness of the informee.

This distinction may resolve many of the questions left inadequately addressed by the literature reviewed and the contemporary paradigm. A complete investigation into its validity and implications is well beyond the scope of this dissertation, but initial ideas and observations are outlined in Appendix 7.

Of particular interest is the difference between physical reality and conceptual reality entailed by this distinction. Physical objects exist in the physical space, and adhere to the physical laws - these are well understood. Conceptual objects[156], however, do not appear to exist in a physical space[157], but rather in conceptual space, only manifesting within the awareness of a perceiving[158] entity. As conceptual objects do not exist in physical space, they do not adhere to the physical laws[159].

So - what are the laws governing conceptual reality?

[156] Concepts? Are concepts the same as conceptual objects?
[157] Or at least do not primarily manifest there. Initial enquiry suggests that each conceptual object is necessarily grounded in some physical manifestation - see Appendix 7.
[158] Also, informational; discerning of meaning.
[159] See Appendix 6 for an exploration of an example conceptual object.

Appendices

Appendix 1: Submodels

On model instantiation a single instance of the `Blockchain` class is created. The specified number of `Sensor` objects are created and the model `Blockchain` object is added to each as an instance variable[160]. `Sensor` instances were then added to the model scheduler - but the `Blockchain` object was not.

Each tick a number of operations were executed.

In an order randomized each step, `Sensor` agents were activated by invoking their `Sensor.step()` methods. If the `Sensor` was active[161], this process began with the invocation of the `Sensor.record()` method. If the sensor was determined to attempt a transmission that tick[162], then a `Sensor.transmit()` method simulated the preparation and signing of a transaction, transmitted to the blockchain network [163].

[160] For these model runs the `Blockchain` object could have been assigned the `Sensor` objects as a class variable. However, this design pattern was chosen to leave room for a model design that includes the simulation of multiple blockchain networks, with different sensors connecting to different chains - or even the same sensor submitting transactions to different chains based on some logic executed on the device. In this way, each sensor attained the ability to invoke multiple `Blockchain` object methods.
[161] In order to enable modeling of edge sensors constrained by limited energy supplies, a `Sensor.mortal` boolean was included. If the `Sensor.battery_life` instance variable was depleted below zero, the `Sensor` instance would not perform on-device actions such as recording and data transmission. These sensors were not simply removed from the scheduler because they might still have unvalidated transactions pending in the mempool, and needed to be accessed via model scheduler for the invocation of `Sensor.confirm_tx()` method by the `Blockchain` object upon validation. This would ensure complete recording of costs. For the simulations analyzed here, however, edge devices were not mortal - i.e. they were connected to a reliable power source, rather than a limited battery.
[162] Based on comparison of its `transmit_freq` variable with either a randomly-generated float value (probabilistic) or the current block number (deterministic, interval-based).
[163] A `Sensor.compute()` method to simulate data reduction by analyzing data on the edge device was defined, but not utilized in the batch runs performed. This code would enable modeling of energy usage at the edge, and would have implications for each of the dependent variables measured in model runs: mining times, financial costs and informational currency. This method could simulate the generation of a hash (perhaps an IPFS Content ID) on board the device, to be registered on chain. This could help system users ensure that edge data integrity has been maintained. Alternatively, summary statistics (or

Within the `Sensor.transmit()` method execution the sensor invoked the instance's `Blockchain.add_to_mempool(tx)` method, simulating the submission of the valid, signed transaction to the blockchain miners to add to the queue of unvalidated transactions.

Within the model `SensorBlockchainNetwork.step()` method, after randomly activating each `Sensor` agent, the `Blockchain.mine_block()` method was called [164], in which a subset of unvalidated transactions was selected from the mempool[165]. Upon transaction validation, the `Blockchain.chain` dataframe was updated such that the appropriate sensor's column reflected the current state of on-chain data availability. Additionally, the appropriate sensor's `Sensor.confirm_tx()` method was invoked, simulating the deduction of the gwei spent for transaction validation from the agent's account.

---

some other information reduction technique) could enable a balance to be struck between edge compute resource usage and wireless bandwidth usage.

[164] This aspect of model design would need adaptation to simulate the operation of multiple blockchain networks. If the activation of all `Sensor` agents prior to the `Blockchain` agents mining action is as important as assumed, the `RandomActivationByBreed` subclass defined in Mesa's `wolf_sheep/schedule.py` example may achieve this (Core Mesa Team 2018).

[165] Transactions were selected for inclusion in a block based on transaction value, to simulate the miners' incentive to mine highest-value transactions first, then block submission time, to prevent transactions arbitrarily remaining unvalidated (getting "stuck"). It is worth noting that transactions can get stuck in the Ethereum mainnet mempool, remaining unvalidated (usually because the gas price is so low that miners do not select it for validation) (McDonald 2017). It was decided to remove this feature to simplify the modeling of network informational currency; it is unknown how its inclusion might affect model behavior. Furthermore, although gas prices were fixed in this research, the modeling of variable gas prices could shed light on the costs of increased informational currency in heterogeneous agent networks. This approach to transaction selection should capably handle such an adaptation.

Appendix 2: Additional Results

Network Sizes

Mining Dynamics

A positive relationship between network size and mean transaction mining times per transaction was observed: larger networks tended to validate transactions more slowly.

Table 7: Summary statistics, Mean mining times per transaction across network size parameter sweep

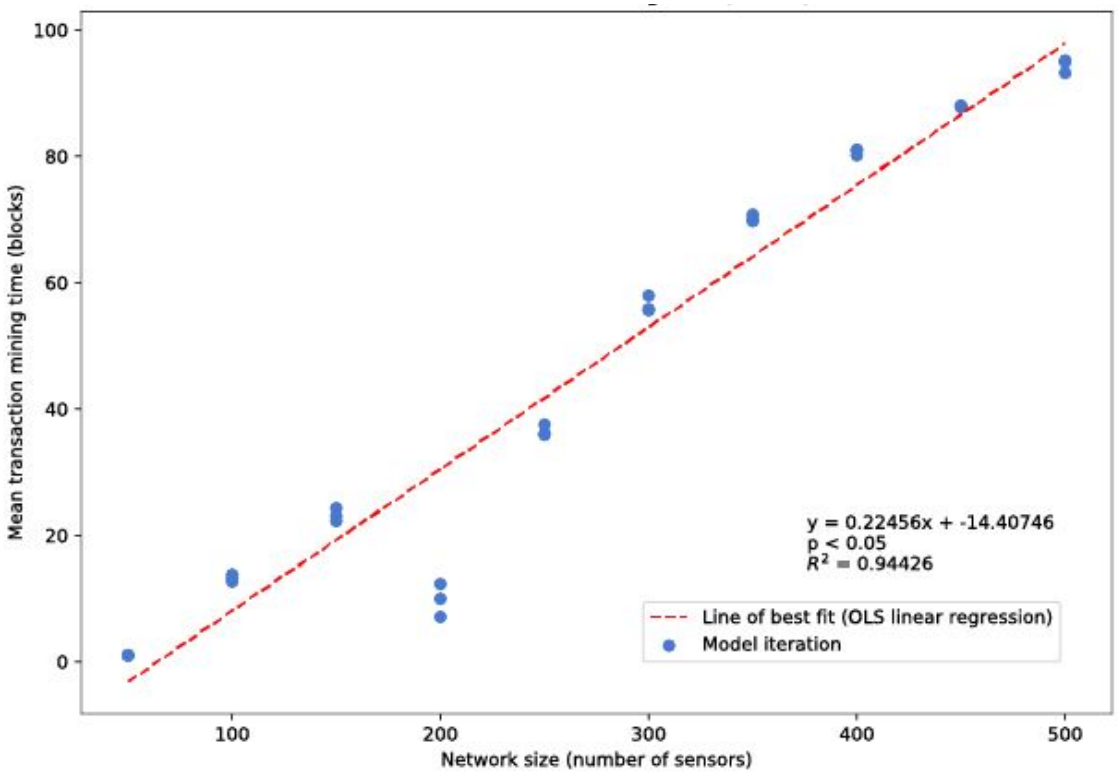
n iterations		30
Transaction mining time	Mean	47.34600
	σ	33.75534
	Minimum	1.00404
	Median	46.59344
	Maximum	95.18947

The observed mean mining times exhibited a positive correlation (Pearson's correlation coefficient r_{xy} = 0.971728). It stands to reason that this correlation is due to the block gas limit, which fixes a limit on the amount of data the network can write each tick. Transactions each specified the same gas price, and data capture volumes and transmit frequencies were homogeneous[166] across sensors, so network size directly correlated with the volume of transactions being submitted to the blockchain network for validation. Because blocks have a fixed limit to the gas that could be consumed - here, the amount of data to be written[167] - in larger networks generating more data, transactions will take longer to be validated.

An ordinary least squares simple linear regression analysis was conducted to produce a line of best fit modeling the relationship between network size and mean mining times.

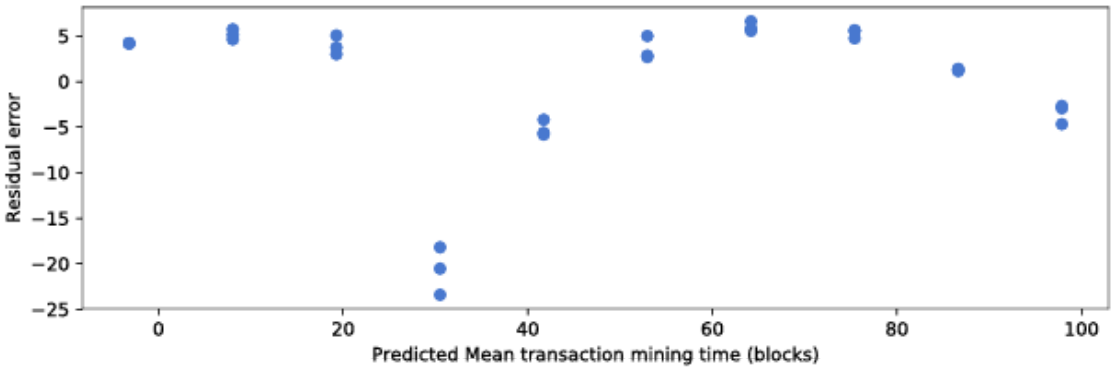
[166] Though stochasticity did cause slight variation.
[167] Since on-chain compute operations were excluded in this model.

Figure 6: The effect of network size on mean mining time per transaction



While the correlation coefficient indicates a strong positive correlation between the variables considered, the residual errors do not appear to be normally distributed (Figure 7), calling into question the validity of the line of best fit calculated using OLS simple linear regression.

Figure 7: Residual errors versus fitted values, OLS simple linear regression on network size versus mean transaction mining times



Curiously, the observed values at a network size of 200 breaks with the trend established in smaller network sizes; mean mining times were, on average, less than 50% of the values observed in sensor networks of 150 sensors. It is unclear if this break in the otherwise consistent positive relationship between the two variables is due to some quirk of the model or true emergent behavior of these systems interacting in reality; the former seems likelier, but without a dataset to validate observations from the simulation this is difficult to assess.

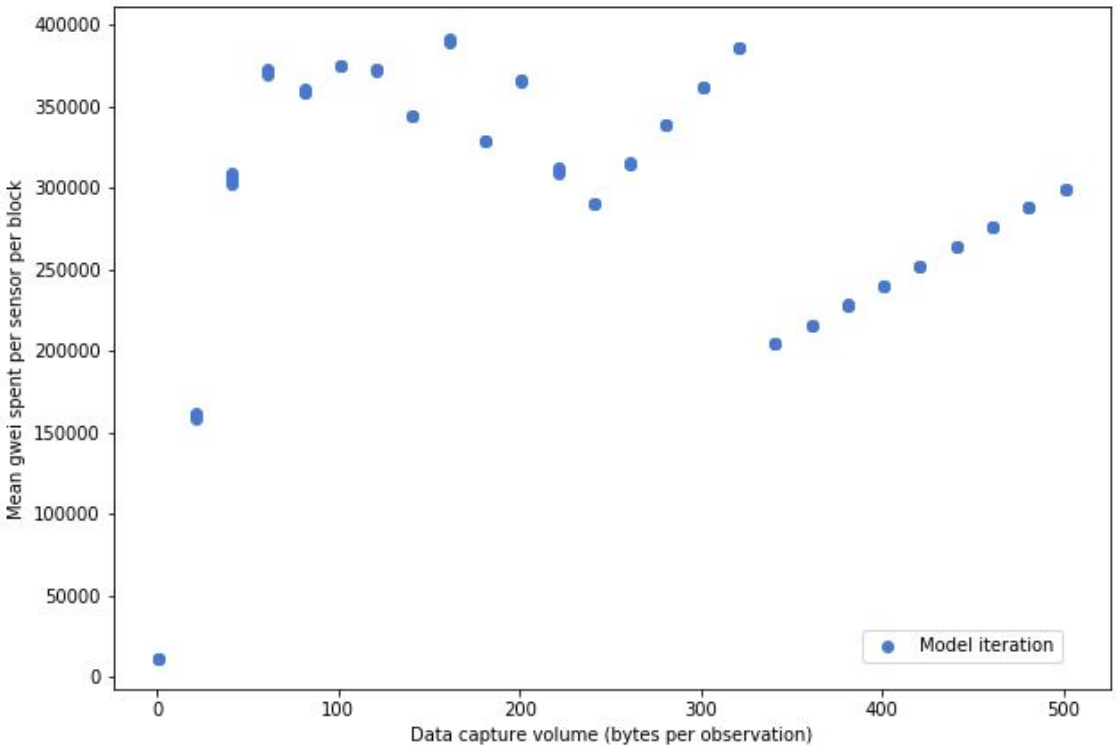
Visual inspection of the scatterplot suggests the possible subsequent establishment of a new, convex positive non-linear relationship for networks of 200 or greater sensors. Further investigation of the nonlinear effects of transaction volumes on mining dynamics is warranted, but beyond the scope of this investigation.

Recording Volumes

Gwei spent

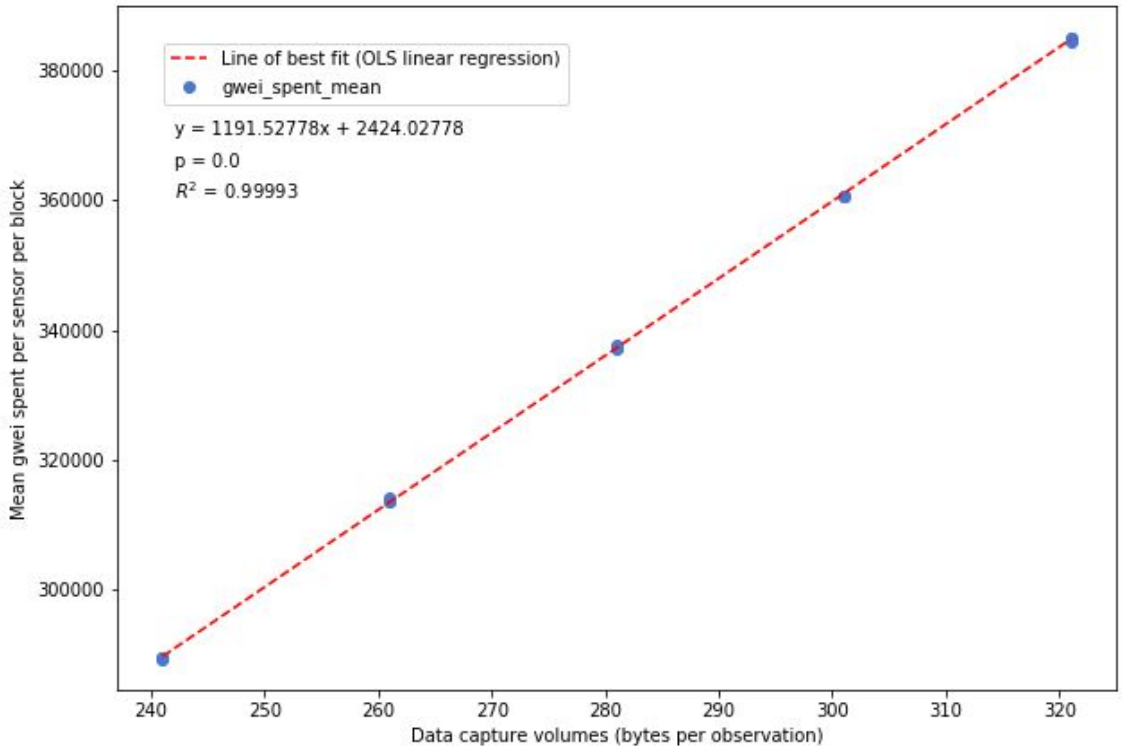
As Figure 24 shows, recording volumes had an unusual effect on gwei spent per sensor. At smaller record volumes, the financial expenditures rose rapidly with the independent variable, quickly leveling off between 350000 - 400000 gwei spent per sensor. These mean recordings appeared to begin to decrease as record volumes approached 241 bytes, at which point two highly uniform segments of data are seen.

Figure 24: Data capture volumes against mean gwei spent per sensor



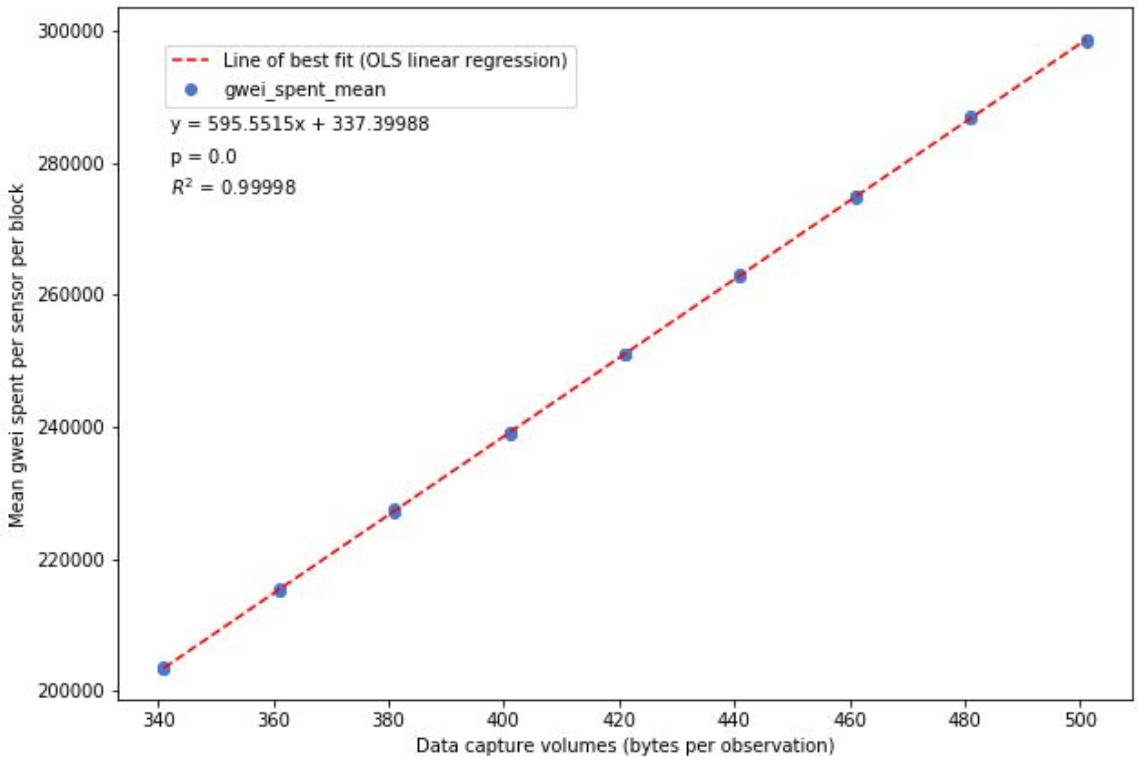
First, in model runs simulating edge recording volumes of 241 - 321 bytes, a very strong positive linear correlation is observed (r_{xy} = 0.99997). Figure 25 depicts this segment of the sample, including the OLS line of best fit - note the R^2 value of 0.99993, indicating that 99.99% of variance is explained by the independent variable.

Figure 25: Data capture volumes (241 <= bytes <= 321) against mean gwei spent per sensor



Between 321 and 341 bytes per record, average gwei expenditures drop dramatically and begin another, less steep upward trend, also a positive linear correlation (Pearson's correlation coefficient r_{xy} = 0.99999) which continues to the end of the parameter sweep. This segment of the sample is shown in Figure 26; an R^2 value of 0.99998 indicates that this positive linear relationship is very strong.

Figure 26: Data capture volumes (341 <= bytes <= 501) against mean gwei spent per sensor



While these results are interesting, it seems unlikely that these near-perfectly correlated relationships, as well as the mid-sweep discontinuity, are due to something other than an unintended aspect of model design. As such, these results are determined to be not useful to this investigation into the dynamics of a scaling blockchain network.

Recording Frequencies

Gwei spent

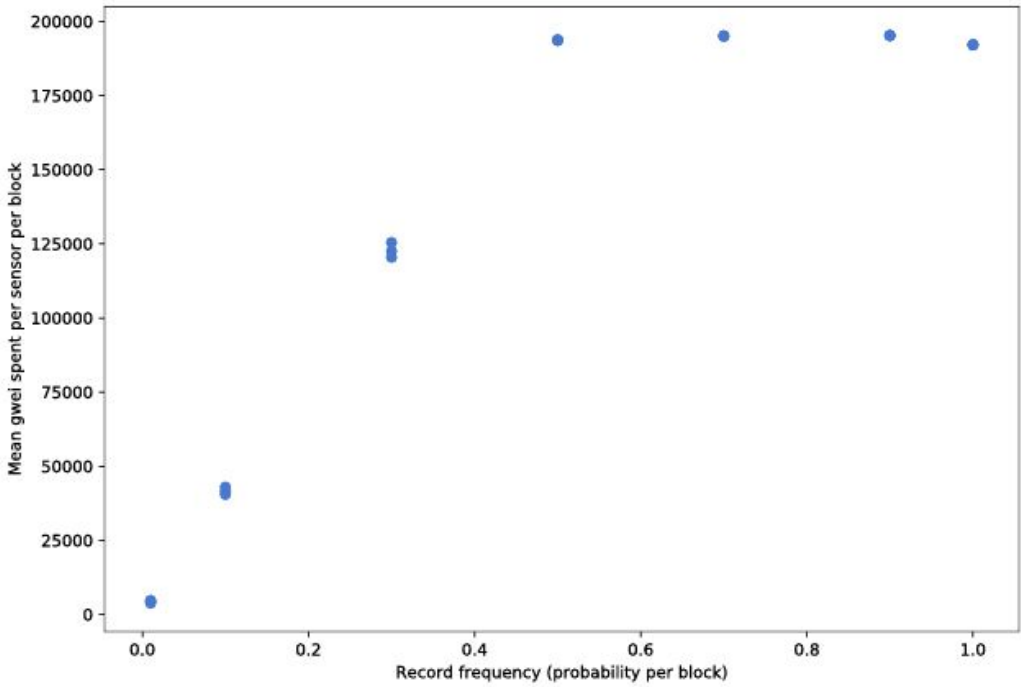
Figure 27 shows that mean gwei spent per sensor per block exhibited a positive linear correlation with record frequencies until a limit was reached, at an approximate record frequency of 0.5, after which the level remained roughly constant. This limit occurred at just below 200000 gwei per sensor per block. (Note the small difference between the median, third quartile and maximum values in Table 12). In a 40 sensor network, this is almost certainly caused by the block gas

limit: higher gwei expenditures would have exceeded the block gas limit, an impossibility according to the blockchain protocol. (Rather than exceed this limit, unvalidated transactions are simply left in the mempool to be validated in a future block.)

Table 12: Summary statistics, Mean gwei spent across record frequency parameter sweep

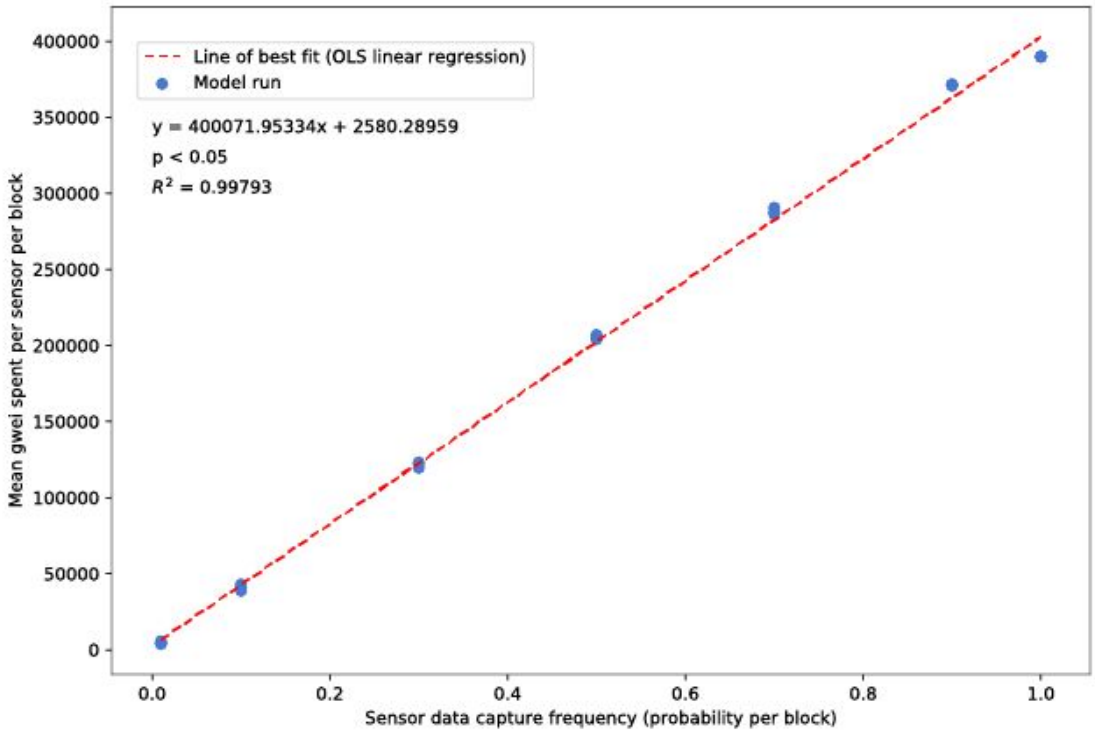
n		21
Mean gwei spent per sensor per block	Mean	135006.0
	σ	77359.1
	Minimum	3818.8
	Quartile 1	42969.8
	Median	192168.8
	Quartile 3	194968.8
	Maximum	195385.4

Figure 27: Record frequencies versus mean gwei spent per sensor per block



An OLS simple linear regression was performed to model the linear relationship between the variables at record frequency values less than 0.5; this plot, including the line of best fit, are depicted in Figure 28. An R^2 value of 0.99924 corroborates the visual assessment of a near-perfect linear relationship.

Figure 28: Record frequency (< 0.5) versus mean gwei spent per sensor per block



Appendix 3: Allen's 10 Principles of Self-Sovereign Identity

Reprinted verbatim from *The Path to Self-Sovereign Identity* (Allen 2016)

1. **Existence.** *Users must have an independent existence.* Any self-sovereign identity is ultimately based on the ineffable "I" that's at the heart of identity. It can never exist wholly in digital form. This must be the kernel of self that is upheld and supported. A self-sovereign identity simply makes public and accessible some limited aspects of the "I" that already exists.

2. **Control.** *Users must control their identities.* Subject to well-understood and secure algorithms that ensure the continued validity of an identity and its claims, the user is the ultimate authority on their identity. They should always be able to refer to it, update it, or even hide it. They must be able to choose celebrity or privacy as they prefer. This doesn't mean that a user controls all of the claims on their identity: other users may make claims about a user, but they should not be central to the identity itself.

3. **Access.** *Users must have access to their own data.* A user must always be able to easily retrieve all the claims and other data within his identity. There must be no hidden data and no gatekeepers. This does not mean that a user can necessarily modify all the claims associated with his identity, but it does mean they should be aware of them. It also does not mean that users have equal access to others' data, only to their own.

4. **Transparency**. *Systems and algorithms must be transparent.* The systems used to administer and operate a network of identities must be open, both in how they function and in how they are managed and updated. The algorithms should be free, open-source, well-known, and as independent as possible of any particular architecture; anyone should be able to examine how they work.

5. **Persistence.** *Identities must be long-lived.* Preferably, identities should last forever, or at least for as long as the user wishes. Though private keys might need to be rotated and data might need to be changed, the identity remains. In the fast-moving world of the Internet, this goal may not be entirely reasonable, so at the least identities should last until they've been outdated by newer identity systems. This must not contradict a "right to be forgotten"; a user should be able to dispose of an identity if he wishes and claims should be modified or removed as appropriate over time. To do this requires a firm separation between an identity and its claims: they can't be tied forever.

6. **Portability.** *Information and services about identity must be transportable.* Identities must not be held by a singular third-party entity, even if it's a trusted entity that is expected to work in the best interest of the user. The problem is that entities can disappear — and on the Internet, most

eventually do. Regimes may change, users may move to different jurisdictions. Transportable identities ensure that the user remains in control of his identity no matter what, and can also improve an identity's persistence over time.

7. **Interoperability.** *Identities should be as widely usable as possible.* Identities are of little value if they only work in limited niches. The goal of a 21st-century digital identity system is to make identity information widely available, crossing international boundaries to create global identities, without losing user control. Thanks to persistence and autonomy these widely available identities can then become continually available.

8. **Consent.** *Users must agree to the use of their identity.* Any identity system is built around sharing that identity and its claims, and an interoperable system increases the amount of sharing that occurs. However, sharing of data must only occur with the consent of the user. Though other users such as an employer, a credit bureau, or a friend might present claims, the user must still offer consent for them to become valid. Note that this consent might not be interactive, but it must still be deliberate and well-understood.

9. **Minimalization.** *Disclosure of claims must be minimized.* When data is disclosed, that disclosure should involve the minimum amount of data necessary to accomplish the task at hand. For example, if only a minimum age is called for, then the exact age should not be disclosed, and if only an age is requested, then the more precise date of birth should not be disclosed. This principle can be supported with selective disclosure, range proofs, and other zero-knowledge techniques, but non-correlatibility is still a very hard (perhaps impossible) task; the best we can do is to use minimalization to support privacy as best as possible.

10. **Protection.** *The rights of users must be protected.* When there is a conflict between the needs of the identity network and the rights of individual users, then the network should err on the side of preserving the freedoms and rights of the individuals over the needs of the network. To ensure this, identity authentication must occur through independent algorithms that are censorship-resistant and force-resilient and that are run in a decentralized manner.

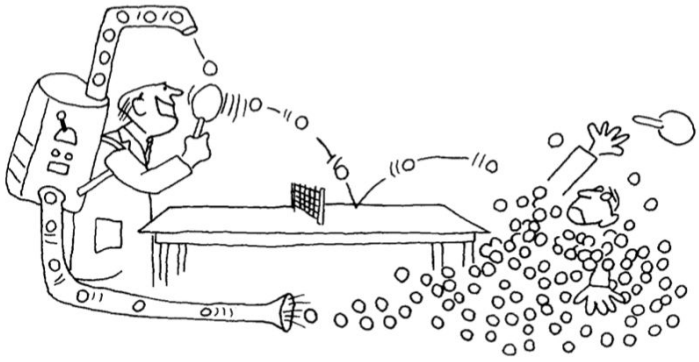
Appendix 4: The Three Tragedies of the Informational Commons

1. The failure to share information with someone who would benefit from it.

2. The abuse of a public channel's openness, rendering it unusable for others.

   Distributed Denial of Service attacks are an example of this (Beal 2019).

   Figure 29: The misuse of a public channel



. . . filibustering destroys communication.

Reprinted from Licklider (1968 pp 35)

3. Not admitting to a security breach from embarrassment or due to reputational cost.

   "Many firms treat hacks like gonorrhoea, an embarrassing affliction no one wants to admit even if speaking about it would stop its spread. Some call it a tragedy of the cyber-commons." (The Economist 2019 Schumpeter).

Appendix 5: History of Computing

A5.1: Analytical machines

Building on the work of Gottfried Wilhelm Leibniz's work on a mechanical calculator of "the four fundamental operations of arithmetic"[168] (Martin 1992 pp 39), Charles Babbage first proposed an automatic computer in 1822, later outlining an even more sophisticated analytical engine, a "general purpose computer" (CrashCourse 2017). Based on Babbage's proposed mechanical computers (Babbage 1822), Ada Lovelace foresaw the potential of computing[169], going so far as to develop software programs for Babbage's still-unbuilt machine (Lovelace 1843, Fuegi 2003, Essinger 2018). Technology was moving toward the reliable mechanization of logical operations.

In the early 20th century, mathematicians, researchers and engineers worked to incorporate the use of digital electronics in computing systems, based on advancements of understandings in the properties of electricity and conductive materials, as well as Akira Nakashima's early work on switching theory, based on two-valued Boolean algebra (Wynn-Williams 1931, Stankovic 2008). These technologies built toward Alan Turing's seminal work.

A5.2: Transmitting information prior to Shannon

As innovations in computing machines were progressing, so were techniques for transmitting information over distances. Signals had been sent on optical and auditory channels for millennia (Gleick 2011); by the early 1800s the semaphore telegraph[170] was in wide use (Burns 2004). The encoding of symbols into electrical

[168] Addition, subtraction, multiplication and division.
[169] It could be argued that Lovelace foresaw such cutting-edge computational techniques as the application of generative adversarial networks to music composition (Engel 2019): "Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent." -Ada Lovelace (Toole 1998, pp 694)
[170] By which messages are transmitted via line-of-sight relay stations on a visual channel.

currents had been discussed for some time (Fahie 1884) prior to “the first working electrostatic telegraph” being built in 1816. Such a system conferred substantial benefits over the optical telegraphs in use at the time[171] (Norman 2019).

These telegraphy systems shared the attribute of providing a communication channel upon which messages could travel rapidly and accurately. They necessarily relied on the establishment of some system of encoding information on the channel - without the ability to interpret the symbols encoded, the data would be meaningless to the informee. Enter Claude Shannon.

A5.3: Early innovation in internetworking

The vision for networked computers was put forward with striking clarity in the late 1950s and early 1960s by computer scientist J.C.R. Licklider (Living Internet 2019): a “galactic network” of interconnected computers (Leiner 1997). Licklider and his contemporaries investigated the concept, including the development of a theory of packet switching, which enabled a channel to be used by multiple traffic sources (Kleinrock 1961). These developments led to the first wide-area network connection being established over a telephone line connected California and Massachusetts in 1965 by Merrill and Roberts (Leiner 1997).

In 1966, the Defense Advanced Research Projects Agency began funding the development of ARPANET, which accelerated the refinement of “the overall structure and specifications” for such a network of computers (Leiner 1997). By the early 1970s, the technical capability for networked computing was established.

However, many of the networks relied on protocols implemented within organizations: because “these protocols have addressed only the problem of

[171] According to inventor Francis Ronalds, the system offered "a mode of conveying telegraphic intelligence with great rapidity, accuracy, and certainty, in all states of the atmosphere, either at night or in the day, and at small expense." Privacy was also greater, as the optical channels semaphores used were public; accessing the electrical signal carried on an electrical conduit would require additional specialized equipment and knowledge. Expense was reduced not least due to the reduction in the required line-of-sight infrastructure and personnel required to relay an optical message.

communication on the same network" (Cerf 1974 pp 1), internetworking remained difficult. In *A Protocol for Packet Network Intercommunication*, Cerf and Kahn (1974) established the Transport Communication Protocol, which "provides reliable, ordered, and error-checked delivery of a stream of octets (bytes) between applications running on hosts communicating via an [Internet Protocol] network" (Wikipedia 2019d). It could be argued that this marked the birth of the Internet, as this protocol is crucial to its functioning.

Appendix 6: Money as a conceptual object

A relevant example of a conceptual object: a monetary unit. It seems that our understanding of money is that it adheres to the physical laws - like physical objects, that there can only exist one of any one instance of a class (i.e., dollar). These units are scarce; it is part of what makes money money. This derives from our worldview regarding the nature of physical objects, and from the origins of money, when units were physical.

However, multiple instances of the same unit of a currency can exist, because they are conceptual objects, not physical ones. This is the basis, in this author's highly limited understanding, of fractional reserve banking: banks lend out a second version of the dollars they are given by depositors. It appears that this is enabled by the fungibility of money (Hodgson 2012) - the "sameness" of its individual units (which is not a physical property, but a conceptual one, see Appendix 7) - depositors do not care if they receive a dollar different to the one they deposited when they go to withdraw funds. If they did - if money were non-fungible[172] - banks could not promise that depositors could withdraw whenever they want and only keep a fraction of deposits on hand (Investopedia 2019).

It seems that this violation of the conceptual integrity of monetary units[173] does not occur in the context of cryptocurrencies such as bitcoin and ether. Though they exist almost entirely within the informational space, there can only be one of each instance[174], represented by the private key required to transfer the unit to another. This is the first time money has had this attribute of conceptual integrity since at least the abandonment of the gold standard, before which monetary units represented physical objects: pieces of gold of a standard weight - though gold is considered to be fungible; bitcoin may be the first conceptual object with complete integrity ever created. Based on this, Bitcoin should not be thought of as a bank where money is stored; it *is* the money ...

[172] Also meaning heterogeneous - as all physical objects are (Appendix 7).
[173] Violating integrity because they are intended to be unreplicable.
[174] Based on the existence of a single authoritative ledger representing ownership, the blockchain.

While this author is skeptical of the sustainability of the current configuration of fractional reserve banking, its implementation in reality should not be confused with it in the ideal. It clearly offers some value, or it wouldn't have been adopted.

Here the potential of the locked coins comes to mind. Of the 17,899,562 BTC in circulation at the time of writing (coinmarketcap 12:27:00 UTC+1, 27 August 2019, Block 591965) - coins that exist, that were mined, are stored at a valid wallet address [175] - it is estimated that 36% of the Bitcoin minted is held in wallets where the private keys represented have been lost by their custodians (Kelso 2018). Based on the current bitcoin price of approximately $10,150, this represents approximately $65,400,000,000 in monetary value that has been captured then lost, as without these private keys to sign a transaction transferring (i.e. spending) these coins, according to the Bitcoin protocol, they can never be used. Furthermore, many account holders store their bitcoin for long periods, rarely accessing the funds.

Could that value, captured in the blockchain, somehow be put to use? As long as holders' coins are always, of course, available for the key holder to spend, as that is a requirement of the system[176] - could that value, lying fallow in the system, generate more value for users, for humanity? It seems clear that this would need to exist at the protocol level, but could leverage a critical difference between physical and conceptual objects, especially the ability for conceptual objects to be in two places at once.

[175] Which means private keys exist in the keyspace that could sign a transaction specifying their transfer (Quora 2014).
[176] Of course the problem is what if suddenly all "lost" keys submitted transactions at once?

Appendix 7: Initial observations[177] on conceptual reality

Observations on conceptual reality

The unmanifest conceptual space is ever present, global[178] and infinite.

The meaning contained within the universe is infinite at every point through time and space[179].

Without sentience conceptual reality does not manifest. Within a sentient entity's private awareness, it does - to the depth that the awareness perceives[180].

Manifestations within conceptual reality are caused by the perception of some data. They do not manifest without this physical cause[181].

Every object has a subject and every subject has an object[182]. Without either the entity would not exist[183].

[177] These are intuitions and beliefs, stated because they provide a paradigm that resolves some previously unresolved issues in this author's worldview. Inspired in no small part by Wilber (2001).
[178] i.e. universal.
[179] Put another way, the informational potential of the unmanifest conceptual space is infinite.
[180] In this author's definition, digital computers are sentient and possess an awareness.
[181] So it appears ... See Footnote 193 for treatment of conception vs perception.
[182] Consider: my body is an object - matter and energy in space. My subject is the meaning about that object - my name, my experiences, my ideals, my aspirations and so on - my self. To complicate matters, it appears that subjects may also be objects in their own right, with subjects of their own ...
[183] This may be because object classes only exist in the conceptual space; sameness only exists within the conceptual space. This claim derives from the recognition that physical reality is entirely heterogeneous: every point in the universe, at every instant, is perfectly unique. Only in interpretation is comparison possible and similarity ascribed. Since discernment of the boundaries required for the identification of discrete objects is necessarily subjective, it must take place within the perceiving entity's subjective awareness (i.e. conceptual space).

Data is tangible and manifests in physical space; information is intangible and manifests in conceptual space.

Data[184] can be stored in matter or transmitted as energy[185].

Data carries informational potential[186] about the qualities of its source when it was emitted, reflected or refracted.

[184] Note: here, "data" refers to analogue data. Analogue data exists on a continuous spectrum; digital information exists on a discrete spectrum. It is only within awareness that the boundaries exist necessary for discrete entities to arise; thus, digital information only exists within an informational entity's awareness. By this reasoning, it is possible that "digital data" is a misnomer. This digital information is, as described in the Literature Review, encoded onto an analogue signal. Investigation of the differences between analogue and digital holds promise for explaining the nature of conceptual reality. Taking light as an example, a key distinction between analogue and digital data is that, in analogue, meaning is typically conveyed by the two dimensional spatial configuration of individual photons with varying qualities changing over time. Our retinas detect these relative variations - dark and light, or colors - and images are discerned. This is what was meant by "parallel channel", Footnote 16. Digital information can be carried on a one-dimensional analogue channel: a stream of single photons, as referenced in Footnote 151.

[185] A clear distinction is necessary between static data - matter, often intentionally inscribed with symbols - and dynamic data, traveling through space borne on a carrier wave. To come into contact with a sensory organ it seems it must be transmitted on such a wave. The ink markings on a book page serve as an example of static data. To be read (perceived), they must be carried to the reader's eye by light - dynamic data. It appears that dynamic data is very closely related to - or perhaps just is - energy. It is emitted or reflected by a source and travels through some "'ether' … the hypothetical invisible medium that permeates the universe" (Buterin 2014b). This "ether" includes whatever it is light travels through; air, as the auditory channel carries data; as well as conductive materials, which can carry data on an electrical signal. Data can also be conveyed through two objects of mass physically contacting each other - Footnote 188.

[186] Contingent upon sensing, perception and interpretation within the awareness of an informee.

1. Data must come into contact with the sensory organ[187] of a perceiving entity in order to be sensed[188].

2. If sensed, by focusing its attention[189] on the incoming data the informational potential carried can become information and interpreted by the informee as meaning[190]. This is qualia[191].

3. Within the conceptual space in the informational entity's awareness[192], if its attention is placed upon the signal carrying the data into its organism[193], perception can occur[194].

4. As data is perceived, information[195] is created, and can be discerned, interpreted and acted upon.

[187] In biological organisms: eye, ear, tongue, skin, cell membrane; in synthetic informational organisms, antenna or conductive connection with the signal origin, or empirical sensor.
[188] A sleeping person is still an informational entity: it can detect a signal - say, a stick poking them or a loud noise - and awaken.
[189] Probably a key to understanding all of this ...
[190] It appears that information only exists within the awareness of a perceiving entity. Physical data holds the *potential* for meaning (information), but it is only upon sensing and perception that this information comes into existence. This trait - that information only exists within the manifest conceptual space (i.e. the awareness) of informational entities - means that it is a necessarily subjective phenomenon.
[191] Generally, qualia is the process of data entering a sentient entity's awareness and becoming information - the objective-subjective interface. Because the physical universe is perfectly unique, each instant of qualia is unique (and infinite), a moment in your lived experience where the universe pours into you. It is by conceptualizing these qualia that we discern information and ascribe meaning to our perceived reality.
[192] Which is local, private and manifest.
[193] Here, focused on perception of external events. It seems likely that the conception of internal events must arise from the detection of internal data entering awareness.
[194] Section 3.2.1 refers to the process of machine qualia in which analogue data is perceived as binary information through the machine's empirical sensor.
[195] A type of conceptual object?

References

@santisiri. (2019). *santi.* Twitter [online]. Available at https://twitter.com/santisiri/status/1163779311221923840 [Accessed 21 August 2019].

@santisiri. (2019b). *santi.* Twitter [online]. Available at https://twitter.com/santisiri/status/1155139202318602240 [Accessed 27 August 2019].

@Snowden. (2019). *Edward Snowden*. Twitter. [online] Available at https://twitter.com/Snowden/status/1165096076799619073

@VitalikButerin. (2018). *Vitalik Non-giver of Ether on Twitter*. Twitter [online]. Available at https://twitter.com/VitalikButerin/status/1051160932699770882 [Accessed 19 August 2019].

Allen, C. (2016). *The Path to Self-Sovereign Identity*. Life with Alacrity [online]. Available at http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html [Accessed 10 May 2019].

Allen, C. et al. (2015). *Decentralized Public Key Infrastructure*. Github [online]. Available at https://github.com/WebOfTrustInfo/rwot1-sf/blob/master/draft-documents/Decentralized-Public-Key-Infrastructure-CURRENT.md [Accessed 7 May 2019].

Anderson, Ross. (2003). *'Trusted Computing' Frequently Asked Questions*. [online]. Available at https://www.cl.cam.ac.uk/~rja14/tcpa-faq.html [Accessed 24 August 2019].

Antonopoulos, A. (2017). *Mastering Bitcoin*. Github [online]. Available at https://github.com/bitcoinbook/bitcoinbook [Accessed 12 August 2019].

Antonopoulos, A. (2017b). *Money as a System-of-Control*. Available at https://www.youtube.com/watch?v=FyK4P7ZdOK8&vl=en [Accessed 20 April 2019].

Antonopoulos, A. (2018). *Mastering Ethereum.* Github [online]. Available at https://github.com/ethereumbook/ethereumbook [Accessed 10 October 2019].

Antonopoulos, A. (2019). *Coinscrum presents :: Andreas M Antonopoulos :: 13/06/19*. Youtube [online]. Available at https://www.youtube.com/watch?v=LlvuKoXRrsU [Accessed 6 June 2019].

Aragon. (2019). *Aragon Network*. Github [online]. Available at https://github.com/aragon/whitepaper [Accessed 20 June 2019].

Babbage, C. (1822). *A Note Respecting the Application of Machinery to the Calculation of Astronomical Tables*. [online]. Available at http://cyn.io/charles-babbage-a-note-respecting-the-application-of-machinery-to-the-calculation-of-astronomical-tables/ [Accessed 18 August 2019].

BBC. (2017). *CryptoKitties craze slows down transactions on Ethereum.* BBC News [online]. Available at https://www.bbc.com/news/technology-42237162 [Accessed 22 August 2019].

Beal, V. (2019). *DDoS attack - Distributed Denial of Service*. [online]. Available at https://www.webopedia.com/TERM/D/DDoS_attack.html [Accessed 25 August 2019].

Becha, H. (2019). *Smart Containers: Real-time Smart Container data for supply chain excellence*. UNECE - UN / CEFACT [online]. Available at https://www.unece.org/fileadmin/DAM/cefact/GuidanceMaterials/WhitePapers/WP-SmartContainers_Eng.pdf [Accessed 20 June 2019].

Benet, Juan. (2014). *IPFS - Content-Addressed, Versioned, P2P File System*. Github [online]. Available at https://github.com/ipfs/papers/raw/master/ipfs-cap2pfs/ipfs-p2p-file-system.pdf [Accessed 21 March 2019].

Berners-Lee, T. (1990). *WorldWideWeb: Proposal for a HyperText Project*. W3.org [online]. Available at https://www.w3.org/Proposal.html [Accessed 21 August 2019].

Bevir, M. (2012). *Governance: A Very Short Introduction*. Oxford: Oxford University Press.

Bilbow, Angela. (2019). *High Court tackles Bitcoin 'property' first*. Available at https://www.cdr-news.com/categories/litigation/10003-high-court-tackles-bitcoin-property-first [Accessed 24 August 2019].

Bitvalley. (2018). *IBISA: Disrupting Agriculture Insurance*. Bitvalley [online]. Available at https://www.ibisa.network/ [Accessed 13 March 2019].

Boyd-Rice, J. (2018). *New A.I. application can write its own code*. Futurity [online]. Available at https://www.futurity.org/artificial-intelligence-bayou-coding-1740702/ [Accessed 23 August 2019].

Braendgaard, P. (2016). *Simple Convention for Human Readable Terms for Smart Contracts*. Stake Ventures [online]. Available at https://blog.stakeventures.com/articles/smart-contract-terms [Accessed 26 August 2019].

Burns, R. (2004). *Communications: An International History of the Formative Years.* The Institution of Electrical Engineers [online]. Available at https://books.google.co.uk/books?id=7eUUy8-VvwoC&pg=PA29&redir_esc=y#v=onepage&q&f=false [Accessed on 14 August 2019].

Buterin, V., Weyl, G. (2018). *Liberation Through Radical Decentralization*. Medium [online]. Available at https://medium.com/@VitalikButerin/liberation-through-radical-decentralization-22fc4bedc2ac [Accessed 26 August 2019].

Buterin, V. (2013). *A Next-Generation Smart Contract and Decentralized Application Platform*. Ethereum [online]. Available at https://github.com/ethereum/wiki/wiki/White-Paper [Accessed 10 October 2018].

Buterin, V. (2014b). *Ethereum Community Forum*. [online]. Available at https://forum.ethereum.org/discussion/comment/3389/#Comment_3389 [Accessed 22 August 2019].

Buterin, V. (2016). *A Proof of Stake Design Philosophy*. Medium. [Accessed at https://medium.com/@VitalikButerin/a-proof-of-stake-design-philosophy-506585978d51]

Buterin, V. (2017). *The Meaning of Decentralization*. Medium [online]. Available at https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274 [Accessed 14 August 2019].

Carney, M. (2019). *The Growing Challenges for Monetary Policy in the current International Monetary and Financial System*. Bank of England [online]. Available at https://www.bankofengland.co.uk/speech/2019/mark-carney-speech-at-jackson-hole-economic-symposium-wyoming [Accessed 27 August 2019].

Cerf, V.G., Kahn, R.E. (1974). *A Protocol for Packet Network Intercommunication 13.* IEEE Trans on Comms, Vol Com-22 No 5 [online]. Available at https://www.cs.princeton.edu/courses/archive/fall06/cos561/papers/cerf74.pdf. [Accessed 20 August 2019].

Chen, J. (2019). *Agency Problem.* Investopedia [online]. Available at https://www.investopedia.com/terms/a/agencyproblem.asp [Accessed 21 August 2019].

Chen, R., Ramsundar, B., Robbins, R. (2019). *Fair value and decentralized governance of data*. Computable. [online] Available at https://github.com/computablelabs/computable/blob/master/whitepaper/computable_whitepaper.pdf [Accessed 26 August 2019].

Choudhury, N. (2014). World Wide Web and its Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies*, Vol. 5 (6) , 2014, 8096-8100.

Clack, C. D., Bakshi, V., Braine, L. (2016). Smart Contract Templates: foundations, design landscape and research directions. *Barclays Bank PLC* [online]. Available at https://arxiv.org/pdf/1608.00771.pdf [Accessed 26 August 2019].

Cohen, B. (2003). *Incentives build robustness in bittorrent*. In Workshop on Economics of Peer-to-Peer systems, volume 6, pages 68–72.

Corda. (2019). *Corda.* [online]. Available at https://www.corda.net/ [Accessed 10 July 2019].

Core Mesa Team. (2019). *Mesa: Agent-based modeling in Python 3*+. Available at https://github.com/projectmesa/mesa [Accessed 10 March 2019].

CrashCourse. (2017). *Early Computing: Crash Course Computer Science #1*. Youtube [online]. Available at https://www.youtube.com/watch?v=O5nskjZ_GoI [Accessed 10 July 2019].

CrashCourse. (2017b). Electronic Computing: Crash Course Computer Science #2. Youtube [online]. Available at https://www.youtube.com/watch?v=LN0ucKNX0hc [Accessed 10 July 2019].

Daemen, J., n.d. (1999). *The Rijndael Block Cipher 47.* Available at csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf#page=1 [Accessed 19 August 2019].

DAOstack. (2019). *An Operating System for Collective Intelligence*. DAOstack [online]. Available at https://daostack.io/wp/DAOstack-White-Paper-en.pdf [Accessed 20 February 2019].

DiNucci, D. (1999). *Fragmented Future*. Print [online]. Available at https://web.archive.org/web/20111110143942/http://darcyd.com/fragmented_future.pdf [Accessed 19 August 2019].

Dume, B. (2012). *Photon shape could be used to encode quantum information*. Ultrafast Science [online]. Available at https://physicsworld.com/a/photon-shape-could-be-used-to-encode-quantum-information/ [Accessed 17 August 2019].

Ehrsam, F. (2017). *Blockchain Governance: Programming Out Future*. Medium [online]. Available at https://medium.com/@FEhrsam/blockchain-governance-programming-our-future-c3bfe30f2d74 [Accessed 10 December 2017].

Engel, J. (2019). *GANSynth: Making music with GANs*. magenta [online]. Available at https://magenta.tensorflow.org/gansynth [Accessed 10 April 2019].

Essinger, J. (2018). *Ada Lovelace: a visionary of computing*. History Extra [online]. Available at https://www.historyextra.com/period/modern/ada-lovelace-stem-women-computing-science-facts-life/ [Accessed 15 August 2019].

Ethereum. (2019b). *On sharding blockchains.* Github [online]. Available at https://github.com/ethereum/wiki/wiki/Sharding-FAQ [Accessed 26 August 2019].

Ethereum Foundation. (2019). *Design Rationale*. Github. [Accessed at https://github.com/ethereum/wiki/wiki/Design-Rationale]

Etherscan.io. (2019). *Ethereum Block Time History*. Etherscan. [online] Available at https://etherscan.io/chart/blocktime [Accessed 20 August 2019].

Fahie, J. J. (1884). *A History of Electric Telegraphy to the Year 1837.* E. & F. N. Spon: London [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=2ahUKEwjn4JHy-qPkAhVkSxUIHUKzAYQQFjADegQIARAC&url=https%3A%2F%2Fwww.princeton.edu%2Fssp%2Fjoseph-henry-project%2Ftelegraph%2FA_history_of_electric_telegraphy_to_the.pdf&usg=AOvVaw20l3iQsyg_Avz0JrT9l2uv [Accessed 14 August 2019].

Fecke, M. (2018). *The Problem of Blockchain Oracles - Interview with Alexander Egberts*. Legal Tech Blog [online]. Available at

https://legal-tech-blog.de/the-problem-of-blockchain-oracles-interview-with-alexander-egberts [Accessed 21 August 2019].

Fernandez, M., Sanger, D. E., Martinez, M. T. (2019). *Ransomware Attacks Are Testing Resolve of Cities Across America*. The New York Times [online]. Available at https://www.nytimes.com/2019/08/22/us/ransomware-attacks-hacking.html [Accessed 23 August 2019].

Floridi, L. (2010). *Information: A Very Short Introduction*. Oxford: Oxford University Press.

Fuegi, J., Francis, J. (2003). *Lovelace & Babbage and the Creation of the 1843 'Notes'.* IEEE Annals of the History of Computing [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiPgu34_aPkAhWltXEKHeLDB6AQFjAAegQIABAC&url=https%3A%2F%2Fwww.scss.tcd.ie%2Fcoghlan%2Frepository%2FJ_Byrne%2FA_Lovelace%2FJ_Fuegi_%26_J_Francis_2003.pdf&usg=AOvVaw1H1JvrMxvmSNoGxeoMJ7-t [Accessed 15 August 2019].

Galperin, E., Hassine, W. B. (2015). *Changes to Facebook's "Real Names" Policy Still Don't Fix the Problem*. EFF [online]. Available at https://www.eff.org/deeplinks/2015/12/changes-facebooks-real-names-policy-still-dont-fix-problem [Accessed 23 August 2019].

Gentry, C. (2010). *Computing Arbitrary Functions of Encrypted Data*. [online]. Available at https://dl.acm.org/citation.cfm?id=1666444 [Accessed 21 August 2019].

Gilbert, N. (2008). *Agent-based models.* London: Sage Publications.

git-scm.com. (2019). *Git.* [online]. Available at: https://git-scm.com/. [Accessed 22 August 2019].

Glaser, A., Barak, B., Goldston, R. (2014). *A zero-knowledge protocol for nuclear warhead verification*. Nature [online]. Available at https://www.nature.com/articles/nature13457 [Accessed 24 August 2019].

Gleick , J. (2011). *The Information: A History, a Theory, a Flood*. Fourth Estate: London.

Grigg, I. (2004). The ricardian contract. In *Proceedings of the First IEEE International Workshop on Electronic Contracting*, pages 25–31. IEEE. http://iang.org/papers/ricardian_contract.html.

Gupta, V., Knight, R., Buchanan, A., Wray, C., Grigg, I., Kuhlman, C., Cimpoesu, M., Mainelli, M., Freedman, C. (2019). *Mattereum Working Paper*. Mattereum [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiZlObctp7kAhX8RxUIHcj7C0QQFjAAegQIARAC&url=https%3A%2F%2Fmattereum.com%2Fupload%2Fiblock%2Faf8%2Fmattereum_workingpaper.pdf&usg=AOvVaw1ZdsGE1x8Ks65p8Y_QEmXR [Accessed 2 February 2019].

Gutierrez, D. (2016). *Why Time-Value of Data Matters*. insideBIGDATA [online]. Available at https://insidebigdata.com/2016/04/08/why-time-value-of-data-matters/ [Accessed 26 August 2019].

gwei.io. (2019). *GWEI.IO*. [online]. Available at https://gwei.io/ [Accessed 18 August 2019].

Harris, M. (2019). *Pentagon testing mass surveillance balloons across the US*. The Guardian [online]. Available at https://www.theguardian.com/us-news/2019/aug/02/pentagon-balloons-surveillance-midwest [Accessed 23 August 2019].

Haselton, T. (2017). *Credit reporting firm Equifax says data breach could potentially affect 143 million US consumers.* CNBC [online]. Available at https://www.cnbc.com/2017/09/07/credit-reporting-firm-equifax-says-cybersecurity-incident-could-potentially-affect-143-million-us-consumers.html [Accessed 23 August 2019].

Heller, J. (1961). *Catch-22*. London: Vintage.

Hodgson, G. (2012). *A revisionist critique of fractional reserve banking*. Positive Money. [online] Available at https://positivemoney.org/2012/01/revisionist-critique-fractional-reserve-banking/ [Accessed 27 August 2019].

Hoopes, J. (2019). *Reimagining "global": Programmable incentivization and implications for personal governance*. Github [online]. Available at https://github.com/robisoniv/rwot9-prague/blob/master/topics-and-advance-readings/reimagining-global-rwot9.md [Accessed 19 August 2019].

Hopkins, N. (2017). *Deloitte hit by cyber-attack revealing clients' secret emails.* The Guardian [online]. Available at https://www.theguardian.com/business/2017/sep/25/deloitte-hit-by-cyber-attack-revealing-clients-secret-emails. [Accessed 23 August 2019].

Huang, A. (2017). *The Origins of Computing: The Sumerians,, The Abacus & Binary Code*. Medium [online]. Available at https://medium.com/@Ashley_Huang/the-origins-of-computing-the-sumerians-the-abacus-binary-code-15289c3ced16 [Accessed 1 August 2019].

Hunter, J. D. (2007). *Matplotlib: A 2D Graphics Environment.*, Computing in Science & Engineering, 9, 90-95, DOI:10.1109/MCSE.2007.55

Ibarra, I. A., Goff, L., Hernández, D. J., Lanier, J., Weyl, E. G. (2018). *Should We Treat Data as Labor? Moving Beyond "Free"*. Data as Labor Papers and Proceedings. [Available at https://ssrn.com/abstract=3093683]

Investopedia. (2019). *Fractional Reserve Banking*. Investopedia [online]. Available at https://www.investopedia.com/terms/f/fractionalreservebanking.asp [Accessed 24 August 2019].

jnnk. (2015). *What is Gas Limit in Ethereum?* Stackexchange.com [online]. Available at https://bitcoin.stackexchange.com/questions/39132/what-is-gas-limit-in-ethereum [Accessed 21 August 2019].

Kelso, C. E. (2018). *BTC: 36% in Circulation Lost, 23% Held by Speculators, US Tax Authority Monitoring*. Bitcoin.com. [online] Available at

https://news.bitcoin.com/btc-36-in-circulation-lost-23-held-by-speculators-us-tax-authority-monitoring/ [Accessed 27 August 2019].

Ker, Andrew D. (2014). *Computer Security*. Oxford: Oxford University Department of Computer Science, [online]. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjQ-ZH636DkAhXZTxUIHfVBD_EQFjAAegQIBBAC&url=http%3A%2F%2Fwww.cs.ox.ac.uk%2Fandrew.ker%2Fdocs%2Fcomputersecurity-lecture-notes-mt2014.pdf&usg=AOvVaw2reitUz7E6ktrlCORJ2vj5 [Accessed 20 October 2018].

Kerckhoffs, A. (1883). *La Cryptographie Militaire (première partie) 37*. Available at https://www.petitcolas.net/kerckhoffs/crypto_militaire_1_b.pdf; translation at https://www.petitcolas.net/kerckhoffs/index.html [Accessed 17 August 2019].

Kim, C. (2019). *Code For Ethereum's Proof-of-Stake Blockchain to Be Finalized Next Month*. Coindesk [online]. Available at https://www.coindesk.com/code-for-ethereums-proof-of-stake-blockchain-to-be-finalized-next-month [Accessed on 20 August 2019].

Kleinrock, Leonard. (1961). *Information Flow in Large Communication Nets*. Massachusetts Institute of Technology Research Laboratory of Electronics [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiE3LaF4JbkAhWoTxUIHTGDDrwQFjAAegQIAhAC&url=https%3A%2F%2Fwww.lk.cs.ucla.edu%2Fdata%2Ffiles%2FKleinrock%2FInformation%2520Flow%2520in%2520Large%2520Communication%2520Nets.pdf&usg=AOvVaw11sqIM4l6SD1VAaJbM3ODH [Accessed 20 August 2019].

Krebs, Brian. (2019). *Facebook Stored Hundreds of Millions of User Passwords in Plain Text for Years.* Krebs on Security [online]. Available at https://krebsonsecurity.com/2019/03/facebook-stored-hundreds-of-millions-of-user-passwords-in-plain-text-for-years/ [Accessed 22 August 2019].

Lee, D. (2018). *Facebook security breach: Up to 50m accounts attacked*. BBC, [online]. Available at https://www.bbc.com/news/technology-45686890 [Accessed 26 August 2019].

Leiner, B., Cerf, V., Clark, D. D., Kahn, R. E., Kleinrock, L, Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. (1997). *Brief History of the Internet.* Internet Society [online]. Available at https://www.internetsociety.org/wp-content/uploads/2017/09/ISOC-History-of-the-Internet_1997.pdf [Accessed 20 August 2019].

Licklider, J. C. R. (1968). *The Computer as a Communication Device*. Science and Technology [online]. Available at http://memex.org/licklider.pdf [Accessed 20 August 2019].

Liu, C. (2008). *The Dark Forest*. Tor, A Tom Doherty Associates Book.

Living Internet. (2019). *J. C. R. Licklider And The Universal Network.* Living History [online]. Available at https://www.livinginternet.com/i/ii_licklider.htm [Accessed 14 August 2019].

Lloyd's. (2019). *Triggering innovation: How smart contracts bring policies to life*. Queen Mary Centre for Commercial Law Studies [online]. Available at https://www.lloyds.com/news-and-risk-insight/risk-reports/library/technology/triggering-innovation [Accessed 11 August 2019].

Long, C. (2019). *Bitcoin, The Dollar And Facebook's Cryptocurrency: Price Volatility Versus Systemic Volatility*. Forbes [online]. Available at https://www.forbes.com/sites/caitlinlong/2019/06/29/bitcoin-the-dollar-and-facebooks-cryptocurrency-price-volatility-versus-systemic-volatility/#1aa3f14d88b8 [Accessed 10 July 2019].

Lovelace, A. (1843). *Notes on L. Menabrea's 'Sketch of the Analytical Engine Invented by Charles Babbage, Esq.'*. Taylor's Scientific Memoirs, *3*(1843), p.1843.

Malan, D. (2014). *Lecture 0: Introduction to Computer Science I*. Youtube [online]. Available at: https://www.youtube.com/watch?v=z-OxzIC6pic [Accessed 13 July 2016].

Malik, N. (2008). *The Tradeoff Between Security And Privacy: How do Terrorists Use Encryption?* Forbes [online]. Available at https://www.forbes.com/sites/nikitamalik/2018/11/07/the-tradeoff-between-security-and-privacy-how-do-terrorists-use-encryption/#368d68862d8c [Accessed 17 August 2019].

Martin, Ernst. (1992). *The Calculating Machines: Their History and Development*. The MIT Press [online]. Available at www.rechenmaschinen-illustrated.com/Martins_book/Ernst%20Martin%20-%20Rechen%20Machinen%20OCR%204.pdf [Accessed 17 August 2019].

McCoy, M. (2015). *6 Notorious Cases of Data Loss All Hosting Providers Can Learn From*. R1 Blog [online]. Available at https://www.r1soft.com/blog/6-notorious-cases-of-data-loss-all-hosting-providers-can-learn-from [Accessed 23 August 2019].

McDonald, J. (2017). *Releasing Stuck Ethereum Transactions*. Medium [online]. Available at https://medium.com/@jgm.orinoco/releasing-stuck-ethereum-transactions-1390149f297d [Accesseed 25 August 2019].

McKinney, Wes. (2010). *Data Structures for Statistical Computing in Python.* Proceedings of the 9th Python in Science Conference, 51-56.

Merkle, R. (2016). *DAOs, Democracy and Governance*. [online]. Available at merkle.com/papers/DAOdemocracyDraft.pdf [Accessed 28 August 2019].

Merritt, R. (2003). *New group aims to secure PCs, PDAs, cell phones*. [online]. Available at https://www.eetimes.com/document.asp?doc_id=1202119 [Accessed 20 August 2019].

Moriya, H. (2018). *How to get Ethereum Block Gas Limit*. Medium. [online] Available at https://medium.com/@piyopiyo/how-to-get-ethereum-block-gas-limit-eba2c8f32ce [Accessed 10 July 2019].

Mullins, R. (2012). *What is a Turing machine?* University of Cambridge Department of Computer Science and Technology [online]. Available at https://www.cl.cam.ac.uk/projects/raspberrypi/tutorials/turing-machine/one.html [Accessed 12 August 2019].

Nair, B. Somanathan. (2002). *Digital Electronics and Logic Design*. Prentice Hall [online]. Available at https://books.google.co.uk/books?id=WK45wLHL-ycC&pg=PA289&dq=%22digital+signals+are%22+%22digital+electronics%22&hl=sv&sa=X&redir_esc=y#v=onepage&q=%22digital%20signals%20are%22%20%22digital%20electronics%22&f=false [Accessed 16 August 2019].

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Bitcoin.org [online]. Available at https://bitcoin.org/bitcoin.pdf [Accessed 23 October 2018].

Nakamoto, S. (2009). *Bitcoin v0.1 released.* The Cryptography Mailing List [online]. Available at https://www.metzdowd.com/pipermail/cryptography/2009-January/014994.html. [Accessed 18 August 2019].

National Institute of Standards and Technology. (2001). *FIPS 197, Advanced Encryption Standard (AES)*, *51*. Available at nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf [Accessed 19 August 2019].

Node.js Foundation. (2018). *Node.js*. [online] Available at: https://nodejs.org/en/. [Accessed 28 January 2018].

Norman, J. (2019). *Francis Ronalds Builds the First Working Electric Telegraph*. Jeremy Norman's HistoryofInformation.com [online]. Available at http://historyofinformation.com/detail.php?entryid=519 [Accessed 15 August 2019].

Nuñez, D. (2018). *Umbral: A Threshold Proxy Re-encryption Scheme*. NICS Lab, University of Malaga, Spain [online]. Available at https://github.com/nucypher/umbral-doc/blob/master/umbral-doc.pdf [Accessed 1 March 2019].

Oliphant, T. E. (2006). *A guide to NumPy.* USA: Trelgol Publishing. Available at https://www.numpy.org/

Orcutt, M. (2017). *A Mind-Bending Cryptographic Trick Promises to Take Blockchains Mainstream.* MIT Technology Review [online]. Available at https://www.technologyreview.com/s/609448/a-mind-bending-cryptographic-trick-promises-to-take-blockchains-mainstream/ [Accessed 26 August 2019].

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. ISBN 978-0-521-40599-7.

Oxford. (2019a). *Cipher*. Lexico [online]. Available at https://www.lexico.com/en/definition/cipher [Accessed 20 August 2019].

Oxford. (2019b). *Algorithm*. Lexico [online]. Available at https://www.lexico.com/en/definition/algorithm [Accessed 20 August 2019].

Paul, M. S. (2018). *Hyperledger — Chapter 3| When to use the Blockchain Technology*. The Startup [online]. Available at https://medium.com/swlh/hyperledger-chapter-3-when-to-use-the-blockchain-technology-a5c414221bdf [Accessed 27 August 2019].

Peaster, W. (2018). *FCoin GPM Listing Competition Acutely Clogs Ethereum*. Bitsonline [online]. Available at https://bitsonline.com/fcoin-gpm-listing-cloggage/ [Accessed 23 August 2019].

Pérez, Fernando, Granger, Brian E. (2007). *IPython: A System for Interactive Scientific Computing*. Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, doi:10.1109/MCSE.2007.53. Available at https://ipython.org.

phant0m. (2013). *What does the ^ (XOR) operator do?* Stackoverflow.com [online]. Available at https://stackoverflow.com/questions/14526584/what-does-the-xor-operator-do [Accessed 19 August 2019].

Prince, M. (2019). *Terminating Service for 8chan*. Cloudflare [online]. Available at https://blog.cloudflare.com/terminating-service-for-8chan/ [Accessed 23 August 2019].

Protocol Labs. (2019b). Filecoin FAQs. [online] Available at https://filecoin.io/faqs/ [Accessed 26 August 2019].

Protocol Labs. (2019). *Content Identifiers (CIDs).* Accessed at https://docs.ipfs.io/guides/concepts/cid/]

Quinn, J. (2008). *Greenspan admits mistakes in 'once in a century credit tsunami'*. The Telegraph [online]. Available at https://www.telegraph.co.uk/finance/financialcrisis/3248774/Greenspan-admits-mistakes-in-once-in-a-century-credit-tsunami.html [Accessed 20 August 2019].

Quora. (2014). *What would happen if I tried to send BTC to a nonexistent bitcoin address*. Quora. [online] Available at https://www.quora.com/What-would-happen-if-I-tried-to-send-BTC-to-a-nonexistent-bitcoin-address [Accessed 27 August 2019].

Quorum. (2019). *Quorum.* [online]. Available at https://www.goquorum.com/ [Accessed 26 August 2019].

Raymond, E. (1999). *Release Early, Release Often*. [online]. Available at http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/ar01s04.html [Accessed 14 March 2017].

Rea, A., Fischer, A., du Rose, J. (2019). *Colony Technical White Paper*. Colony [online]. Available at https://colony.io/whitepaper.pdf [Accessed 20 August 2019].

Rouse, M. (2019). *Access control.* [online]. Available at https://searchsecurity.techtarget.com/definition/access-control [Accessed 18 August 2019].

Ryan, D. (2017). *Calculating Costs in Ethereum Contracts*. Hackernoon [online]. Available at https://hackernoon.com/ether-purchase-power-df40a38c5a2f [Accessed 3 July 2019].

Sabanal, P. (2016). *Thingbots: The Future of Botnets in the Internet of Things*. SecurityIntelligence [online]. Available at https://securityintelligence.com/thingbots-the-future-of-botnets-in-the-internet-of-things/ [Accessed on 21 August 2019].

Safford, D. (2003). *Take Control of TCPA*. Linux Journal [online]. Available at https://www.linuxjournal.com/article/6633 [Accessed 23 August 2019].

Seabold, S., Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python*. Proceedings of the 9th Python in Science Conference.

Shannon, C. (1949). *Communication Theory of Secrecy Systems*. Bell System Technical Journal [online]. Available at https://archive.org/stream/bstj28-4-656#page/n5/mode/2up [Accessed 21 August 2019].

Shannon, C. E., (1938). *A symbolic analysis of relay and switching circuits*. Trans. Am. Inst. Electr. Eng. 57, 713–723. https://doi.org/10.1109/T-AIEE.1938.5057767

Shannon, C. E. (1948). *A Mathematical Theory of Communication*. The Bell System Technical Journal Vol XXVII No. 3 pp 379-423.

Shirer, M, MacGillivray, C. (2019). *The Growth in Connected IoT Devices Is Expected to Generate 79.4ZB of Data in 2025, According to a New IDC Forecast*. [Accessed at https://www.idc.com/getdoc.jsp?containerId=prUS45213219]

Simonsen, S. (2018). *Quantum Communication Just Took a Great Leap Forward*. Singularity Hub [online]. Available at https://singularityhub.com/2018/12/26/quantum-communication-just-took-a-great-leap-forward/ [Accessed 27 August 2019].

Stackoverflow. (2016). *Is there a limit for sha256 input?* Stackoverflow.com [online]. Available at https://stackoverflow.com/questions/17388177/is-there-a-limit-for-sha256-input [Accessed 27 August 2019].

Stankovic, R., Astola, J. (2008). *Reprints from the Early Days of Information Sciences: On the Contributions of Akira Nakashima to Switching Theory*. Tampere International Center for Signal Processing [online]. Available at http://ticsp.cs.tut.fi/index.php/Reprints_from_the_Early_Days_of_Information_Sciences.html [Accessed 16 August 2019].

State of Victoria. (2017). *What is sentience?* Victoria State Government [online]. Availabe at http://agriculture.vic.gov.au/pets/care-and-welfare/animals-and-people/what-is-sentience [Accessed 26 August 2019].

Szabo, N. (1994). *Smart Contracts*. Nick Szabo [online]. Available at http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html [Accessed on 17 June 2019].

Taleb, N. N. (2013). *Antifragile: Things That Gain From Disorder*. Penguin.

Techopedia. (2019). *Cryptographic Key*. technopedia [online]. Available at https://www.techopedia.com/definition/24749/cryptographic-key [Accessed 20 August 2019].

The Economist. (2017). *A clever way to transmit data on the cheap.* The Economist [online]. Available at https://www.economist.com/science-and-technology/2017/09/16/a-clever-way-to-transmit-data-on-the-cheap [Accessed 16 August 2019].

The Economist. (2019). *Schumpeter: The Exxon-Valdez of Cyberspace*. Available at https://www.economist.com/business/2019/08/08/the-exxon-valdez-of-cyberspace. Accessed 24 August 2019].

The Linux Foundation. (2019). *Hyperledger.* [online]. Available at https://www.hyperledger.org/ [Accessed 20 May 2019]

Thunberg, G. (2019). *No One is Too Small to Make a Difference*. UK: Penguin.

Toole, B. A. (1998). *Ada, The Enchantress Of Numbers: Prophet Of The Computer Age*. Strawberry Press [online]. Available at http://www.cs.yale.edu/homes/tap/Files/ada-lovelace-notes.html [Accessed 16 August 2019].

Torvalds, L. (2005). *Initial revision of "git", the information manager from hell.* Github [online]. Available at https://github.com/git/git/commit/e83c5163316f89bfbde7d9ab23ca2e25604af290 [Accessed 22 August 2019].

Turing, A. M., (1937). *On Computable Numbers, with an Application to the Entscheidungsproblem*. Proceedings of the London Mathematical Society s2-42, 230–265. https://doi.org/10.1112/plms/s2-42.1.230

van Rossum, G. (1995). *Python tutorial, Technical Report CS-R9526.* Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

Varian, H. R. (1998). *Markets for Information Goods*. University of California, Berkeley. Available at http://people.ischool.berkeley.edu/~hal/Papers/japan/ [Accessed 22 August 2019].

Veness, C. (2019). *SHA-256 Cryptographic Hash Algorithm*. Movable Type Scripts [online]. Available at https://www.movable-type.co.uk/scripts/sha256.html [Accessed 20 August 2019].

Vincent, J. (2019). *Bitcoin consumes more energy than Switzerland, according to new estimate*. The Verge [online]. Available at https://www.theverge.com/2019/7/4/20682109/bitcoin-energy-consumption-annual-calculation-cambridge-index-cbeci-country-comparison [Accessed 18 August 2019].

Vincent, S. (2018). *Willis Towers Watson launches reef insurance project*. insuranceday [online]. Available at https://insuranceday.maritimeintelligence.informa.com/ID1122617/Willis-Towers-Watson-launches-reef-insurance-project [Accessed 25 August 2019].

Vu, T. M. (2015). *Are the differences between Public Good and Common Pool Resource too blurred?* Researchgate [online]. Available at https://www.researchgate.net/post/Are_the_differences_between_Public_Good_and_Common_Pool_Resource_too_blurred [Accessed 25 August 2019].

Vyas, Tanvi. (2016). *No More Passwords over HTTP, Please!* Mozilla [online]. Available at https://blog.mozilla.org/tanvi/2016/01/28/no-more-passwords-over-http-please/ [Accessed 20 August 2019].

Waters, R. (2019). *Three ways that Big Tech could be broken up.* Financial Times, [online]. Available at: https://www.ft.com/content/cb8b707c-88ca-11e9-a028-86cea8523dc2 [Accessed 26 August 2019].

web3.js. (2019). *Callbacks Promises Events — web3.js 1.0.0 documentation*. [online]. Available at: https://web3js.readthedocs.io/en/v1.2.1/callbacks-promises-events.html. [Accessed 21 February 2019].

Weigert, M. (2015). *The surveillance state is inevitable*. Meshed Society [online]. Available at https://meshedsociety.com/the-surveillance-state-is-inevitable/ [Accessed 25 August 2019].

Wikipedia Contributors. (2019a). *Data type*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Data_type [Accessed 27 August 2019].

Wikipedia Contributors. (2019b). *Message*. Wikipedia [online]. Available at: https://en.wikipedia.org/wiki/Message [Accessed 12 August 2019].

Wikipedia Contributors. (2019c). *Homomorphic encryption*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Homomorphic_encryption [Accessed 10 May 2019].

Wikipedia Contributors. (2019d). *Transmission Control Protocol*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Transmission_Control_Protocol [Accessed 19 August 2019].

Wikipedia Contributors. (2019e). *History of free and open-source software*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/History_of_free_and_open-source_software [Accessed 19 August 2019].

Wikipedia Contributors. (2019f). *Zero-knowledge proof*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Zero-knowledge_proof [Accessed 26 August 2019].

Wikipedia Contributors. (2019g). *Government*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Government [Accessed 25 August 2019].

Wikipedia Contributors. (2019h). *Treaty*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Treaty [Accessed 23 August 2019].

Wikipedia Contributors. (2019j). *Checksum*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Checksum [Accessed 24 August 2019].

Wilber, K. (2001). *A Brief History Of Everything*. Gill & Macmillan.

Williams, J. (2018). *Stand Out Of Our Light: Freedom and Resistance in the Attention Economy*. Cambridge University Press [online]. Available at https://www.cambridge.org/core/books/stand-out-of-our-light/3F8D7BA2C0FE3A7126A4D9B73A89415D [Accessed 14 January 2019].

Wood, G. (2014). ĐApps: What Web 3.0 Looks Like. Insights into a Modern World [online]. Available at http://gavwood.com/dappsweb3.html [Accessed 17 August 2019].

Wood, G. (2019). *Ethereum Yellow Paper.* Github [online]. Available at https://github.com/ethereum/yellowpaper [Accessed 15 October 2019].

Wu, H., Wang, F. (2014). A Survey of Noninteractive Zero Knowledge Proof System and Its Applications. *Scientific World Journal* [online]. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032740/ [Accessed 26 August 2019].

Wynn-Williams, C. E. (1931). *The use of thyratrons for high speed automatic counting physical phenomena*. The Royal Society Publishing [online]. Available at https://doi.org/10.1098/rspa.1931.0102. [Accessed 15 August 2019].

Xu, L.D., He, W., Li, S. (2014). *Internet of Things in Industries: A Survey*. IEEE Trans. Ind. Inf. 10, 2233–2243. https://doi.org/10.1109/TII.2014.2300753

Zamfir, V. (2017). *Against on-chain governance*. Medium [online]. Available at https://medium.com/@Vlad_Zamfir/against-on-chain-governance-a4ceacd040ca [Accessed 25 May 2019].

Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M. (2014). *Internet of Things for Smart Cities*. IEEE Internet Things J. 1, 22–32. https://doi.org/10.1109/JIOT.2014.2306328

Acknowledgements

This work would not have been possible without input and support over countless coffees and calls with an incredible group of experts, advisors, mentors and friends who shared their time and invaluable perspectives with me. An incomplete list:

From UCL: my supervisor Dr Duncan Wilson, Dr Sarah Wise, Dr Kira Kempinska, Dr Thomas Oléron-Evans, Kristian Lunow Nielsen, Katie Jamieson, Dagmar Ellefsen, Pili Mayora, Aude Vuilliomenet, Katherine Coates and Dr Abel Maciel.

From the blockchain community - through London Blockchain Labs: Yakko Majuri, Ilya Kleyner, Lorcan Delaney, Alex Terenda, Amber Rignell, Isaac Sultan, Alex Zakharov and Amanda Tan, as well as Jon Geater, Ben Bennett and James Worthington, Anthony Beaumont and Jean-Yves Rotté-Geoffroy. Lawrence Tilli and the BitBoysAndGirls crowd, and Vu Tien Khang.

From the peace, security and development community, Dr Curtis Bell, Dr Conor Seyle, Dr Sarah Glaser, Admiral Sir James Burnell-Nugent, Larry Sampler, Maisie Pigeon, Simon Williams, Dr Kuzi Charamba, John Filitz, Dr Clayton Besaw, Sean Duncan, Greg Clough, Melissa Hanham, Saskia Westhof and Camilo Casas, along with Dr Ian Stewart, Catherine Dill, Simon Ring, Mike Lewis, Adrian Wilkinson, Beth Clark, Mariyana Radeva Berket, and Janne and Andrew Kaiser-Tedesco.

From the London insurance world, Dr.rer.Pol. Magdalena Ramada Sarasola, Walid al Saqqaf, Aqua Sanfelice and Joe Mellen.

And from elsewhere - Jordan Bouley, Olivia Rhodes, Rollie Williams, Nicola Henderson, Zach Gorman, Chris Cote, Brandon Rattiner, Philipp Friemann, Ed Bayes, Lloyd Harrison, Caroline Clark, Lyle Barton, Jess Murray, Anne Milton, Helen Giles, Jen Carswell, Gemma Lyons and Michael Parratt.

And of course my family: Jocelyn, Jack, Emily, Jesse and Darlene. Thank you - 🙏.

‘And anything worth dying for,’ answered the sacrilegious old man, ‘is certainly worth living for.’

Joseph Heller
1961

We can possess nothing—neither thing nor thought—absent a border between self and State. Each individual right is derived from this line, which we call privacy.

@Snowden
24 August 2019

Abstract

As connected sensors capture and process ever greater amounts of information about their local vicinities, growing networks of these devices create the potential for improving situational awareness and, therefore, the efficiency and virtue of resource management. However, due to the technical, commercial and legal realities of these systems in the present day, much of the informational value these networks represent to humanity remains unaccessed. Concurrently, advancements in information technologies - especially cryptography - are enabling the emergence of a decentralized web paradigm designed to preserve individual autonomy and privacy. Most notably, blockchains are emerging to provide the public with transparent, open and neutral information computing and storage resources that operate independently of any central authority, and are highly secure and antifragile. This dissertation seeks to identify and define concepts key to connected sensor networks and the decentralized web. The fundamentals of enabling technologies including computing, cryptography and digital communication are reviewed, and relevant decentralized technologies such as smart contracts, decentralized autonomous organizations, content-based addressing, self-sovereign identities and Ricardian contracts are defined. An agent-based model simulating the operation of connected sensors connecting to a public blockchain network was developed; quantitative results of parameter sweeps of three input parameters are reported and analyzed. Considerations concerning the connection of edge sensors with smart contracts are presented. The technical and political feasibility and ethics of high impact applications of such systems are discussed. Finally, a possible limitation in our current paradigm is contemplated.

Figures

Tables

Statement of Ethics

I engage in this work with an ethical mindset, and hope that this writing conveys that I strive to consider deeply not only the technical, but also the moral, social, economic and political implications of these subjects. Generally, I am oriented toward improving human dignity and supporting us in our role as stewards of the planet.

But unless we recognize the overall failures of our current systems we most probably don’t stand a chance.

Greta Thunberg
2019

… to create a world that better preserves the autonomy of the individual …

Vitalik Buterin
2016

1 Introduction

Objects exist in space. These entities interact over time. We know this because they come into our awareness, moment to moment. We[1] are objects that perceive the world, observing it through these moments while we are alive.

This dissertation is a reporting back of the current state of my understanding of these topics. My overall interests include the patterns and relationships of our lived experience, including the natures of both matter and meaning, and the relationship between the two. My angle is the simplest instance of which I am aware of informational entities communicating: connected sensors - computing nodes interpreting, transmitting and receiving binary data. My goal is to investigate the intricacies of these systems of connected sensors in the current moment - 2019.

The relevance is difficult for me to overstate - I am drawn to pursue this research because I feel it might help us to understand the enormous risks and opportunities inherent in the technologies we are just beginning to invent and discover. My hope is that if we can understand these risks we might be able to mitigate them, just as if we do the same for the opportunities they might be maximized toward the promotion of the values we share. I believe this is a pivotal moment[2] for us, here on Earth today. It seems we have the capacity - and perhaps the capability - to finally achieve what so many just people have spent their lives pursuing: the eradication of violence and the peaceful coordinated thriving of life[3].

The risks of this endeavor cannot be overstated. What I intend to learn will imbue me with great power, the ability to help create the services that people and machines may use to live in the 21st century - an ability that confers moral responsibility. I accept this, but cannot say I understand it - I am only beginning to grasp the ethical implications of this emerging reality. For this reason the work will

[1] As in, us humans.
[2] Perhaps every present moment is a pivotal one? This one seems especially so.
[3] I am not a technological solutionist but I do believe that, as a tool, technology can help us achieve our intentions.

describe a number of ethical principles, adopted or adapted or arrived at in my studies.

Specifically, this research explores the interaction of networks of connected sensors with blockchain networks. In order to adequately define terms and frame the research, the necessary technologies will be reviewed and clear[4] explanations provided. A middle-range agent-based model simulating key features of edge sensors, quasi-Turing-complete smart contract platforms and their interaction was built, as described in the Methodology section. Model runs produced simulation output of parameter sweeps of three independent variables; three dependent variables were measured and analyzed, presented as Results.

In the Discussion I briefly review the model findings, and propose the outlines of a few system configurations that utilize the potential of these systems in a just and inclusive way.

The Conclusion will raise some of the questions I encountered that were not satisfactorily answered in the current scientific paradigm, as I understand it.

A distinction needs to be made, between people that are[5] alive, and everything else; it seems important - paramount - that place higher value in the former. If each person is equal[6], this humanist ideal must be ingrained in us and in everything we create.

[4] Hopefully ...
[5] And will be.
[6] In value, yes. In qualities? Never.

1.1 Research Questions

General

What is information? What is data? How are they related? How are they produced, stored, transported and used? How are they governed?

Specific

What opportunities, risks, constraints and limitations are inherent in the use of edge sensors as oracles for smart contracts?

Agent-based model

What are the effects of network size and data capture and transfer dynamics on public blockchain validation patterns and costs?

2 Literature Review

2.1 Computers

Humans have relied on physical objects - tools - to help perform logical operations since the origin of writing and mathematics (Huang 2017). A common feature of these tools (tally sticks, abacuses, and so on) is their ability to store symbolic representations of numerical values so the operator did not have to remember the values. With the aid of these tools, human understanding of the patterns of mathematics developed.

Leibniz and Babbage[7] paved the way for Alan Turing to create his (now known as) Turing complete “a-machine” (Turing 1937), a mathematical model of computation capable of “simulat[ing] any computer algorithm (Mullins 2012).

On this basis, modern computing arose. Computers apply binary operations to sequences of bits[8], represented in contemporary computers as electrical voltages stored in capacitors. A bit sequence is computed by the processing unit according to its data type[9]; an identical sequence carries different information for the processor can be interpreted and processed differently[10] based on this metainformation (Malan 2014).

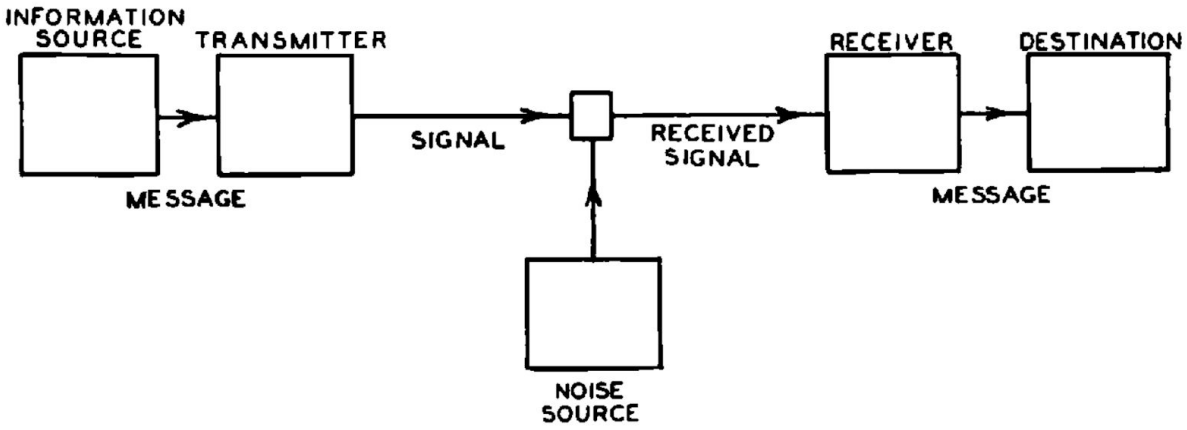
2.2 Communication

Building on knowledge of how information is transmitted on an electrical or broadcast signal[11], as well as the community’s emerging understanding of the behavior of electrical circuits and its implications for data processing and the

[7] See Appendix 5.1 for more on the history of computing.
[8] One byte is 8 bits, one kilobyte is 1000 bytes, and so on, according
[9] I.e. Real, integer, boolean (Wikipedia 2019a).
[10] i.e. different binary operations applied.
[11] See Appendix 5.2 for more on the history of communication.

encoding of information on a serial channel[12], Claude Shannon published his master's thesis[13], *A symbolic analysis of relay and switching circuits* (1938). A decade later, in *A mathematical theory of communication*, Shannon (1948) defined the fundamental principles of the transfer of a message from source[14] to destination[15]: "the theory that lies behind any phenomenon involving data encoding and transmission" (Floridi 2010)[16].

Figure 1: Schematic diagram of a general communication system



Reprinted from Shannon (1948 pp 381)

Shannon's paper described how a binary message is encoded on a signal[17] by the information source, or informer. The signal is transmitted on a (perhaps noisy) channel[18], then detected by the receiver and interpreted by the information destination, the informee.

[12] A one dimensional, serial channel, as opposed to a two dimensional, parallel channel - see Appendix 7.
[13] Just one of many factors contributing to this author's sense of intellectual inadequacy …
[14] Also informer; origin.
[15] Also informee, recipient.
[16] For clarity, this theory pertains to the measurement and communication of units of information: by Shannon's definition, a message consists of a sequence of symbols - Shannon information (Floridi 2010 pp 38). "If base 2 is used, the resulting units may be called binary digits, or more briefly bits, a word suggested by J. W. Tukey" (Shannon 1948 pp 380). In the Conclusion and Appendix 7 the limits of this theory of digital communication will be explored, contingent upon on a definition of "message" broader than Shannon's. This is not to discredit the brilliance and profound impact Shannon's work has had on the world by applying a rigorous conceptual framework to the mathematical treatment of binary information.
[17] A carrier wave of energy propagated through a medium.
[18] Either cable or broadcast.

Digital information is encoded into a digital signal by changing some quality of the carrier wave - modulating the amplitude, frequency or phase of fixed-duration pulses of electrical current (Nair 2002 pp 289) or increments of light (Dume 2012, Economist 2017). The information receiver, able to detect the regular modulations of the signal, can decode the binary sequence from the signal upon reception.

Figure 2: Binary information encoded on a signal

By Mcanet - Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=6024833

Shannon's work formalized the necessary framework for the encoding and transmission of binary information on a digital signal; computers executing instructions encoded in binary were developing in sophistication and usage. From these advancements, networked computing arose.

2.2.1 Cryptography

In the context of digital computing, cryptography applies binary operations and arithmetic to modify these binary sequences - messages intended to carry meaning to an informee - in predictable and sometimes reversible ways. "A message is a discrete unit of communication intended by the source for consumption by some recipient or group of recipients" (Wikipedia 2019b).

To maintain privacy[19] on an open network, message senders need to be assured that:

- only intended recipients can access the contained information - they are "authorized";
- intermediary relayers cannot access the contained information - the message is "confidential"; and
- the message received is identical to the message sent - it has "integrity".

To provide these three assurances - confidentiality, authorization, and integrity - cryptographic algorithms and protocols have been developed, enabling the establishment of secure communication connections on open[20] channels. A basic familiarity with their purpose and function is important to understanding subsequent concepts.

Table 1: Cryptography Key Terms

cipher	"a secret or disguised way of writing; a code" (Oxford 2019a)
plaintext	an unencrypted message, in which the information contained is visible[21]
ciphertext	an encrypted plaintext, in which the information contained is obscured
algorithm	"a process or set of rules to be followed in calculations or other problem-solving operations" (Oxford 2019b)
key	"a sequence of bits" (Techopedia 2019), sometimes with certain mathematical properties
protocol	"a set of agreed-upon conventions" (Cerf 1974)

[19] "Privacy is a human right. It is also the fountainhead for all of the other human rights. If you don't have privacy, you don't have freedom of expression, you don't have freedom of association, you don't have freedom of assembly ... without financial privacy you don't have political rights." (Antonopoulos 2019 57:34). Significant further discussion of the ethics of privacy and the extent of individuals' rights to it is left for future work.
[20] i.e. public.
[21] To anyone familiar with the encoding scheme.

2.2.1a Integrity

Hashing Algorithms

Cryptographic hashing algorithms accept a message of arbitrary length. Based on this unique bit sequence, the algorithm calculates a cryptographic hash (or digest), which is a fixed-length sequence computed from the input value. Most often this hash is substantially shorter than the input data[22] - for example, the SHA-256[23] algorithm returns a 256-bit (32-byte) hash for an input of effectively[24] any length (Veness 2019).

To be useful, cryptographic hashing algorithms must have a few crucial features. They must be deterministic: the same input sequence will always yield the same output hash. Extending this, changing even a single bit in the input message must result in a different output hash[25]. Also, they must be one-way, or trapdoor, functions: it must be extremely difficult to calculate the input data from the message digest[26].

Cryptographic hashes serve as a kind of message "fingerprint": they are effectively[27] unique to the message. This means that, if a message is transmitted along with its hash, the recipient can confirm that the message's integrity is uncompromised by hashing the message on their computer and comparing the output hash with the one included in the message[28]. Additionally, producing and sharing a hash can prove an entity has a specific datum[29] without sharing the datum[30] (Ker 2014).

[22] i.e. it is a "compression function".
[23] "Secure Hashing Algorithm", returning a 256-bit hash. Other common hashing algorithms or algorithm families include MD5, SHA1 and Keccak.
[24] Up to 2,091,752 terabytes (Stackoverflow 2016)
[25] And, ideally (for security purposes), a dramatically different hash - "diffusion".
[26] This is termed "pre-image resistance".
[27] Though not exactly: "note that the compression property implies that the function must be many-to-one, because the domain is infinite and the codomain finite, so infinitely many collisions exist" (Ker 2014 pp 48). In collision resistant hashing algorithms "it should be computationally infeasible to find any" collisions, where different inputs yield the same output.
[28] Often hashes are used as checksums to ensure that a file has downloaded in full and uncompromised (Wikipedia 2019d).
[29] Such as a file.
[30] As long as the recipient knows the hash or has the datum themselves, and can compute the hash.

Figure 3: SHA-256 hashing algorithm Invocation[31]

```
>>> sha256("ucl")
"43c34ee9af7bfaccca6b3bd5d2af0d96bab09732aa5a3dc63a5eaa7015f2a8ce"[32]
```

2.2.1b Confidentiality and Authentication

Symmetric Key Ciphers

Symmetric key ciphers allow communicators to maintain confidentiality even if a message is viewed by an unauthorized entity. The message plaintext is encrypted with a secret key by the informer in a private, secure computing environment prior to transmission on the open network. In this process, an arbitrary-length binary plaintext and a secret key are passed into an algorithm, which returns the encrypted message - the ciphertext. To decrypt, the ciphertext and the key are passed into an algorithm, which returns the plaintext, thereby unobscuring[33] the information represented[34] (Ker 2014).

Asymmetric Key Cryptosystems

While symmetric key ciphers provide communicators confidence in message confidentiality, they are constrained by the need for both parties to have the same secret key. It is infeasible for every pair of communicators establishing a secure connection on a public channel to meet and exchange keys. Transmitting a plaintext key on an open network leaves risk of unauthorized intermediaries detecting it;

---

[31] If the `sha256()` function has declared with the SHA-256 algorithm.
[32] Any online SHA256 hash calculator computing the three letters "ucl", all lowercase will yield the output shown, if the encoding is the same (utf-8).
[33] This phenomenon provides one of the best opportunities to understand the relationship between matter and meaning, in this author's view. This will be cursorily explored in the Conclusion.
[34] While explanation of the mathematical processes specific to various key cipher algorithms is beyond both the scope of this paper and the intellectual capacity of this author, it should be noted that these algorithms rely heavily on the application of the XOR ("exclusive or") binary operator, which returns 1 for differing input operands ($1 \oplus 0 = 1$ and $0 \oplus 1 = 1$) and 0 if input operands are the same ($0 \oplus 0 = 0$ and $1 \oplus 1 = 0$). Because of this operation's properties - it is "commutative, associative and self-inverse" (phant0m 2013) - the application of the operation to the ciphertext with the same key outputs the plaintext. To improve security, more sophisticated symmetric key ciphers have been developed, such as the Advanced Encryption Standard (Daeman 1999, National Institute of Standards and Technology 2001).

attackers could, if aware of the algorithm used[35], decrypt every intercepted ciphertext encrypted with that key.

To resolve this, asymmetric key ciphers and key exchange protocols were developed, based on the work of Ellis, Cocks, Diffie and Hellman, Rivest, Shamir and Adleman, and Merkle (Ker 2014). The algorithms and protocols developed and defined enabled two parties "to communicate confidentially after transmitting a key which is not confidential" (Ker 2014 pp 71).

The distinguishing feature of asymmetric key ciphers is the creation of two keys: a public and a private key. These numbers are calculated using a key generation algorithm, and are linked based on their mathematical relationship.[36]

Two pairs of algorithms characterize asymmetric key cryptosystems, based on these public-private keypairs: encrypt / decrypt and sign / verify.

Encrypt / Decrypt

Asymmetric key ciphers include algorithms by which messages can be encrypted and decrypted, as in symmetric key ciphers. However, a critical difference exists: encryption is performed by passing the plaintext and the message recipient's[37] public key into the encryption algorithm. This is a trapdoor function, one "easy to perform and difficult to invert" (Ker 2014). The resultant ciphertext can be decrypted by passing it, along with the recipient's private key, into the decryption algorithm, which returns the plaintext.

By disaggregating these two functions, the creators of public key cryptosystems created a way to establish a secure connection on a public channel: communicators could simply transmit their public keys to each other in plaintext form, while

[35] which is always assumed, according to Kerckhoffs's Principle, concisely stated by Shannon as "the enemy knows the system" (Kerckhoffs 1883, Shannon 1949)

[36] For example, key generation in the RSA cryptosystem, directly from Ker (2014 pp 74):

(1) Choose two prime numbers each b/2 bits in size, call them p and q. Compute their b-bit product n = pq and the so-called totient $\varphi = (p - 1)(q - 1)$.
(2) Choose an integer e which is coprime to φ.
(3) Publish the public key, which is the pair pk B = ⟨n, e⟩.
(4) Find an integer d such that $ed \equiv 1 \pmod{\varphi}$.
(5) The private key is the pair skB = ⟨n,d⟩2.

[37] NOT message sender's.

maintaining the secrecy of their private key. The message sender could then encrypt the message with the recipient's public key, confident that it could only be decrypted with the corresponding private key - which only the intended informee had[38].

Sign - Verify

Due to the mathematical properties of these public-private keypairs, a second pair of algorithms was possible in the cryptosystem: sign and verify. This pair of algorithms provides message recipients confidence that a signed message was sent by the holder of the private key. If private keys are properly managed[39], this allows informational entities to prove their identity (the basis of authentication) by proving they possess the private key, without ever sharing that key.

The signing algorithm accepts a signer's private key and a message to be signed; the function returns a signed message, or digital signature.

```
>>> signed_data = sign( data, private_key )
```

The verify algorithm, when provided the signed message and the signer's public key, enables the message recipient to confirm that the message was signed with the signer's private key (Ker 2014).

```
>>> verify( signed_data, public_key )
returns True if signed_data was signed with private_key,
        False if not[40]
```

This sign-verify functionality of public key cryptosystems is the enabling technology behind the decentralized web and, in this author's view, represents a profound opportunity to improve the systemic justice of the Internet.

[38] This is, of course, a significant simplification of the complexity of these cryptosystems, but hopefully captures the relevant features of these algorithms for the purposes of this paper.
[39] A big "if", and a primary, and valid, critique of decentralized web technologies. If a private key is lost, in these systems it is effectively impossible to retrieve the assets encrypted or controlled by it.
[40] This is an abstraction.

2.2.1c Cryptographic Protocols

By combining the functionalities of cryptographic hashing algorithms and symmetric and asymmetric key cryptosystems, standardized procedures (i.e. protocols) have been established enabling communicators to verify sender authenticity, ensure message confidentiality[41], and confirm message integrity[42] using technical mechanisms (Ker 2014).

2.2.1d Homomorphic encryption

Although a thorough treatment is well beyond this paper's scope and author's comprehension, "homomorphic encryption is a form of encryption that allows computation on ciphertexts, generating an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on the plaintext" (Wikipedia 2019c, Gentry 2010). Crucially, the plaintext information encoded in the encrypted message is never revealed to the computer processing it, meaning these computations can be executed on untrusted devices while preserving the privacy of its subject.

2.2.1e Proxy re-encryption

In these schemes a message sender can direct a third party proxy node to re-encrypt a ciphertext that has been encrypted for one recipient to be decrypted by another without revealing plaintext to the node. This is significant especially in the context of resource-constrained edge devices; re-encryption for multiple recipients can be outsourced to better-resourced computing environments[43] (Nuñez 2018).

[41] via Message Authentication Codes, MACs.
[42] via digital signatures and keyed hashes.
[43] like a cloud server.

2.2.2 Networked computers

As general purpose analytical machines, Turing-complete computers processing binary information began to be adopted by the military, industry and academia (CrashCourse 2017b) to efficiently and accurately compute and catalogue (relatively) large volumes of data. System developers began to see that the transfer of binary information from one computer to another would enable informational resources to be shared amongst them, thereby enabling the execution of more complex applications, as well as the ability to access data stored on other computers[44].

Innovation in techniques for connecting distant computers and transmitting data culminated in the publication of *A Protocol for Packet Network Intercommunication* (Cerf and Kahn 1974). The paper established the Transport Communication Protocol, which "provides reliable, ordered, and error-checked delivery of a stream of octets (bytes) between applications running on hosts communicating via an [Internet Protocol] network" (Wikipedia 2019d).

These protocols established a standard format for data packets[45] to be transmitted to a destination address, included in the message. Critically, in his design of the protocols Kahn adhered to the principle that "There would be no global control at the operations level" (Leiner 1997). Rather than relying on a centralized authority to organize traffic - which Kahn likely understood to be unscalable - the authority would reside in the rules of the protocol, to which system developers would adhere - an early instance of decentralization in computing technology.

2.2.2a The Web (1.0)

With the establishment of the Transport Control Protocol and the Internet protocol suite, local area networks using heterogeneous operating systems had a standardized way to structure data to transmit to external networks. While this represented an enormous leap forward in information accessibility, it was only with the creation of the World Wide Web and web browsers - alongside the increase in

[44] For a more thorough history of internetworking, see Appendix 5.3
[45] binary messages

public access to computers - that these advancements were made more broadly accessible.

The Web 1.0 was characterized by static web pages connected to resources stored elsewhere on the network by hyperlinks: “a single user-interface to large classes of information” “to link and access information of various kinds as a web of nodes in which the user can browse at will” (Berners-Lee 1990).

2.2.2b The Web 2.0

Concurrent improvements in web communications technology, advancements in usability and design, and increases in web-enabled computer access resulted in the emergence of the “Web 2.0” (DiNucci 1999), characterized by interactivity and a social aspect. Services came to provide more utility, entertainment value and customizability for users (Choudhury 2014). Due to the high capital and labor costs required to provide reliable data storage and access services used by the public, economies of scale, and the network effects such platforms created, and competitor acquisitions, large corporations came to dominate different swathes of the web.

In order to sustain operations and deliver a return to shareholders, their assets needed to be monetized. Charging users was rarely viable - users seem to feel that information and information services should be free. Instead, web-connected devices became channels to serve advertising to users. And, due to the information users were sharing with such platforms, advertising could be targeted to a degree unimaginable prior to the digital era. A positive feedback mechanism has emerged by which more data yields better services[46], which yields more data (Ibarra 2017).

In 2019, calls to break up “Big Tech” abound (Waters 2019); public criticism of the ethical and security practices of these technology providers is rising (Williams 2018, Lee 2018).

[46] Due to “the role [they] play in enabling [machine learning]” (Ibarra 2017 pp 4).

The Internet of Things

Increasing connectivity and decreasing equipment costs has led to the inclusion of computing and data transfer capabilities on a widening array of devices. Embedding sensors and connected computing nodes on vehicles, infrastructure, appliances, surveillance equipment (on Earth and in orbit) and so on enables device controllers to access data stored on and captured by the device and, possibly, the ability to actuate device operation (Xu 2014). Devices are increasingly aware of their surroundings and location in space, and able to interpret their environment and adapt (Zanella 2014). Battery-operated, wirelessly connected computing devices capable of transmitting data representing measurements taken from sensors or other on-board data are being deployed on huge scales; it is projected that by 2025, "there will be 41.6 billion connected IoT devices … generating 79.4 zettabytes (ZB) of data" (Shirer 2019).

2.2.2c The decentralized web

As the records underpinning society's operation have been digitized, the responsibility for maintaining and updating those databases remained where it had historically resided: with governments, banks, medical institutions and other centralized authorities. Providing this service incurred substantial cost in terms of capital equipment and operating and labor costs[47], and these authorities continued to be trusted to securely and accurately maintain the records, and to act in the interests of the users, the subjects of the data in their custody.

This presented an agency problem[48] for many of these trusted data custodians: what was in their (often financial) interest sometimes diverged from the interests of their users. Additionally, their centralized data repositories represented a high value target for attackers seeking to gain access to large quantities of valuable data. "Corporations can argue that data are trickier to manage than oil … The hacker only has to be right once to penetrate a system. Defenders have to parry every jab, all

[47] i.e. Servers and computing hardware; energy, water, real estate; legal and software development talent.
[48] "A conflict of interest inherent in any relationship where one party is expected to act in another's best interests" (Chen 2019).

the time; one misstep and they lose" (The Economist 2019). This value of these data lakes rose as Web 2.0 companies began to aggregate more and more data about their users.

A peer-to-peer electronic cash system

In October 2008 - shortly after Federal Reserve Chairman Alan Greenspan's perspicacious observation in "shocked disbelief" of a "once in a century credit tsunami" (Quinn 2008) - a pseudonymous entity called Satoshi Nakamoto published *Bitcoin: A Peer-to-Peer Electronic Cash System* (Nakamoto 2008). The paper described key features of a peer-to-peer network of computers maintaining a ledger representing electronic cash, configured to prevent double-spending. The system would not rely on trust between participants.

On 8 January 2009 Nakamoto released a C++ program implementing the functionality described in the white paper, inviting any member of the public[49] to run a node by executing the program, thereby participating in the consensus network (Nakamoto 2009). By creating this ledger (blockchain) and the system required to maintain it, Nakamoto presented created an alternative money system fulfilling the three functions of money[50] not reliant on a trusted authority.

While a description of the complexities of the operation of the Bitcoin consensus network are beyond the scope of this paper[51], broadly, it is composed of a network of computers competing to win the right to validate blocks of transactions indicating the transfer of funds from holders to recipients. The winner updates the database state and the competition starts again, based on the new state.

From the money spender's perspective, their funds are held at "wallet address" that they created[52]. This address is derived from a public key, created with a matching

[49] Really, the readers of metzdowd.com, an obscure cryptography message board.
[50] 1. A store of value. 2. A medium of exchange. 3. A unit of account. (Antonopoulos 2017b)
[51] "Bitcoin can be thought of as a state transition system, where there is a 'state' consisting of the ownership status of all existing bitcoins and a 'state transition function' that takes a state and a transaction and outputs a new state" (Buterin 2013)
[52] The bitcoin enters their wallet either by being sent there, or as a result of that wallet owner mining a new block, at which point they add a number of bitcoin to their wallet as a reward per the protocol.

private key by a key generation algorithm executed on their computer. To transmit funds to another wallet address, the user must generate a valid transaction, which includes a digital signature created with their private key. This transaction is then transmitted to the Bitcoin network.

Table 2: Example Bitcoin Private Key, Public Key and Wallet Address[53]

Private key	Hexadecimal	43c34ee9af7bfaccca6b3bd5d2af0d96bab09732aa5a3dc63a5eaa7015f2a8ce
	Binary	1000011110000110100111011101001101011110111101111111010110011001100101001101011001110111010101110100101010111100001101100101101011101010110000100101110011001010100101110100011110111000110001110100101111010101001000111100000001010111110010101010000110011100000001010111110010101010000110011101000000101011111001010101000011001110
Public key	Hexadecimal	04aca6b60b848e3bb6da4fee5b8e8be30a7acef0ed82ef82e63fe3ba5d56525729fddc63723b5a269a5facd9a5316b47da24191757d3c54e9044f29249e65f3fc4
	Binary	1001010110010100110101101100000101110000100100011100011101110110110110100100111111101110010110111000111010001011110001100001010011110101100111011110000111011100000101110111110000010111001100011111111100011101110100101110101010110010100100101011100101001111111011101110001100011011100100011101101011010001001101001101010011110111001010100111010010000010001001111001010010010010010011110011001011111100111111111000100
Wallet address[54]	Base58Check	19ZvdpcrQdS8fPQQjw7UHCvnTGBoAixL4E

[53] See *Mastering Bitcoin* Chapter 4 and Chapter 5 (Antonopoulos 2017) for a thorough accounting of Bitcoin key and wallet generation. These keys were generated using the bitcoin python library; the private key is the SHA256 hash of the string "ucl" in utf-8 encoding, 01110101 01100011 01101100 (Figure 3).

[54] A bitcoin address is "derived from the public key through the use of one-way cryptographic hashing"; it is Base58Check encoded for readability and error protection (Antonopoulos 2017 Ch 4).

Miners gather together batches of transactions into a block[55] and assembles a "block header" which, crucially, contains a hash of the previous block[56], as well as a hash representing the transactions included[57]. The miner then "repeatedly hash[es] the header of the block and a random number [nonce] with the SHA256 cryptographic algorithm" (Antonopoulos 2017). When an output hash less than a certain number[58] is found, the miner broadcasts the solution and block to the network, updating the state[59].

The system hinges on a small feature of the protocol: when a miner finds the solution nonce, entitling it to update the state, with the state update *it puts new bitcoins into a wallet that it controls*. This is how a bitcoin is minted: as a reward for participating in the competition to win the right to update the blockchain to its new state. Miners are incentivized to maintain the computing infrastructure required to sustain such a system; users are in full control of their funds. By combining this insight into human nature with the technical implementation to leverage it, Nakamoto's was "arguably one of the highest-leverage actions in human history" (Ehrsam 2017).

---

[55] Required to be smaller than 1 megabyte, the "block size".

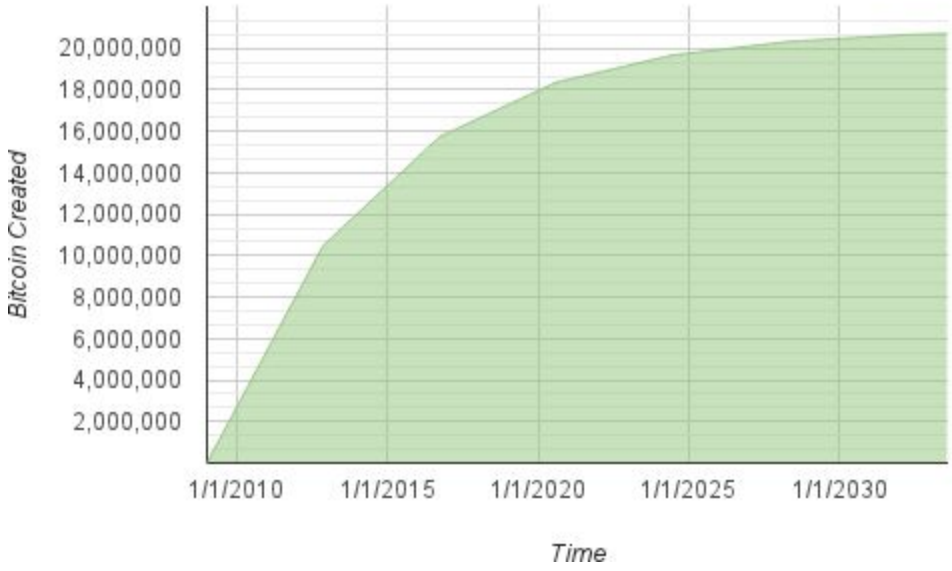
[56] This inclusion of the prior block hash, due to the properties of cryptographic hashing algorithms described, means that once a block is mined changing it would entail re-mining all subsequent blocks, including the sequence of new block hashes. This is the basis of the term "blockchain", which is, in the strictest sense, simply a database in which the current state is cryptographically linked to the prior state, providing a guarantee of immutability. The term has come to be understood, however, as the combination of "the state transition system with a consensus system in order to ensure that everyone agrees on the order of transactions" (Buterin 2013).

[57] The inclusion of this Merkle root hash further strengthens the security of the system. If a dishonest node changes a single bit (say, giving itself more bitcoin), the deviation is detected by the rest of the network, and rejected.

[58] ""Proof-of-work involves scanning for a value that when hashed, such as with SHA-256, the hash begins with a number of zero bits" (Nakamoto 2008); i.e. the output hash has a number of leading 0s - formalized as the "block difficulty". By adjusting this block difficulty the network can raise or reduce the difficulty of finding a nonce yielding an output hash, thereby self-adjusting to the amount of computing power operating on the network. The hash rate of the Bitcoin network is a measure of its security (Long 2019) meaning, sadly, that the security of the network - a strength - is proportionate to its energy consumption - a substantial drawback (Vincent 2019).

[59] This so-called Proof-of-Work consensus mechanism serves to effectively randomize which validator node wins the right to update the database state, and puts in place a substantial computational cost to change prior states. Solution nonces for all subsequent blocks would need to be computed, as changing the prior data would most likely result in the hash of subsequent block headers not fulfilling the block difficulty requirement to be accepted as valid, according to the protocol.

Figure 4: Bitcoin money supply



Every 210,000 blocks the block reward is reduced by 50%[60]; "in approximately[61] 2140, almost 2,099,999,997,690,000 satoshis, or almost 21 million bitcoin, will be issued. Thereafter, blocks will contain no new bitcoin, and miners will be rewarded solely through the transaction fees." (reprinted from Antonopoulos 2017, Ch 10)[62]

A quasi-Turing complete world computer

In 2013 Vitalik Buterin recognized that while the Bitcoin protocol provided some capability to execute scripts upon transaction validation - "a weak version of a concept of 'smart contracts'" - "the scripting language as implemented in Bitcoin [had] several important limitations", including a "lack of Turing-completeness".

Buterin's key insight was that while performing the state update required to validate transactions, arbitrary computations could be performed by the validating computer, and arbitrary data written to the ledger. He proposed Ethereum, "a blockchain with a built-in Turing-complete programming language, allowing anyone to write smart contracts and decentralized applications where they can create their

[60] According to the protocol. Block solutions broadcast to the network that violate the protocol will be rejected by the rest of the network.
[61] The exact date is uncertain due to stochasticity in block mining times.
[62] Bitcoin is inherently deflationary.

own arbitrary rules for ownership, transaction formats and state transition functions" (Buterin 2013).

With this adaptation Buterin created a system that, in this author's view, is a technological innovation representing one of the greatest opportunities to improve the dignity of the human condition in history[63]. The decentralized web - Web 3.0, or Web3 - is arising from the realization "that entrusting our information to arbitrary entities on the internet [is] fraught with danger" (Wood 2014). This movement toward decentralization has been occurring, in some ways, since the earliest days of networked computing, and even before (Wikipedia 2019e)[64].

The Ethereum Protocol

Based on a Proof-of-Work consensus mechanism similar to Bitcoin[65], the Ethereum protocol extended the functionality of the blockchain network to enable users to deploy and store executable computer programs - termed "smart contracts"[66] (Szabo 1994) - on the blockchain.

Transactions specifying the transfer of funds from one account (the sender) to the wallet address[67] of another (the recipient) were of course possible, as in Bitcoin, but compiled program bytecode could also be deployed and made available for execution in the Ethereum Virtual Machine (EVM)[68]. Upon deployment[69], the program's "contract address"[70] returned. Transactions subsequently sent to a

[63] Hyperbolic? Only time will tell ...
[64] For example, the establishment of a (necessarily) open protocol - TCP/IP - by Cerf and Kahn (1974) decentralized the ability to develop software compliant with that protocol. This sort of precompetitive coordination is of deep interest to this author.
[65] Initially - though an alternative consensus mechanism, Proof of Stake, is being developed for Ethereum 2.0 (Kim 2019).
[66] Smart contracts - "computerized transaction protocol[s] that execute the terms of a contract" - were first conceptualized by Nick Szabo (1994), who envisioned sophisticated systems of ownership and custodianship subject to conditions implementable in code. Buterin has since expressed "regret [at] adopting the term 'smart contracts'. I should have called them something more boring and technical, perhaps something like 'persistent scripts'." (@VitalikButerin 2018).
[67] Externally-owned account, sometimes EOA.
[68] A "transactional singleton machine with shared-state" (Wood 2019).
[69] Deploying a contract requires a special "contract creation transaction" to be sent to the zero address (0x0), with compiled bytecode included (Antonopoulos 2018 Ch 7).
[70] A second type of address in the Ethereum system, contract addresses reference the contracts deployed there. They do not have a private key - it "in fact does not exist - we can say that smart contract accounts own themselves" (Antonopoulos 2018 Ch 7)

contract address will result in the EVM attempting "to execute the contract, … [trying] to call the function named in the data payload of your transaction" (Antonopoulos 2018). These smart contracts were capable of computing data, as well as writing data to the blockchain[71], providing the necessary functionality for decentralized applications (dApps).

The costs of being public

As a public, permissionless service, the Ethereum system needed to disincentivize wasteful or malicious behavior. This is achieved by requiring a fee to be paid by any users seeking to transfer funds, or invoke smart contract functions. Each computational or write operation has a corresponding "gas" cost, representing the work the network would need to do to perform the operation (Wood 2019 Appendix G).

As smart contracts run, the gas costs accumulate; once program execution completes, or a limit is reached[72], the total cost of transaction validation[73] is calculated and deducted from the transaction originator's account[74]. This fee is paid in ether, the native[75] currency maintained by the system[76]. Each block has a total limit to the amount of gas that can be consumed, "to keep block propagation and processing time low, thereby allowing for a sufficiently decentralized network" (jnnk 2015, Ethereum Foundation 2019). The gas mechanism protects the Ethereum

[71]A note fundamental to the premise of this dissertation: the system is agnostic to the data being passed into these functions or written to the chain; so long as the data type is compliant, it will be accepted. It makes no judgments about the veracity of the information the data represents (which is a quality). Within the blockchain vocabulary, external entities that provide data to a blockchain are called "oracles"; these entities represent a potential weakness in the system (Fecke 2018).
[72] As specified by the transaction originator, or system block gas limit.
[73] To enable the market to determine the value of performing operations on the Ethereum Virtual Machine, and to act "as a buffer between the (volatile) price of Ethereum [sic] and the reward to miners for the work they do", account holder submitting transactions specifies the price (in a subunit of ether - "wei") they are willing to pay per unit of gas (Antonopoulos 2018 Ch 13). This allows miners, who receive transaction fees of validated blocks, to select which transactions will provide them the highest rewards.
[74] Externally-owned accounts must have some ether in them for any transactions submitted to be validated successfully.
[75] Withing smart contracts decentralized application developers can create their own cryptocurrencies, which have many of the properties of the ones described, namely, that provision of the right digital signature authorizes transfer.
[76] And the currency of the block reward for participating in the validation process.

system by making users “pay proportionately for the computational, bandwidth, and storage resources that they consume”; “thereby disincentivizing attackers” (Antonopoulos 2018 Ch 13).

Three forms of decentralization

In this research three forms of decentralization that characterize the decentralized web have been identified . These systems realize the benefits of openness and mitigate the risks through game theoretic incentivization of system participants. They seek to eradicate any “single point of failure”, instead exhibiting antifragility (Taleb 2013), the quality of being self-healing: they tend to strengthen, on balance, in response to stress, shock and volatility.

1 - Open Source

The first form of decentralization is in the authority to propose and make changes to the code and protocols. This - the open-source model of system and software development - opens access to source code to public scrutiny. In doing so, the pool of potential contributors to a project is expanded, and the likelihood of discovering unintentional errors or intentional abuses included rises[77].

Tools such as git (Torvalds 2005) - a “distributed version control system” (git-scm.com 2019) have radically improved the ability of groups of loosely-coordinated developers to work on the same codebase. In theory, an open-source project can be a pure meritocracy. Ideas - code update proposals - are judged by the community on their quality; if worthy, they are accepted and incorporated. Without a central authority directing efforts, participants are free to pursue solutions they think appropriate[78]. Openness improves security[79], diversity and adaptability.

---

[77] Linus’s Law: ”Given enough eyeballs, all bugs are shallow.” (Raymond 1999)

[78] An example demonstrating the benefit of this diversity is the development of multiple implementations of a blockchain protocol - in the Ethereum ecosystem, geth, parity, pyeth, cpp-ethereum, etc. (Antonopoulos 2018 Ch 3) - which reduces the impact of a common mode failure (Buterin 2017) in any one of those implementations, enabling a network to respond to such a failure without going offline.

[79] Of course, open source code can be examined and exploited more easily by malicious entities. It seems that on balance open source code is more secure, but counter-examples surely exist; some systems are more secure if source code remains private.

2 - Distributed Ledgers

Second is the decentralization of responsibility to maintain and update informational assets: the so-called "distributed ledger". This form of decentralization is enabled by the creation of a strict protocol enabling nodes to confirm block validity by recomputing each transaction: "each node verifies the results of each transaction" (Ryan 2017).

3 - Private Key Custodianship

Third, and most importantly, is the decentralization of the responsibility to hold the private keys needed to interact with the systems. This aspect of the decentralized web is enabled by the asymmetric key ciphers described earlier, namely, that a message recipient (in this case, a blockchain miner verifying the validity of a transaction), given a message and a public key, can mathematically confirm that the message was signed by the corresponding private key. This means that all interaction with these systems can be done by a user without ever transmitting the private key over the Internet. The only way to access the funds in a Bitcoin or Ethereum wallet, or to confirm identity as necessary in the invocation of some Ethereum smart contracts, is to present a valid transaction, including a digital signature. This signature is impossible to generate with the private key (Ker 2014, Antonopoulos 2018, Buterin 2013).

This represents a significant break from the centralized database technologies of the prior web, in which the secrets used to authenticate users (passwords) are stored on a centralized server, and must be transmitted[80] to that server for storage. In a simplified model, when a user tries to log in, the server-side software confirms that the password matches the one that it has in its records[81]; if so, the user is considered authenticated, and granted appropriate rights.

Based on Wood's (2014) observation on trusting arbitrary entities on the Internet, this model carries risks: the authorities managing a user's data might mistakenly delete or alter it[82], reveal it to some malicious attacker[83], or may choose to deny

[80] Albeit on an encrypted channel like HTTPS - sometimes (Vyas 2016)
[81] Hopefully stored securely … though sometimes, even in the case of highly skilled and resourced firms, not (Krebs 2019).
[82] as in the cases of DreamHost, several government agencies …
[83] … JPMorgan Chase, Equifax …

access or revoke service provision[84] (McCoy 2015, Haselton 2017, Hopkins 2017, Fernandez 2019, Galperin 2015, Prince 2019)[85].

Technologies enabled

Decentralized Autonomous Organizations

Smart contracts enable a broad range of decentralized applications to be developed and deployed on a blockchain network. Of particular interest is that of the DAO: "decentralized autonomous organization", an organizational structure enabled by smart contract technologies. As with many of these concepts, the collective understanding of DAOs is nascent and few are operating successfully at the time of writing, but generally these entities operate according to rules and data encoded on a blockchain. When deployed on public blockchains, total transparency as to the governance (Merkle 2016)[86], data, membership and activity of the DAO is available for public review[87].

Due to the versatility of smart contracts these rules can enforce any policy, but DAOs can be conceived of as providing members an interface through which they can interact with the informational and financial assets necessary to coordinate organizational projects[88]. At this early stage DAOs appear to hold enormous promise to improve organizational efficiency, thereby enabling the provision of services to unserved market segments[89]. They also hold the potential to enable as of yet unforeseen organizational structures (Aragon 2019, DAOstack 2019, Rea 2019).

---

[84] … Facebook, and Cloudflare, to name a few.
[85] While a topic of intense fascination, an in-depth discussion of the ethics of the denial of access to informational services is beyond the scope of this dissertation. Web3 purists seem to err toward a complete rejection of the morality of authorities deciding to prevent behavior they deem harmful. This author takes a more nuanced view, but overall would rather see those (dis)incentivization mechanisms (such as censorship, sanctions and so on) built atop a system that is fundamentally oriented toward individual liberties and rights.
[86] i.e. smart contract source code.
[87] Of course, some organizations require privacy for ethical, legal or business reasons. Due to its constraints, this paper cannot thoroughly explore the mechanisms and ethics of organizational privacy, nor mechanisms for preserving privacy on public blockchains.
[88] A common use case would be the collection of funds and collective decision-making about their disbursement, based on some voting mechanism (DAOstack 2019).
[89] For example, IBISA is a project building a decentralized autonomous organization offering crop insurance to smallholder farmers in developing countries. Their customers are unserved by traditional insurers due to the high costs of offering products relative to the low revenue opportunity the customers represent (Bitvalley 2018).

Private Blockchains

Public, permissionless blockchains incentivize members of the public to utilize or contribute the computing resources necessary to sustain the system; to prevent abuse high fees are charged for use of compute and data storage resources. The community develops the protocol and software and decides on which features to implement. Furthermore, all data - including accounts and balances - on a public blockchain can be accessed by anyone with an Internet connection[90].

For use cases where system developers are not willing to sacrifice control and privacy, blockchain protocols and clients designed to be deployed on private, permissioned networks exist, such as Hyperledger (The Linux Foundation 2019), Corda (2019) and Quorum (2019).

If configured properly - say, with validating nodes hosted by competing firms, or loosely coordinated firms situated across a value or supply chain, where each would benefit by the existence of some shared database - private blockchains can provide many of the benefits conferred by decentralization[91] without requiring the participants to rely on public networks. Furthermore, operating costs can be substantially reduced if network participants trust each other to some degree, or have legal mechanisms in place to disincentivize malicious behavior.

Ricardian contracts

Blockchains and the cryptocurrencies and smart contract systems they enable have emerged independent of the direction of any sovereign government, regulatory agency or monetary authority. While this feature attracts some, others identify an opportunity to utilize smart contracts within the existing regulatory and legal framework.

First proposed by Grigg (2004), Ricardian contracts contain both human- and machine-readable elements, a legal (prose) and smart (code) contract. This allows them to document the intent of the agreeing parties, be interpreted in a court of

[90] It should be noted privacy preserving techniques on public blockchains are in development.

[91] Persistence, immutability, security, among others. Web3 purists may disagree with this point.

law, and delegate certain functions to smart contract platforms (Braendgaard 2016, Clack 2016).

Zero-knowledge proofs

Zero-knowledge proofs provide a verifier an assurance that prover knows a piece of information without revealing anything more. It is a privacy-preserving method with applications in authentication systems, identity applications, nuclear disarmament (Glaser 2014), and has been proposed as a privacy-preserving mechanism on public blockchains (Orcutt 2017, Wu 2014, Wikipedia Contributors 2019f).

Content Addressing

In location-addressed systems, resources are referenced by their address, often a URL referencing a location in a web server's directory structure, or containing parameters for a database query enabling the server to respond with the requested data.

This, however, represents a single point of failure[92]: if a file is moved from its location, the URL does not resolve and the web server returns a "404 Not Found" error; this resource is inaccessible. In the serverless Web3 paradigm[93], this problem is solved by addressing data by its content. The technology is built largely on hashing algorithms and their ability to prove the "sameness" of two equal-length binary sequences. Files are identified by a unique and deterministic hash; a user requesting a specific file could receive segments from various peers, re-assembling it on their local machine[94].

Juan Benet (2014) proposed the InterPlanetary File System (IPFS), "a peer-to-peer distributed file system" based on content addressing. The use of content addresses enable short hashes[95] to reference larger files[96]. Due to the cost of writing to a blockchain, and integrity guarantees, IPFS and other content addressing systems

[92] Commonly experienced by this author.
[93] In the spirit of removing single points of failure, and improving performance.
[94] BitTorrent (Cohen 2003) forms some of the basis of Benet's solution.
[95] In IPFS, v0 Content Identifiers are 46 bytes in length (Protocol Labs 2019)
[96] And, relatedly, "*Proof-of-Replication* was developed to solve the specific problems of a verifiable, decentralized storage network that could incentivize and reward file storage", thereby improving on the fragility of the location-addressed client-server model (Protocol Labs 2019b).

and decentralized file storage systems are technologies crucial to the decentralized web.

Self-Sovereign Identities

> "ed: The vulnerability that is being exploited in all systems is identity."
>
> @santisiri, 2019

In the emerging vision of the decentralized web, digital identities are controlled and owned by the individuals and organizations they represent: "users [are] the rulers of their own identity". While the community's understanding of the idea is nascent and developing, Allen defined 10 principles of self-sovereign identity (2016), oriented toward individual autonomy and privacy (Appendix 3). These developments are enabled by the advancements in networked computing and cryptography described - self-sovereign identities rely on users being able to create and manage private keys securely.

Sovereign Sensors

Humans interact with the Internet through computers, which detect inputs and establish connections with other computing nodes. In the most literal sense, every personal computer is a sensor: they "sense" keyboard or touchscreen inputs, as well as incoming transmissions from other computers, either wirelessly via an antenna or through some wired input. As such, for the purposes of this research the definition of "connected sensors" includes any digital computing node with the ability to receive and transmit data.

However, a distinction should be made by computers that are controlled by direct human interaction[97] and ones that are autonomously controlled by software agents installed by human developers[98]. This research was inspired by a peculiar question: what if these autonomous agents had self-sovereign identities of their own? What if

[97] Desktops, laptops, tablets and smartphones, generally.
[98] To this author's knowledge no computer exists that is controlled entirely by software that was developed and deployed by another computer, though it is not unfathomable (Boyd-Rice 2018). It seems obvious, however, that all future software and hardware technology ultimately originated in a human action.

computers had private keys that only they had access to? How might such sensors serve society, and how might they threaten its functioning?

Such a capability has been an area of active research for some years, primarily led by the Trusted Computing Group, including "AMD, Hewlett-Packard, IBM, Intel and Microsoft" (Merritt 2003). Detailed discussion of the technical aspects of trusted computing is well beyond the scope of this paper[99], but the fundamental premise is of a chip that contains in non-volatile memory a "public and private key pair, … created randomly on the chip at manufacture time [that] cannot be changed" (Safford 2003) and cannot be accessed by any external entity[100]. With this technology, it seems that computers have a right to privacy[101].

With this technology, ethical concerns abound (Anderson 2003), but if configured properly and governed transparently, such a system could offer myriad ways to improve the security in the emerging Internet[102], and the ability for computer networks to monitor environmental, physical and information security conditions.

The notion of computers in sole and secure possession of private keys attains a new level of significance in the context of the decentralized web[103]. As sole controllers of wallet addresses, edge devices can be used as oracles in smart contracts without any point of human interference, and a computer can hold funds accessible only to

[99] As well as the cognitive capacity of this author.
[100] If the device securely maintains custody of the private key (i.e. it has the only instance of that private key in existence), all digital signatures with that key must have been performed on the device.
[101] Should they?
[102] Security and privacy are often viewed as a trade-off (Malik 2018): increased surveillance erodes privacy but increases the likelihood of authorities detecting threats to public safety. Acknowledging the constraints entailed by the computational intensity of homomorphic cryptography, this author is still wondering: could homomorphic algorithms analyze data about some entity to detect a threat, and only reveal the contents of the data to authorities if a threat is detected? Could software agents analyzing data about human activity be configured as a sort of benevolent panopticon - an artificial intelligence oriented toward minimizing harm and damage to life and property?
[103] In the Ethereum Yellow Paper, Wood recognized the potential of machines providing inputs to smart contracts: "A transaction (formally, T) is a single cryptographically-signed instruction constructed by an actor externally to the scope of Ethereum. While it is assumed that the ultimate external actor will be human in nature, software tools will be used in its construction and dissemination. [Footnote 1] Notably, such 'tools' could ultimately become so causally removed from their human-based initiation—or humans may become so causally-neutral—that there could be a point at which they rightly be considered autonomous agents" (Wood 2019 pp 4).

it. A vehicle could detect necessary on-board maintenance - and release funds to perform it. A weather buoy could submit data to a smart contract - data that could only have originated on the buoy. Access to the imagery captured by a satellite with a synthetic aperture radar or a surveillance camera installed in an urban environment or a balloon (Harris 2019)[104] could be managed by a DAO governed by an open and transparent set of rules, subject to identity-based, spatial and temporal conditions.

2.3 On resource governance

"Governance refers to all processes of social organization and social coordination" (Bevir 2012). In this context - of computers - governance is closely associated with the concept of access control, the processes by which access to and use of the informational resources on a computer system is managed (Rouse 2019). It also refers to decision-making about the direction of blockchain software and protocol development (Ehrsam 2017, Zamfir 2017)

A thorough exploration of the theory and practice of governance is beyond the scope of this work. Broadly, however, if governance is conceived of as "processes of rule" (Bevir 2012), the principles underlying governance design, as well as the structures manifesting such processes[105] hold great bearing on the justice of the system: its efficacy in equitably governing the subject community.

[104] "'We do not think that American cities should be subject to wide-area surveillance in which every vehicle could be tracked wherever they go,' said Jay Stanley, a senior policy analyst at the American Civil Liberties Union." Leaving aside the apparent inevitability of this, this author believes that such a capability is not inherent unethical. If such information is stored, analyzed and accessed by an open source system designed to respect the individual's privacy, with appropriate anonymization, it is believed that such a system has the potential to fundamentally improve physical security in the service of the public interest. The key phrase: "designed to respect the individual's privacy"; likely incorporating zero-knowledge and range proofs, as well as homomorphic cryptographic techniques, etc. If "the surveillance state is inevitable" (Weigert 2015), then the nature of the state must be adapted so human rights are respected.

[105] Government, as in "the system or group of people governing an organized community" (Wikipedia 2019g). Note that this conception of government is much broader than the common state-based understanding of the term; it is "a means by which organizational policies are enforced, as well as a mechanism for determining policy" (Wikipedia 2019g).

The governance of the physical commons is fairly well understood, most clearly conveyed in the eight "design principles" of common resource governance identified by Ostrom (1990). However, these principles seem to primarily apply to the governance of physical resources; no clear consensus on the nature and intricacies of informational resource governance appears to exist. On one hand, information is "a good - … an object of economic transactions", "typically non-rival and sometimes nonexcludable"[106] (Varian 1998). On the other, "numerous English authorities have affirmed that information or data is not property" (Bilbow 2019)[107].

It seems that informational objects exist in a space with different laws to physical space. Given the commitment to investigating the factors pertaining to the governance of informational resources, the differences between these two spaces in which reality manifests is of acute importance to the research agenda. This will be explored further In the Conclusion.

---

[106] This author is critical of the general economic definition of "non-rivalrous", which forms the basis of the distinction between public goods and common goods (Vu 2015). The difference is appreciated: the use of a rivalrous good decreases the value of the asset, thereby making that value unavailable to future users (i.e. when a tree is cut down and lumber removed no future person can use that specific tree for lumber), whereas the utility of non-rivalrous goods is not diminished with use (i.e. a waterway can be used to transport a shipment without reducing the waterway's utility for future users). The point missed within the definitions reviewed is the temporal rivalry of many public goods: *while I am using this waterway*, availability to others is reduced. Informational resources, especially in the decentralized, content-addressed web, may (almost) be the first non-excludable resources in history (as even high-bandwidth servers have a limit to the number of users who can access an informational resource in a given time). True non-excludability may be impossible - except perhaps in the case of an informational asset in the custody of its user.

[107] "can data be owned?" (@santisiri 2019b)

3 Methodology

3.1 Overview

3.1.1 Purpose

An agent-based approach was employed to investigate the dynamics of sensor networks recording data on a public blockchain. As the scale, resolution and connectivity of the Internet of Things grows, so does its potential to improve situational awareness (Shirer 2019). If leveraged properly, enormous gains in the efficiency and prevention of harmful or malicious activity might be realized in the networks monitored by providing system operators a more accurate understanding of systemic patterns and dynamics - with substantial social, environmental, military and commercial implications.

As outlined, public blockchains have unique limitations when compared with traditional database management systems, cloud servers and private distributed ledger implementations. This modeling effort was conducted to explore emergent dynamics of connected sensor networks reporting measurements from the edge to a public blockchain. By understanding these dynamics and the trade-offs inherent to using public and permissionless blockchain architectures, these findings might contribute to helping system designers to take a more informed approach to architecting such systems, thereby balancing the costs and benefits of decentralization.

This model is focused on investigating factors related to network costs and transaction processing times. These scaling questions are important in the context of connected sensor networks, which are often characterized by high data capture and transmission volumes, and provide the most value to system operators if the data captured at the edge is made available for analysis rapidly and is available long term (Gutierrez 2015).

3.1.2 State variables and scales

Using Python 3's (van Rossum 1995) Mesa framework (Core Mesa Team 2019), two subclasses inheriting Mesa's `Agent` class were defined: `Sensor` and `Blockchain`. A `SensorBlockchainNetwork` subclass of Mesa's `Model` class served to instantiate an individual simulation, configure model schedulers and collect data from model reporters. Parameter sweeps were performed using Mesa's `BatchRunner` tool in an iPython notebook (Pérez 2007); collected data was analyzed and visualized, also with iPython, using the `numpy` (Oliphant 2007), `pandas` (McKinney 2010), `statsmodels` (Seabold 2010) and `matplotlib` (Hunter 2007) packages.

3.1.2a Sensors

Sensor agents represented edge devices capable of recording empirical data about their surrounding environment[108]. Logic designed to simulate on-device data computation and transmission operations was included within the class definition as methods.

[108] The type of recording is abstracted; image sound, image, temperature, air quality, rotation, acceleration, etc.

Table 3: Sensor agent elementary properties

Name	Value	Initialized Parameter
`battery_life`	Energy level	`10000`
`mortal`	Boolean, whether sensor should be removed from scheduler upon battery depletion	`False`
`dead`	NaN, replaced with block number on battery depletion (if mortal)	`NaN`
`record_cost`	Energy consumed with each recording	`1`
`gas_price`	Price (gwei[109]) per unit of gas, required in transaction (Ryan 2017)	`20`
`stochasticity`	Arbitrarily introduced variation to mimic reality	`0.05`
`gwei_spent`	Gwei spent that tick	`0`
`last_sync`	Block number of previous transaction submission	`0`
`nonce`	Iterator for valid transaction generation	`0`
`db`	Block-indexed array with bytes recorded each tick	Empty array

3.1.2b Blockchains

`Blockchain` agents were designed to simulate key features of the quasi-Turing complete public blockchain networks described, with the Ethereum platform serving as the primary inspiration. The `Blockchain` class defined for the model here is based on key features of the Ethereum platform[110] (Antonopoulos 2018).

[109] Gwei stands for gigawei; 1 wei = 10^{-18} ether, 1 gwei = 10^{-9} ether = 1 nanoether (gwei.io 2019). Think pence to pounds or cents to dollars. It must be noted that the term "gwei" is simply borrowed from the Ethereum lexicon; here it represents a theoretical cost, and should not be thought of in terms of value in fiat currency. Rather than attempt to estimate true costs, this dimension was modeled in an effort to understand trends in total transaction costs per sensor across the parameters swept.

[110] Of note: the `Blockchain` class does not have a `step()` method, nor was it added to the Mesa scheduler. This was intentional, to emulate the inherent reactivity of smart contracts executed on blockchain networks. All smart contract computations can be traced back to an origin externally-generated transaction sent to a contract address. Tasks cannot be scheduled on Ethereum (Antonopolous 2018 Ch 13).

In the `Blockchain` class (which inherits `Agent)`, attributes including the gas limit per block in gas and gas cost per byte written in gwei, were defined upon initiation. These integer values were fixed both through a single model run[111] and across model runs.

Class methods defined allowed for the blockchain network to perform necessary actions, including adding submitted transactions to its mempool (`Blockchain.add_to_mempool(tx)`) and mining a block (`Blockchain.mine_block()`).

The latter method included logic to select the highest value and most recent transactions from the mempool[112], validate them - including invocation of that appropriate sensor's `Sensor.confirm_tx(tx)` method - and update `Blockchain.chain` dataframe's state.[113]

[111] Note that this simplifies this dynamic when compared with Ethereum, in which block gas limits are adjustable (Antonopoulos 2018 Ch 13). While the dynamic nature of these parameters are critical to the adaptability and scalability of blockchain systems, these models focused on variation in the size and behavior of the networks connecting to the blockchain in each model run.

[112] Thereby simulating an economically rational miner's selection of transactions from the mempool, designed to maximize block rewards.

[113] The `Sensor.confirm_tx(tx)` method was designed to emulate the response sent from the blockchain to the client upon transaction confirmation. This aspect of smart contract behavior is the reason for the asynchronous functionality of the web3.js library, in which JavaScript's async await syntax is employed to pause execution of client-side code reliant on transaction confirmation and, possibly, values returned from smart contract computations (web3.js 2019). If edge devices were running NodeJS programs (Node.js Foundation 2018) to execute on-device computations and transmit data, the web3.js library would likely be used to interact with the Ethereum blockchain. This would be done by establishing a connection to the network (local, test or main net) and invoking methods of the contract instance, referenced by the contract address. This would be done in a local, private environment as one would need to sign the transaction with the private key of the wallet containing the ether necessary to pay for the gas costs.

Table 4: Blockchain agent elementary properties

Name	Value	Parameter / Range
`block_gas_limit`	Block limit of computation or data write volumes in gas	`8000000`[114]
`gas_per_byte`	The gas cost to write one byte of data to the chain	`625`[115]
`chain`	A dataframe of boolean values representing whether sensor recordings have been validated on-chain	Empty dataframe
`tx_ct`	An iterator so each transaction submitted gets a unique id	`0`
`mempool`	A dataframe containing transactions, including 'mined' boolean column	Empty dataframe

In the middle-range model developed (Gilbert 2008 pp 42) blockchain complexity is limited to write operations: no on-chain compute operations were simulated. This notable exclusion of a key feature of smart contract platforms was deemed necessary to reduce model complexity[116].

3.1.2c Higher-level entities

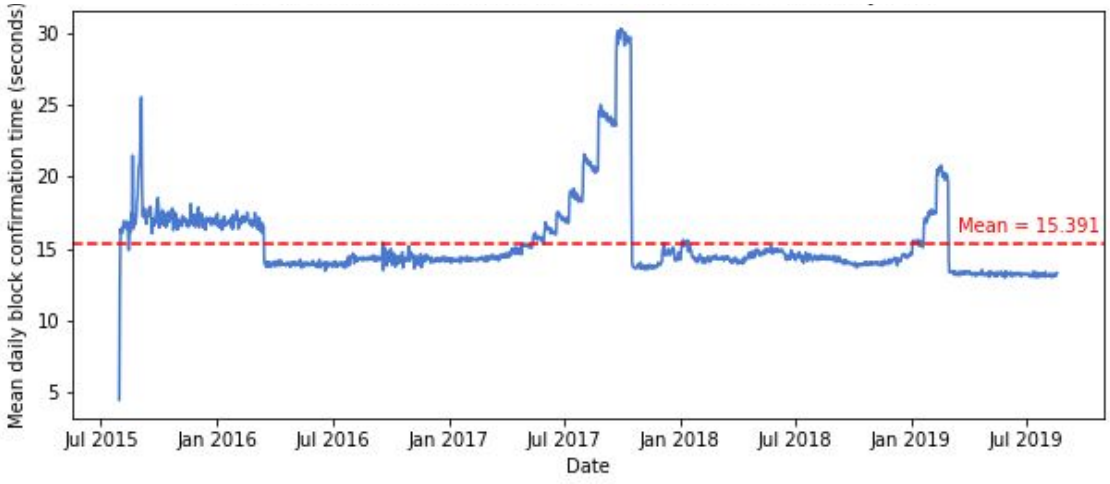
No higher-level groupings of agents exist, though much opportunity to model heterogeneous networks exists within the modeling framework developed.

3.1.2d Scales

The physical location of model agents is not specified. A single tick in the model represents a block confirmation. For context, Figure 5 depicts block times on the Ethereum mainnet since network launch: 15.391 seconds, on average.

[114] (Moriya 2018).
[115] (Ryan 2017).
[116] Methods were included in the `` `Blockchain` `` subclass to simulate on-chain compute operations, namely, gas costs.

Figure 5: Ethereum mainnet block confirmation times



(Etherscan.io 2019)

Model runs were executed for 300 ticks[117]. Scales were selected based on a series of initial model runs. Visual inspection of the results of these model runs confirmed the intuition that network block confirmation times and informational currency measures would stabilize around the optimal values of 1 block and 1.0 respectively. Increasing network loads through adjustment of record volume, record frequency and network size parameters resulted in these measurements degrading rapidly at a tipping point, representing the network reaching some operational performance threshold. Fixed parameters, and the ranges of swept parameters, were chosen to enable investigation into network dynamics around this threshold.

However, development of a facsimile model exactly simulating each of these aspects was well beyond the scope possible given time, computational and analytical constraints. While such a model would be difficult to build, it could provide researchers the opportunity to simulate unanticipated stresses on blockchain networks, as well as the emergent complexities of on-chain interactions such as DAO behavior, voting and governance mechanism performance high-demand periods such as token listings (Peaster 2018), dApp launches (BBC 2017), asset ownership transfers (Gupta 2019), and so on.

[117] This substantially but necessarily abstracts a blockchain network's complexity: the value of such a platform is largely found in its persistent operation.

3.1.3 Process overview and scheduling

To model the operation of a network of sensors connected to a blockchain, key steps in the process were defined for simulation, while others were abandoned in abstraction. Judging the relevance of features of a complex system is a primary point where bias can be injected into such an enquiry; good faith efforts were made to retain important aspects.

Actions can be conceived of as taking place in two computing environments: on board the edge sensor and on the computing nodes forming the blockchain's consensus network, where transactions are validated and the state is updated, including execution of any smart contract code. The data follows a serial process (Table 5).

Table 5: The path of data in the network simulation

Step	Action
1	The edge device takes an empirical recording from the edge and stores it in on board memory, such as an SD card.
2	When ready to transmit data to the blockchain, the edge device gathers all un-submitted data[118], includes it in a transaction, signs the transaction and transmits it to the blockchain network.
3	Upon receipt of the transaction data by the miners, the transaction is added to the mempool, a local data store where transactions await validation.
4	Each block (tick), the network mines the next block by selecting a set of transactions requiring less gas to execute than the block gas limit. Upon validation, the blockchain updates its state to reflect the inclusion of the data (edge recordings) contained in those transactions.
5	A message is transmitted back to each edge sensor confirming transaction validation. This enables sensors to keep a record of the financial expenditures required to perform write operations[119].

To simulate the lack of coordination in reality amongst edge nodes, `Sensor` agents were activated in random order. The tick was completed upon block confirmation, at which point the next transactions in the mempool move up the queue; for this reason the final action taken in each model step was the invocation of the `Blockchain.mine_block()` method.

Each tick represented one block validation by the blockchain network, during which the blockchain's state was updated to incorporate data included in each transaction mined.

[118] i.e. data recorded since the prior transaction submission.
[119] Note: This is not an accurate representation of reality: in Ethereum the accounts and balances are stored on the blockchain itself, and when an externally-owned account's transaction is validated the transaction costs are deducted from that account's balance. No communication is required with the submitting edge node to complete transaction validation. However, this structure was used for conceptual simplicity, meaning the Sensor objects kept track of expenditures - relevant for Mesa's reporting procedures. This meant that costs were measured by adding costs incurred to the `Sensor.gwei_spent` value, rather than seeding sensor-controlled accounts with currency and subtracting from it each transaction validation, as would happen if such a system were actually developed for the Ethereum network.

3.2 Design concepts

3.2.1 Sensing

As informational entities, the agents as defined and simulated sense their environment in a few ways. Edge nodes convert the qualities of the analogue data contacting their sensory interface into some quantitative representation of that quality, represented as binary information[120]. The moment of this analogue-to-digital (qualitative-to-quantitative) conversion is explored more thoroughly in Appendix 7, as understanding it is critical, in this author's view, to understanding the nature of information and thus the governance of informational resources. However, this is largely abstracted in this model, with only the "volume" of data captured in bytes being explicitly referenced in model execution.

All other agent "sensing" pertains to the detection of signals encoded with binary information, carried on some channel. Specifically, these include the detection of a transaction submitted to the blockchain network[121], and the subsequent detection of the message from `Blockchain` to `Sensor` confirming transaction validation[122].

3.2.2 Interaction

Agents interact by transmitting data to a recipient - `Sensor` agents submitting transactions to `Blockchain` agents, and `Blockchain` agents sending confirmation of transaction validation to `Sensor` agents. Exploration of peer-to-peer data transfer systems - in which interactions occur between `Sensor` agents and between different `Blockchain` agents - holds potential to improve system scalability, but was beyond the scope of this research.

3.2.3 Stochasticity

A degree of stochasticity was included at some points in the model design to simulate variations arising in system functioning. `Sensor` agents used the

[120] A number, array, array of arrays, and so on.
[121] via invocation of `Blockchain.add_to_mempool(tx)` within the `Sensor.transmit()` method body.
[122] `Sensor.confirm_tx(tx)`, called within the `Blockchain.mine_block()` method. Again, the specifics of this process are almost entirely abstracted. This was necessary to conduct this study within its constraints, but the dynamics of information transfer and dissemination through these networks of computing nodes is very far from irrelevant.

stochasticity measure to introduce variability in the number of bytes captured when a new recording was taken. Additionally, if instance variables `Sensor.record_freq` or `Sensor.transmit_freq` were assigned float values between 0 and 1.0 exclusive (rather than unsigned integers), a probabilistic conditional test was applied before the `Sensor.record()` or `Sensor.transmit()` operation was executed. This was meant to inject variability in the frequency of these actions taking place, simulating some on-device logic or intermittent Internet connectivity. The degree of randomness is quite system-dependent; if a more accurate model were being built a more thorough consideration of the process and measure of stochasticity introduced would be necessary.

3.2.4 Collectives

No higher-level groupings of agents were simulated in this version of the agent-based model. However, much about system performance could relate to these collectives and their interaction. Indicating the owners or manufacturers of devices might enable simulation of changes affecting subsets of the network such as a corporate decision-maker - or a hacker exploiting a software vulnerability - removing a number of devices from participation in the network, or changing their behavior somehow. This is of special concern if edge devices are in custody of private keys controlling financial assets, trusted hardware, as outlined. A malicious actor co-opting a botnet of such sovereign devices could carry significant financial consequences (Sabanal 2016).

3.2.5 Observation

Upon completion of all agent activation, on each tick agent-level data was collected for further analysis.

The mean gwei spent per sensor was calculated for each model run to gain insight into the effects of network size on financial costs to system participants. Transaction costs disincentivize malicious or wasteful behavior; the higher marginal costs represent a primary way public blockchains differ from private implementations and traditional data storage and cloud computing resources.

A metric representing the proportion of information captured that has been submitted, validated and represented on-chain was developed, termed "informational currency". [123] The metric based on a rolling window of the most recent 30 blocks mined, $IC = N_{records\ mined} / N_{records}$.[124]

Additionally, summary statistics describing transaction mining times was collected upon termination of the model run and a mean mining time was calculated by averaging the difference between block submitted and block mined for all transactions.

3.3 Details

3.3.1 Initialization

Upon initialization, a `Blockchain` object is instantiated with an empty mempool. A specified number of `Sensor` agents are also instantiated with instance variables specified, including reference to the `Blockchain` instance present. Sensors are added to the model scheduler. Swept variable parameter ranges are shown in Table 6.

Table 6: Initialization variables - swept parameters

Object class	**Variable**	**Parameter value or range**
`SensorBlockchainNetwork` (`Model` subclass)	`num_sensors`	{s: s = n * 50, n ∈ [1, 2, ... 10] }
`Sensor` (`Agent` subclass)	`record_freq`	[0.01, 0.1, 0.3, 0.5, 0.7, 0.9, 1]
	`record_bytes`	{r: r = 1 + n * 20, n ∈ [0, 1, 2, ... 25] }

[123] Calculation of mean informational currency for each iteration excluded the warm-up period of 30 ticks. The metric quantified the proportion of data captured at the edge reflected on chain, having completed the full process of recording the data, transmitting a transaction, and that transaction being validated through mining.
[124] Note that record size was not incorporated; more sophisticated informational currency metrics will be necessary if heterogeneity in edge recording dynamics is simulated.

3.3.2 Input

No environmental inputs were introduced during model execution; all activity was in response to pre-programmed agent behaviors and state variables as initialized.

To move toward a facsimile model of connected sensor networks, the programming of responsiveness of sensors to environmental conditions represents an interesting line of further research[125].

3.3.3 Submodels

See Appendix 1 for details of submodels.

Models, data and iPython notebooks can be found at https://github.com/robisoniv/sovereign-sensors.

---

[125] For example, some sensor monitoring coral reef health could transmit a transaction in the event of reef damage, triggering the release of funds to some local repair crew. This type of parametric insurance is being considered by the insurance industry; Willis Towers Watson recently launched a reef insurance project for sites in central America (Vincent 2018). It is unknown, however, if such parametric policies are currently connected to smart contracts holding funds for disbursement, or how parametric triggers are monitored and policies paid in the event of conditional thresholds being exceeded. A recent Lloyd's of London report suggests it is unlikely that such a system has been connected to smart contracts on a blockchain mainnet holding real funds (Lloyd's 2019).

4 Results

Analysis of data captured during model runs revealed interesting system dynamics, though questions remain regarding whether the patterns observed reflect system operation in reality.

This investigation sought insight into the effects of three independent variables on three dependent variables. Interesting and unexpected results will be reported; all other results are included in Appendix 2.

4.1 Network size

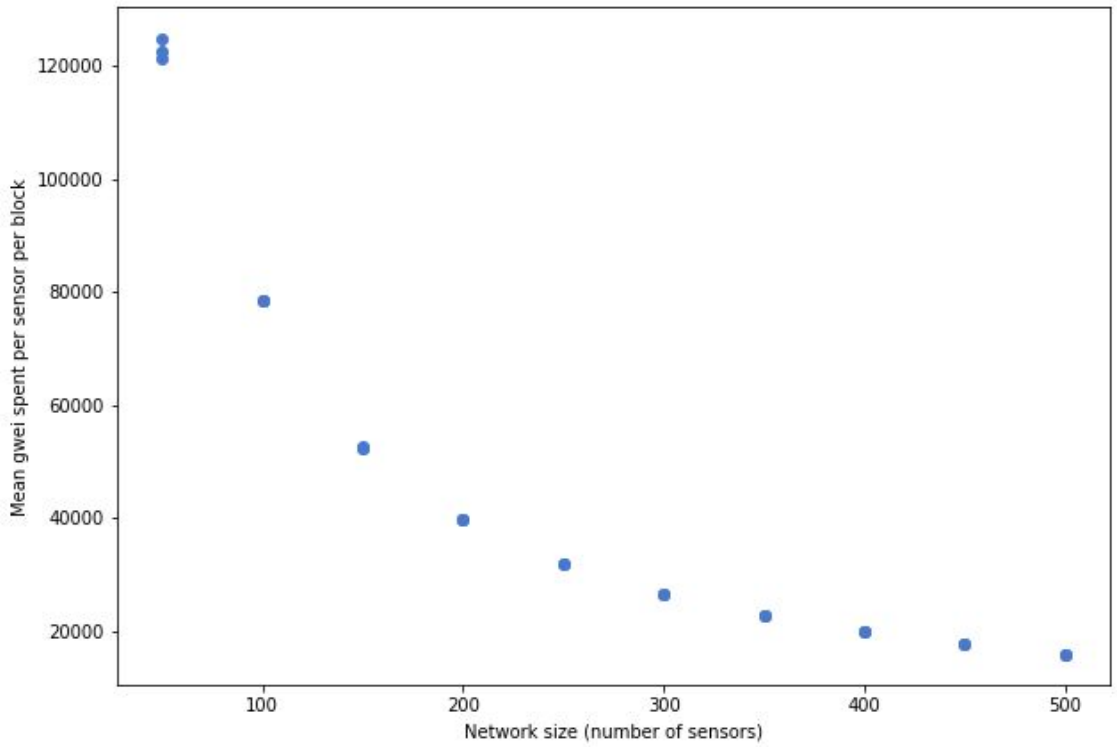
The network size was defined as the number of sensors submitting transactions to a blockchain. Model simulations were executed ranging from 50 to 500 sensors, on intervals of 50.

4.1.1 Gwei spent

Table 8: Summary statistics, Mean total gwei cost per sensor across network size parameter sweep

n iterations		30
Total gwei cost per sensor	Mean	10.44038
	σ	0.64924
	Minimum	9.67701
	Median	10.27899
	Maximum	11.73449

Figure 8: Network size against mean gwei spent per sensor per block



In Figure 8, a concave negative non-linear relationship between network size and mean gwei spent per transaction is visible.

A plot of a log transformation of both variables yielded an apparent negative linear relationship. An OLS simple linear regression analysis produced a model with an R-squared value of 0.9938, indicating that the network size strongly predicts mean gwei spent per sensor. Figure 9 visualizes the scatter plot of the transformation and the calculated line of best fit.

Figure 9: Log of network size against mean gwei spent per sensor per block

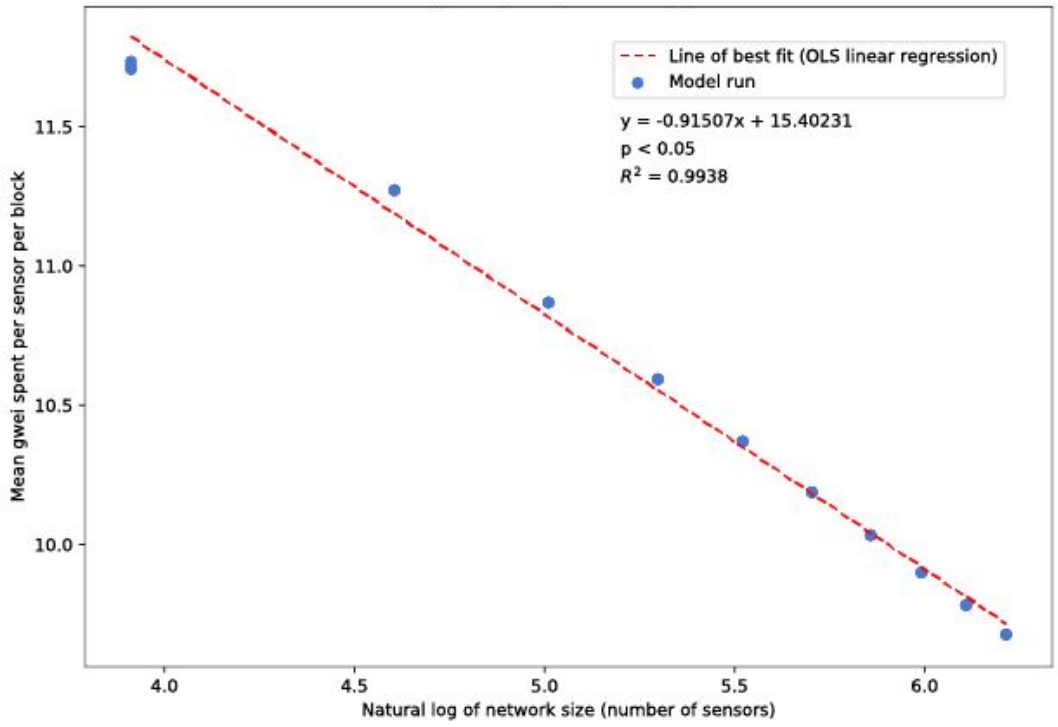
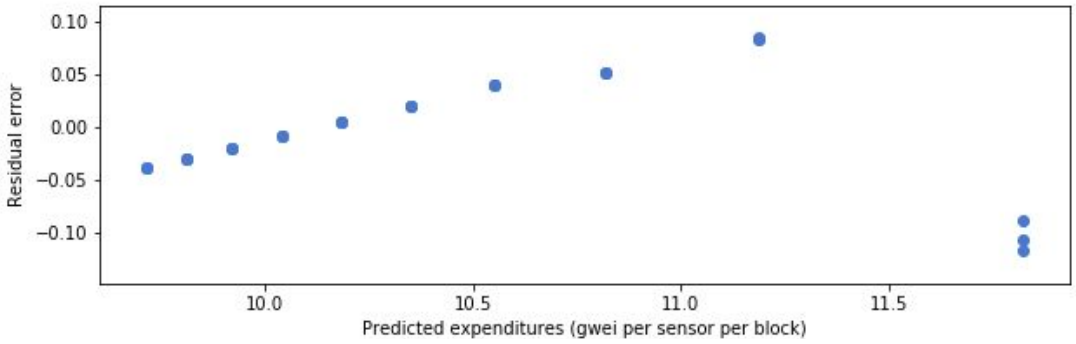


Figure 10: Residual errors versus gwei spent per sensor per block predicted from network size



Visual inspection of the residual errors plotted against predicted values based on the model (Figure 10), however, suggests that the errors are not normally distributed, invalidating the OLS simple linear regression as a method for assessing the significance of the relationship between the variables. Still, the negative relationship between the variables is irrefutable.

As with mining dynamics, the decrease in gwei spent per sensor per block observed with increasing network sizes is caused by the static block gas limit. As the number

of sensors attempting to submit data to be stored on chain increases, the likelihood of each sensor's transactions being validated, contained data being written, and gas costs incurred, decreases. Costs in gwei are only deducted from the sensor's externally owned account upon transaction validation.

Oddly, in model runs with networks larger than 150 sensors, mean gwei costs per sensor were the same across the three iterations executed at each network size simulated, while some, albeit minor, variation was observed in smaller networks (Table 9). Due to the inclusion of the stochasticity metric, as well as a probabilistic record frequency value, this was not expected. Regardless, it almost certainly is a quirk of model design and not an indicator of actual system behavior.

Table 9: Standard deviation of mean gwei spent per sensor per block across identical model runs

Network size (sensors)	**σ - mean gwei spent / sensor / block**
50	0.014425
100	0.000423
150	0.000339
200	0.0000
250	0.0000
300	0.0000
350	0.0000
400	0.0000
450	0.0000
500	0.0000

4.1.2 Informational currency

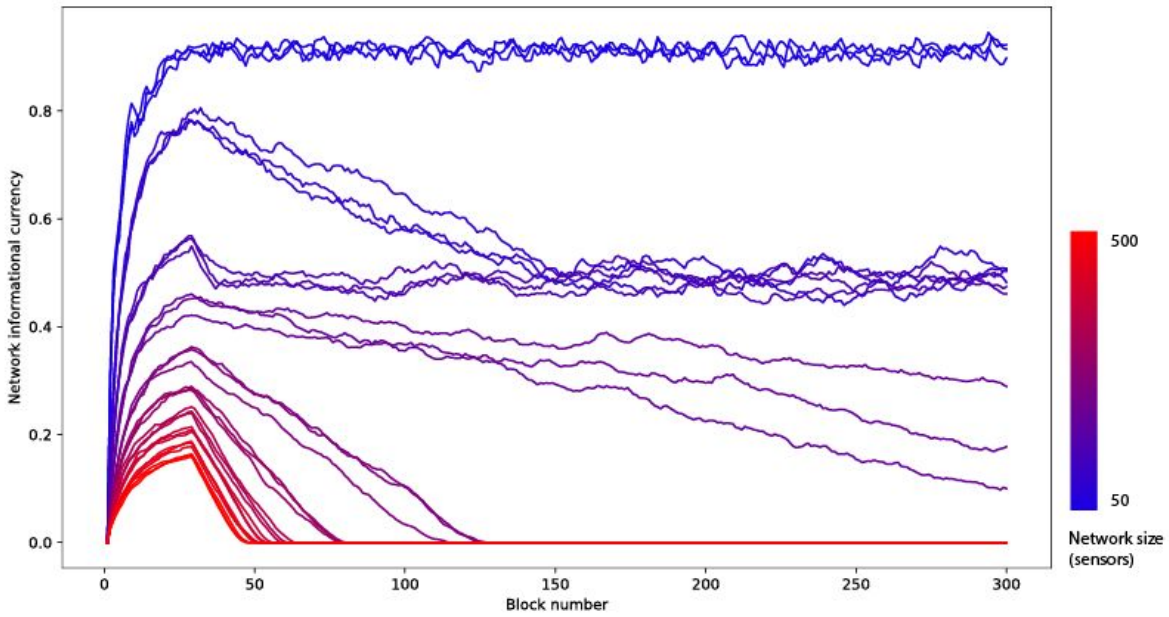
Interesting and unexpected patterns of network informational currency were observed across the parameter sweeps - and across the time series of individual model runs. A linear negative correlation between network size and mean

informational currency values was expected, as was stationarity within model run time series plots. While the pattern expected was observed across model runs, this was not the case within individual iterations.

4.1.2a Time Series

Figure 11 depicts time series plots of mean network informational currency by tick for each of the model iterations recorded, colored by network size. In the smallest networks (50 sensors), informational currency rapidly rose in early ticks and - prior to tick 30, at which point the rolling window meant older blocks were excluded from the measure calculations - reached a maximum value of > 0.9.

Figure 11: Informational currency over time through network sizes



After the warm-up period for these 50-sensor networks, mean network informational currency was a stationary process, as Augmented Dickey-Fuller test results show (Table 10).

Table 10: Augmented Dickey-Fuller test results, mean informational currency over time for 50-sensor networks

Iteration	ADF Statistic	p-value
1	-7.2910	< 0.05

2	-5.3041	< 0.05
2	-6.3199	< 0.05

p-values < 0.05 indicate it is unlikely that these time series do not have a unit root; they are stationary.

Networks larger than 50 sensors exhibited unusual behavior: after achieving a maximum mean informational currency after warm-up, the measures decayed over time until a threshold was reached, at which point the process became stationary. The greater the maximum measured value, the slower the decay rates to the floor threshold. Specifically, ~0.5 appeared to serve as a threshold support level; any networks that exceeded this mean currency receded to this point over time, then achieved stationarity there. Networks of 100 and 150 sensors exhibited this behavior. In networks of 250 or more sensors, which did not reach an informational currency of ≥ 0.5 at any point, after the initial 30-tick warm-up period the measure diminished, ultimately receding to 0. The closer these larger networks got to this threshold level, the longer it took for them to recede to 0 - and the slower the rate of decay.

The eventual establishment of stationary processes around these threshold mean informational currency levels was unexpected; it is again unclear if these behavior patterns are due to a quirk of model design or represent a valid emergent dynamic of these complex systems. The decay in the measurements is likely due to the blockchain block gas limit: if more data is being submitted each time step than can be recorded on the ledger, and because mining prioritizes earlier transactions over more recent ones (given equal gas prices), transaction mining times will increase as the number of unvalidated transactions in the mempool grows. Higher mining times and lower informational currency measures can be attributed to the common cause of block gas limits.

However, the threshold mean informational currency of 0.5 observed in smaller networks remains unexplained. Despite thorough review of model source code, no obvious point causing such behavior was found. Further investigation into this unexpected emergent dynamic would help to identify its cause, and to understand if

it represents a result representative of actual system behavior or simply a quirk of model design.

4.2 Recorded Data Volumes

Empirical sensors operating at the edge record data representing some quality of their environment; the information contained in these data has value to other informational entities[126] seeking insight into conditions in the vicinity of that sensor.

Here the effects of changes in data volumes recorded in each observation on the dependent variables of mean transaction mining times (in blocks), financial costs (in gwei) to each sensor and network informational currency are analyzed.

4.2.1 Mining dynamics

A positive relationship between data volumes recorded per observation and mean transaction mining times is visible in Figure 12: as sensors capture more data per observation, transactions tend to take longer to mine (r_{xy} = 0.74258). Excluding two outliers, for which mining times were exceeded the upper Tukey fence[127], a Pearson's correlation coefficient of r_{xy} = 0.78448 was calculated. Summary statistics for the sample excluding outliers are shown in Table 11.

[126] Humans, and other computing nodes, including blockchain networks.
[127] $Tukey_{upper} = Q_3 + 1.5 * (IQR)$; IQR = 4.8671.

Table 11: Summary statistics, Mean transaction mining times across edge record volumes parameter sweep

n iterations		76
Transaction mining time (blocks)	Mean	7.8928
	σ	3.9433
	Minimum	1.0010
	Quartile 1	5.9778
	Median	7.8253
	Quartile 3	10.712
	Maximum	16.535

Figure 12: Sensor data capture volumes versus transaction mean mining time

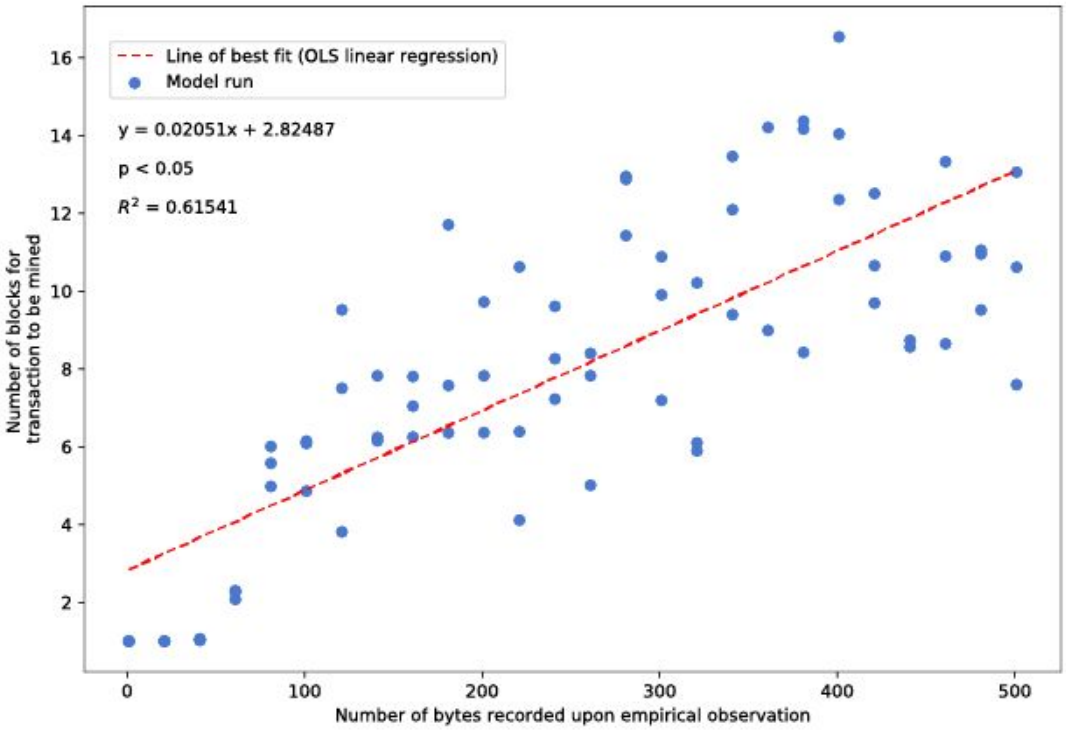


Figure 13: Residual errors versus mean mining time predicted from network size

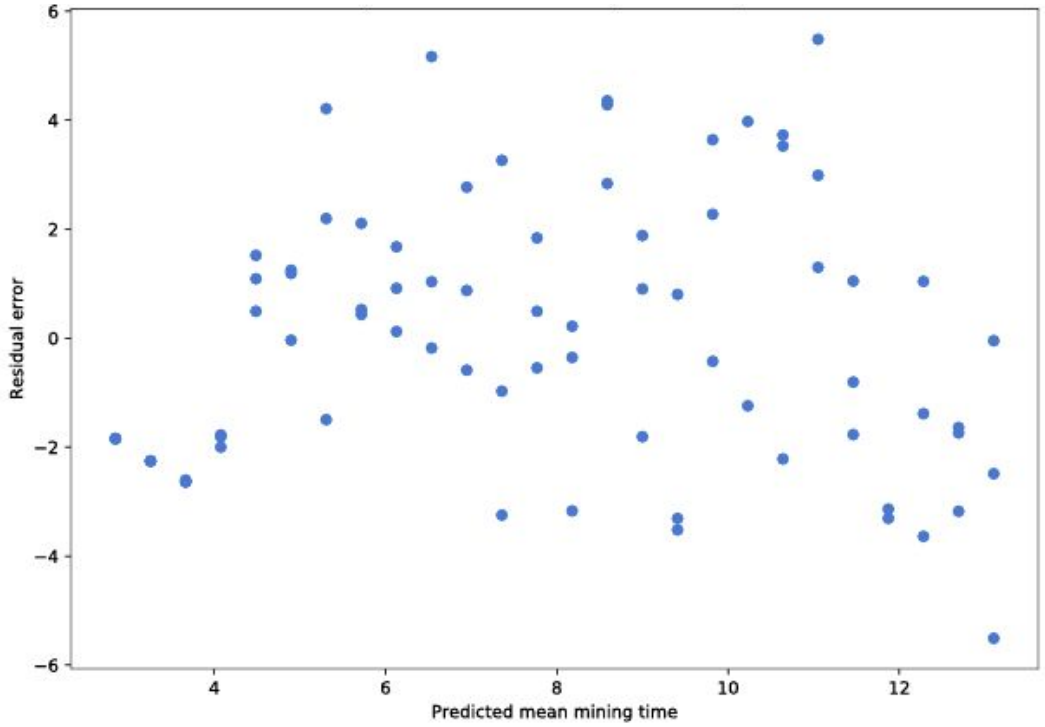
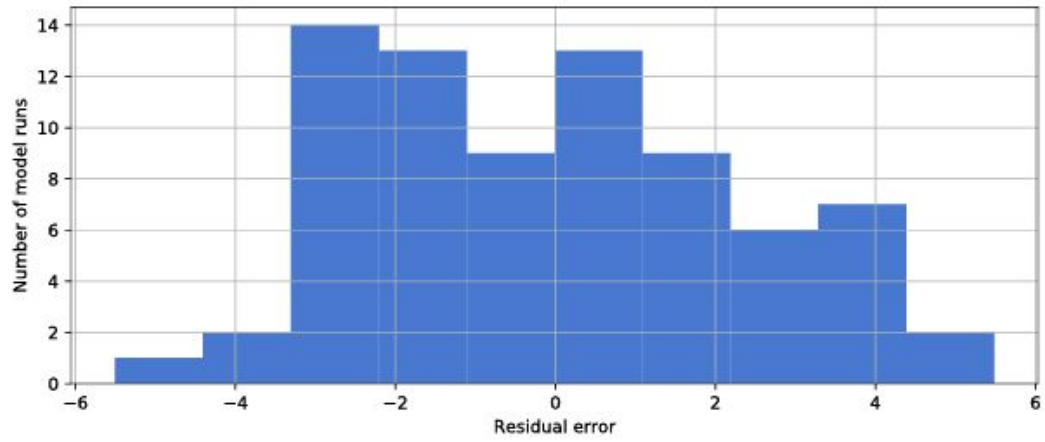


Figure 14: Distribution of residual errors of mean mining times

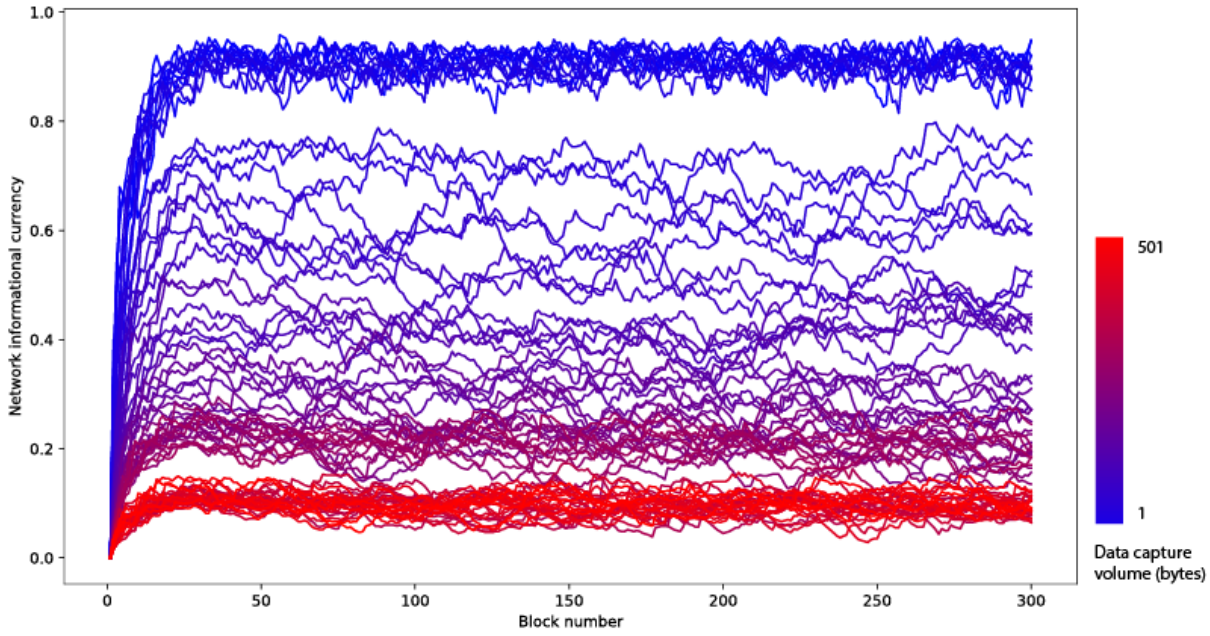


Figures 12 depicts the plot of edge recording volumes against mean transaction mining times with the line of best fit. The calculated R^2 value of 0.61541 indicates that the independent variable accounts for ~61.5% of the variance observed in the dataset. Visual inspection of Figures 13 and 14 suggest that residual error is normally distributed. The mean mining time increased by 0.02051 blocks for each additional byte of data recorded per sensor observation.

4.2.2 Informational currency

As record volumes increased, informational currency was expected to decrease, due to limitations in the amount of data that could be written to the blockchain each tick - the block gas limit. As shown in Figure 15, this was observed: after the warm-up period, in each model iteration a mean level of informational currency was established. Variations around this mean are explained by the stochasticity introduced and the probabilistic conditional transaction transmission frequency.

Figure 15: Informational currency over time across a range of data capture volumes

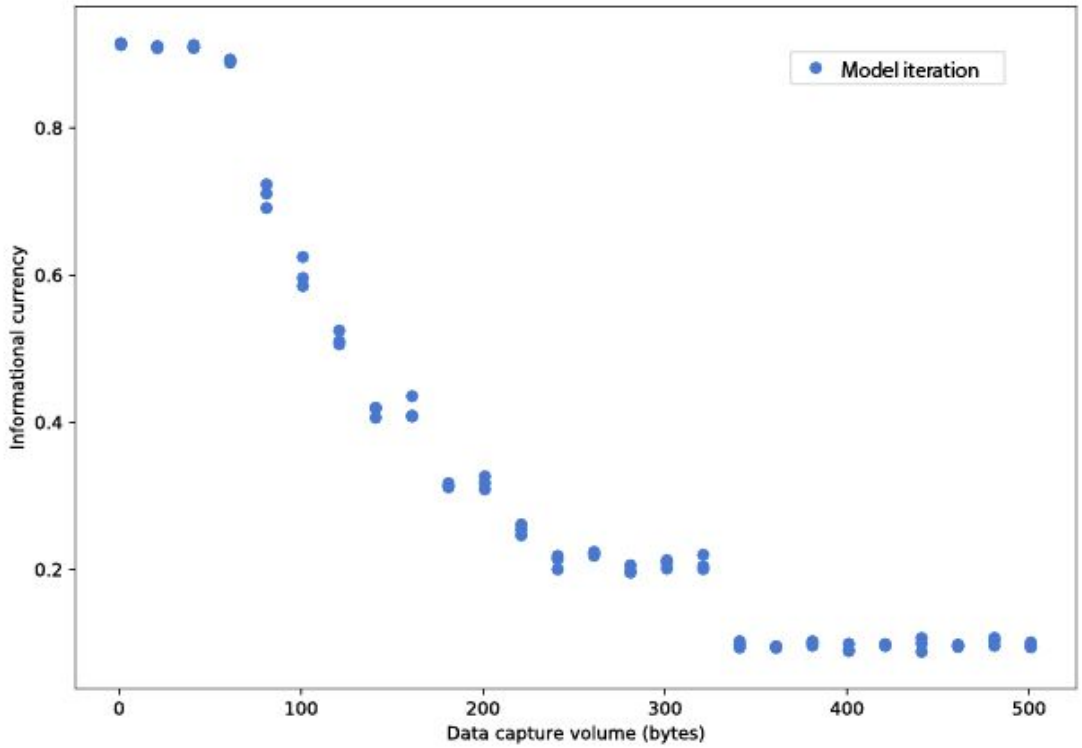


Calculating mean informational currency values for each model run (excluding the warm-up period) enabled the visualization of the effects of data capture volumes on the stable informational currency levels depicted in Figure 16.

A non-linear negative relationship is visible - perhaps a sigmoid curve. Interestingly, as the curve approaches its lower horizontal asymptote at IC ≈ 0.2, a point is reached between 321 and 341 bytes per record where the dependent variable drops

to a new stable level of ~ 0.09. The cause of the discontinuity at this threshold is unclear.

Figure 16: Mean informational currency across data capture volume parameter sweep



4.3 Sensor observation frequency

Sweeping the frequency with which edge sensors recorded empirical observations about their environment was intended to yield further insight into the effects of network activity on blockchain behavior.

Initial examination of results obtained from sweeping record frequencies using fixed parameters as in other independent variable sweeps, it was evident that the model behaved predictably up until near the limit of the sweep. Notably, with fixed network sizes of 20 sensors, the maximum gwei spent per sensor per block of 390000 meant that the block gas limit of 8000000 gwei was never reached.

Since these models were intended to simulate the challenges blockchains might encounter while scaling, and block gas limits are one of the primary constraints to network scalability, the parameter sweep was performed after doubling the number of sensors simulated in each iteration. In these sweeps, block gas limits were reached in the median swept values, enabling the analysis of model behavior as the network transitioned from underutilized to oversubscribed. This is noteworthy because for the analysis below, fixed parameters were identical to investigations into the independent variables' effects conducted above, with the exception of the number of sensors in each network iteration.

4.3.1 Mining dynamics

Initial examination of the scatter plot of mean mining times across record frequencies indicated that for record frequencies below 0.2, transactions tended to be mined immediately after submission. For simulations in which sensors took more frequent recordings, an exponential positive relationship was visible; this segment of the sample was analyzed. An OLS linear regression of record frequencies and the square root of mean mining times yielded a line of best fit shown in Figure 17, with an R^2 value of 0.99243.

Figure 17: Record frequency versus square root of mean transaction mining time

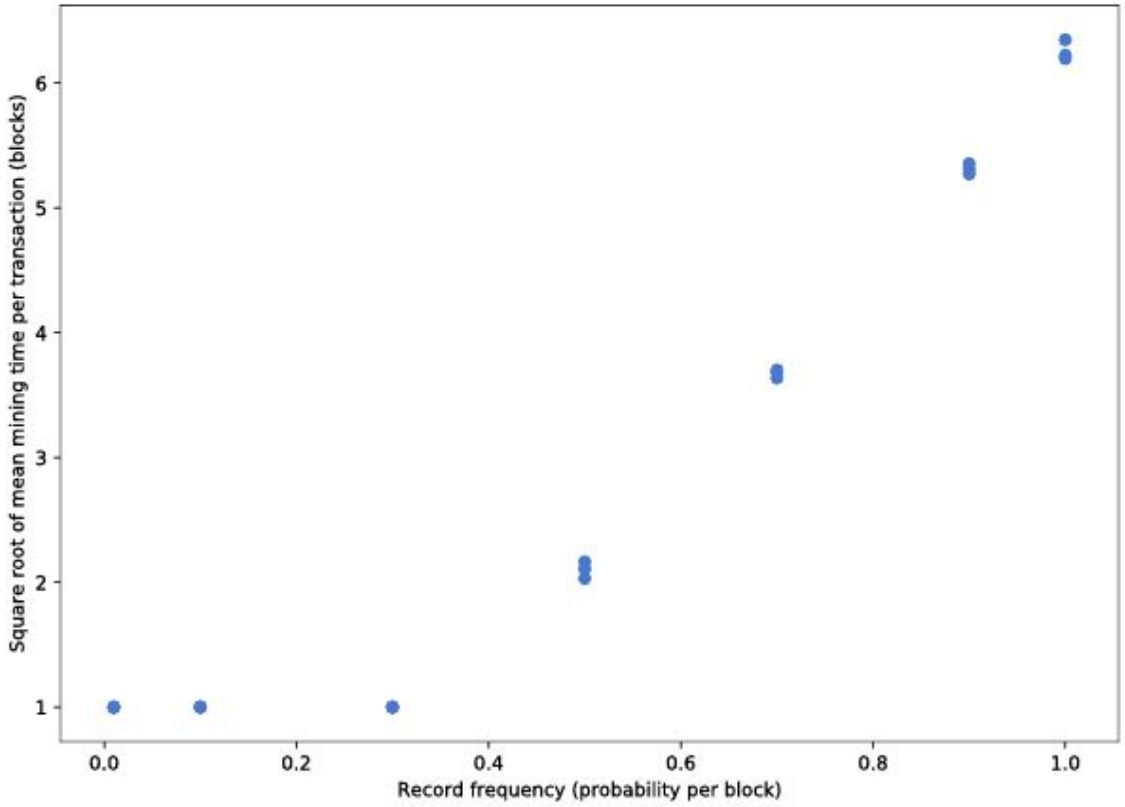
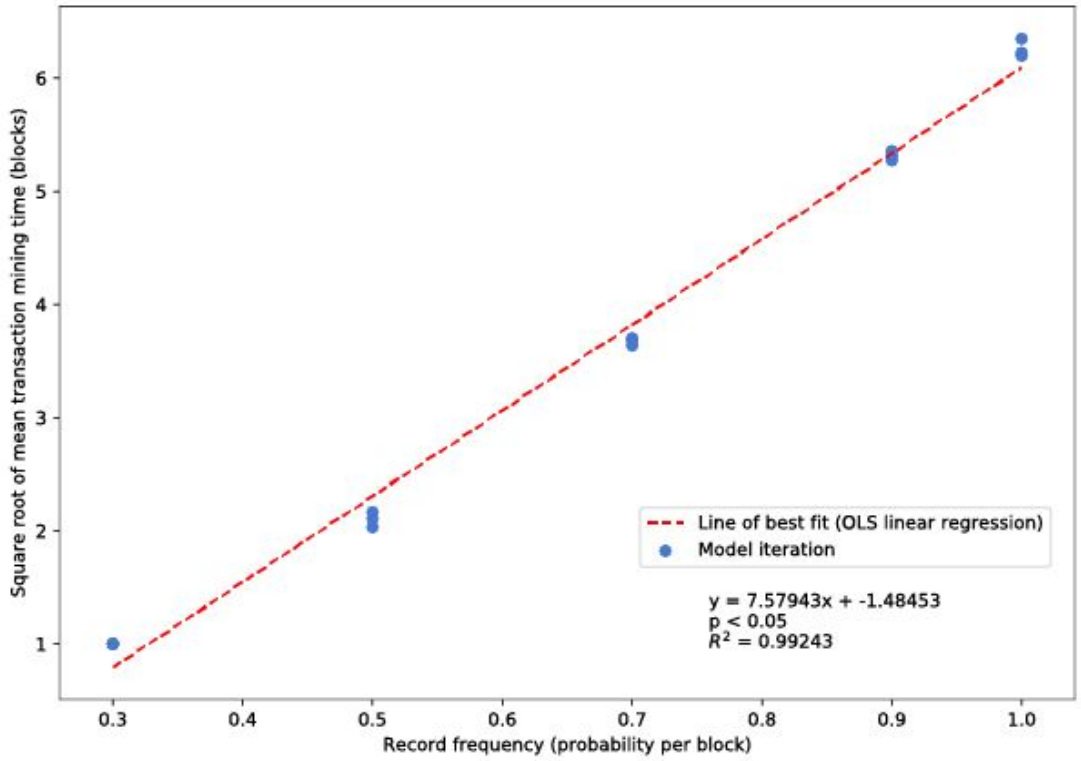
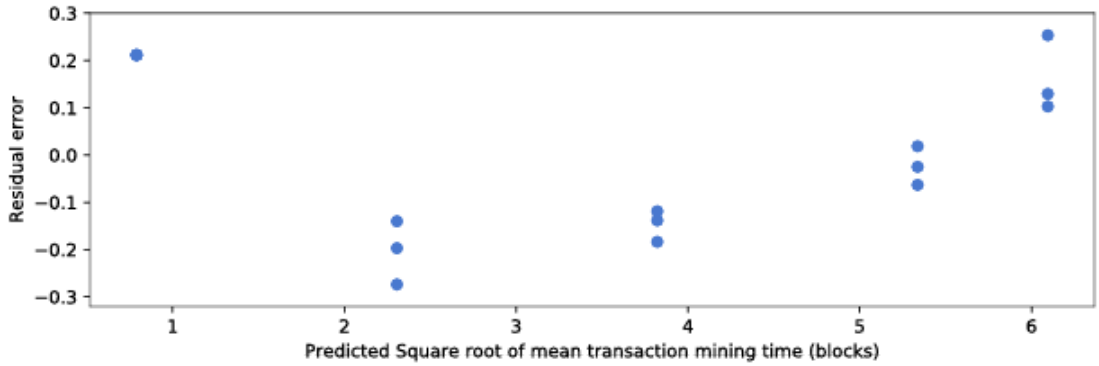


Figure 18: Record frequency (probability per block > 0.2) versus square root of mean transaction mining time (blocks)



Analysis of this plot, and of the residual errors (Figure 19) reveals that the square root transformation did not fully straighten the positive nonlinear relationship observed in the untransformed data.

Figure 19: Residuals versus predicted square root of mean transaction mining time



Performing an additional square root transformation yielded a distribution of residual errors nearer to normal, though uncertainty remains as to the true nature of the nonlinear relationship observed in the untransformed dataset (Figures 20 and 21). Deeper investigation of these nonlinearities is beyond the scope of this analysis.

Figure 20: Record frequency (probability per block, > 0.2) versus double square root of mean transaction mining time

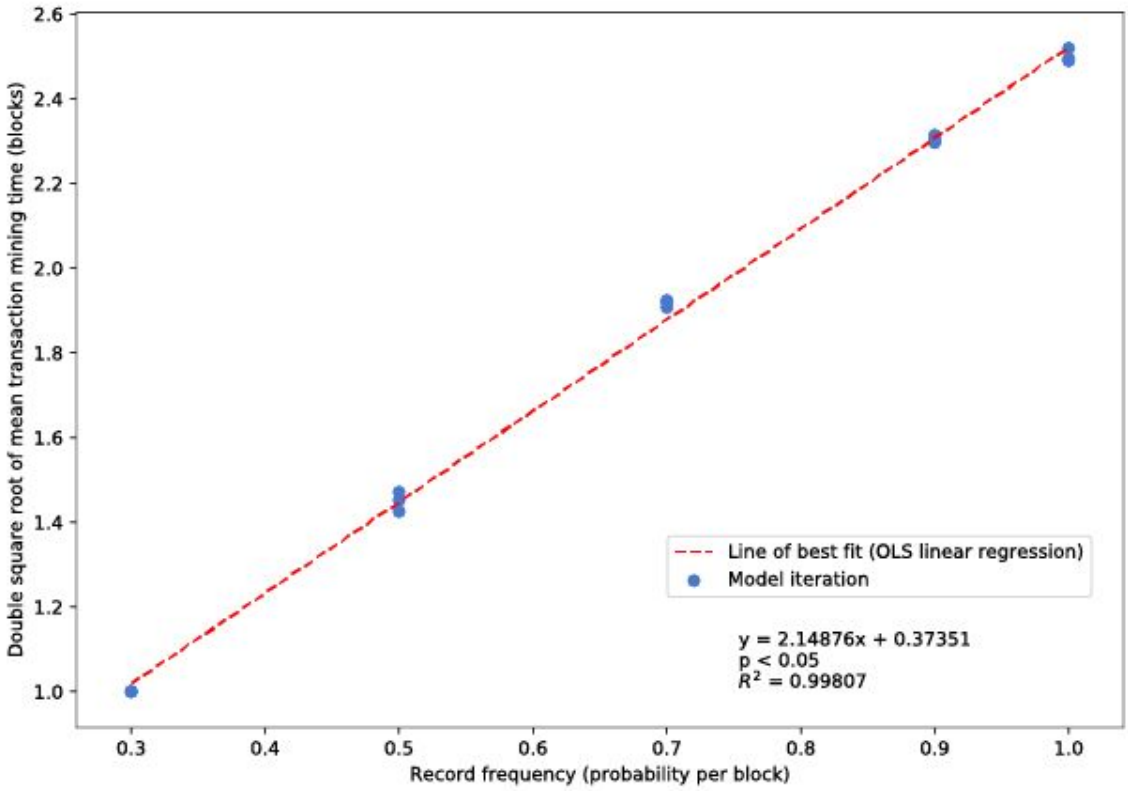
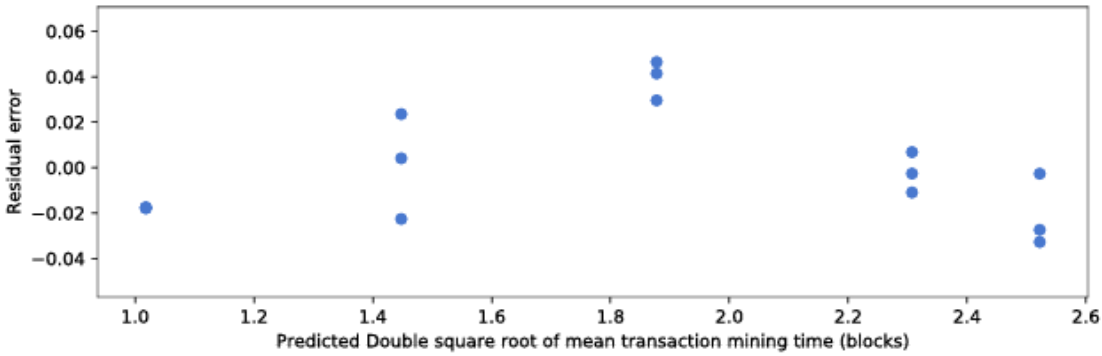


Figure 21: Residuals versus predicted double square root of mean transaction mining time



It does seem clear, however, that once a certain threshold level of transaction activity per block is reached, mining times per transaction rise in a nonlinear fashion. This result is important, and deserves further investigation, as it indicates that blockchain update performance would diminish more and more rapidly as demand increases. Substantial research into mechanisms for managing these scaling challenges due to the inevitable fluctuations in network demand is ongoing [128]; a solution to seems necessary if blockchains are to achieve their potential to form the critical informational infrastructure of the web.

4.3.2 Informational currency

Measures of informational currency collected from model runs across the parameter sweep revealed that a recording frequency of ~0.5 yielded resulted in the highest values. Less frequent updates meant the blockchain was not updated often enough to result in high measures of the metric; as record frequency (and therefore transaction volumes) increased above this optimal level increasing mining times adversely affected network informational currency.

[128] For example, sharding (Ethereum 2019b).

Figure 22: Informational currency over time across recording frequencies

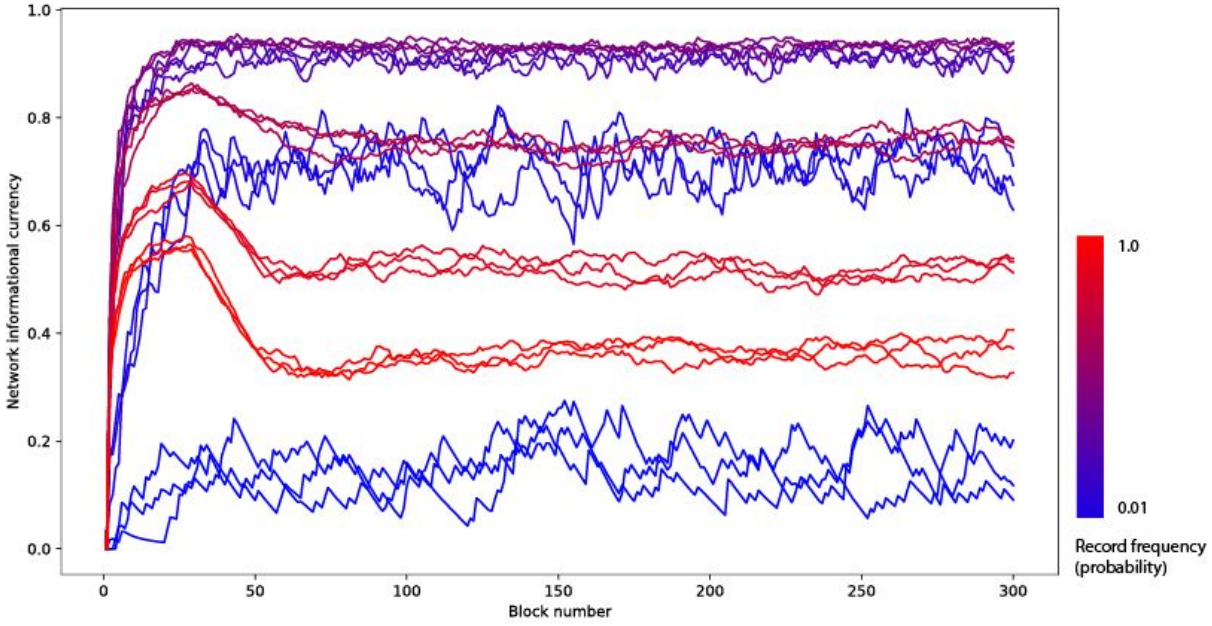
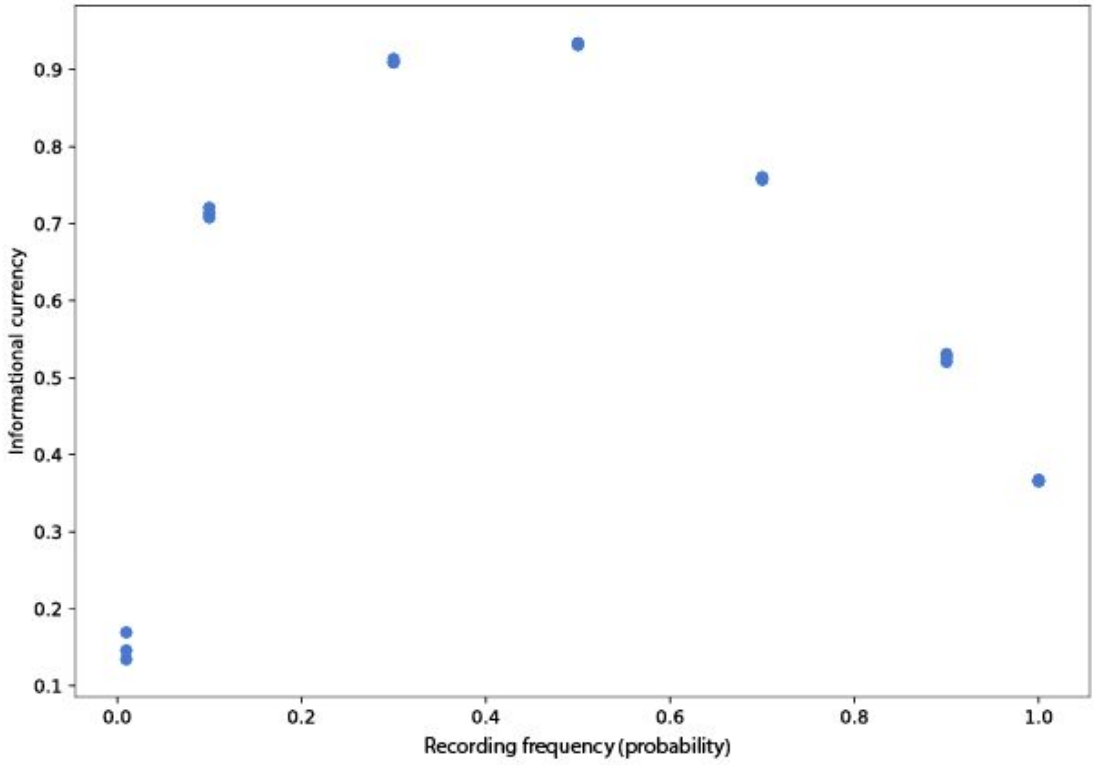


Figure 23: Record frequency vs mean model run informational currencies[129]



[129] Excluding a 30-tick warm-up period.

Interestingly, the model results depicted in Figure 22 show that at higher record frequencies, from a time series perspective within model runs informational currency measures tended to achieve their maximum toward the end of the warm-up period (equal to the window used for measuring the metric), then decayed to a stable level, at which point a stationary process was established. This pattern only occurred with record frequencies greater than 0.5.

It seems that stable states may exist at different transaction volumes. As with other results, distinguishing between authentic system behavior and unintended behavior resulting from model design is difficult. Nonetheless, the result is interesting, and warrants further research.

---

5 Discussion

This research has explored the technologies constituent to a system of trusted sensors acting as oracles to smart contracts. A middle range agent-based model of such a system was developed; initial results were analyzed. The insights gleaned from these simulations and analyses, along with a broader conversation of the implications and opportunities of such systems, are presented here.

5.1 IoT + Blockchain

Public blockchains can only function if the costs associated with utilization of the public resource deter excessive, lazy or malicious use: they represent “an expensive but decentralized and very high-assurance store of data and computation” (Buterin and Weyl 2018). Sensors embedded on devices generate large quantities of data. At first glance, blockchains do not seem a sensible informational architecture to store or compute IoT data. This intuition was validated by the observation of diminishing informational currency measures and increasing costs associated with increasing network loads.

However, techniques such as edge computing and content addressing can dramatically reduce data volumes transmitted from edge sensors without necessarily reducing the user’s access to informational value. Due to the benefits of decentralization described, the computing services blockchains offer should be considered by IoT system designers, especially when the connected sensor network might monitor the public space[130], or external actors might benefit from accessing the information it captures.

While the ideological values of public blockchain systems generally align with the author’s, private blockchain implementations should not be overlooked, as they represent a way for system designers reluctant to relinquish full control to an

[130] This research has raised many questions about the current and future status of data ethics regulation. This author believes strongly that the value captured in a sensor recording should benefit not only the device owner, but also the entities represented in (i.e. subjects of) the recording. The technical barriers to this are significant, but a range of context-dependent solutions can be conceived of.

independent organization[131] to still capture many of the benefits of decentralization, at substantially lower cost. System configuration and governance will largely determine the extent to which such private instances realize these benefits.

Designers of IoT systems connecting to blockchain networks will need to balance the trade-offs between edge resource consumption, financial costs, informational availability and control over data and computing resources. In the near future, focus should be placed the highest-value use cases for connecting edge devices to smart contracts, where the cost could be justified. Parametric insurance products, which trigger a payout if some threshold condition[132] is reached, seem an optimal initial test case (Lloyd's 2019). While this technology is still in its infancy experiments should be conducted such that the harm caused by failure or unexpected contract behavior would be inconsequential.

5.2 Proposals and Recommendations

5.2.1 Ricardian treaties

Trusted sensors connected to blockchain networks might provide a transparent, secure and trusted way for loosely coordinated entities[133] to maintain situational awareness. This author is especially interested in applications of the technologies for sovereign nations holding one another to account. He imagines an agreement between sovereign nations - a valid, legal treaty[134] - that expressly delegates some authority to a smart contract. Contract code would be open to inspection by all. Parties could be required to submit digitally-signed evidence of treaty compliance; violation of terms would automatically trigger agreed-upon enforcement mechanisms. Such technical mechanisms[135] incentivizing proper behavior could

[131] i.e. a blockchain developer community, mining network and so on.
[132] Flood waters exceed a certain height in a building, say, or a flight is delayed by more than a number of minutes.
[133] Especially ones with no trusted central authority.
[134] "A treaty is a formal written agreement entered into by actors in international law, namely sovereign states and international organizations. A treaty may also be known as an international agreement, protocol, covenant, convention, pact, or exchange of letters, among other terms. Regardless of terminology, all these instruments may be considered treaties subject to the same rules under international law." (Wikipedia 2019k).
[135] Such as the forfeiture of staked funds, or the automated enforcement of sanctions on participating states that fail to adhere to agreement terms.

supplement legal ones, which can be slow and difficult to implement. The infinite programmability of such Ricardian treaties could enable the application of a sophisticated system of microincentives to states (Hoopes 2019), perhaps integrated with central bank digital currencies such as the Synthetic Hegemonic Currency proposed by Bank of England governor Mark Carney at the Jackson Hole Symposium in August 2019. It could be an opportunity to "sort out the deep flaws in the international monetary and financial system" - to "change the game" (Carney 2019).

Arms control represents a particularly high impact application of this system of oversight.

As an example, if secured in containers embedded with trusted sensors (Becha 2019), a protocol requiring intermittent data transmissions including location and local environmental characteristics could provide governments confidence that their adversaries are behaving in compliance with mutually-agreed rules.

Of course, privacy and security concerns about the entities authorized to access this sensitive information are valid; it seems that a private, permissioned blockchain implementation[136] would be more appropriate and palatable to sovereign nations considering such an arrangement[137]. Many applications beyond arms control are conceivable, including in citizenship, immigration, international finance, military intelligence sharing, environmental monitoring, international trade, disaster response, conflict mitigation, and so on. Such Ricardian treaties could bring many of the benefits of decentralization to international affairs, and improve sovereign nations' ability to leverage the improving capacity to maintain situational awareness around the world. After all, it could be argued that in the international arena no clear central authority exists, involved parties "have conflicting incentives", and "there is a need for a shared common database" - a context for which blockchains are well-suited (Paul 2018).

[136] Likely with nodes hosted by adversarial governments, as well as entities with a range of incentives - a consortium bound by legal agreements and incentives.
[137] This may be another situation in which homomorphic cryptographic techniques and zero-knowledge and range proofs could provide assurance to adversarial states without revealing additional strategic information. For example, Glaser's (2014) work on "A zero-knowledge protocol for nuclear warhead verification" and other zero-knowledge proof protocols.

5.2.2 A DAO for access control

This work was inspired in part by the author's recognition of an enormous opportunity being missed. Legions of sensors are being deployed and connected to the Internet - in theory[138], able to provide information feeds to people who could derive value from that information. However, due to the current configuration, based on the client-server web paradigm and an approach that is often required that private interests are placed over public ones[139], much of the value to humanity being created is lost as the relevance of the data fades as it is left siloed on its device or in a proprietary relational database.

This is, in this author's view, a tragedy of the informational commons[140]: the failure to share information with someone who would benefit from it. There are certainly valid instances where sharing such information would provide a competitive advantage to the informee, but in many cases this is not so; squandering this value is tragic.

As a solution to this, a peer-to-peer data access protocol is imagined[141]. Edge nodes would need to be visible to a central administrator - in this conception, a DAO hosted on a blockchain. Data consumers could - for a fee, or under certain conditions, or possibly if assessed as deserving by some[142] review mechanism - connect to edge networks and retrieve data for analysis. This could provide sensor owners an additional source of revenue - and they could price the data according to their perception of its value - and it would provide the public an opportunity to extract value from the unprecedented capability for situational awareness emerging.

A thorough exploration and feasibility study of this system is beyond this paper's scope, but it seems that work related to decentralized public key infrastructures (Allen 2015), content addressed storage (Benet 2014), and proxy re-encryption

[138] As in, technically (very often).
[139] Many sensors would not be installed if their owners could not monetize them.
[140] One of three identified. See Appendix 4 for the three tragedies described.
[141] Based on this author's limited understanding, likely employing IPFS content addressing technology, or as recently proposed by Chen, Ramsundar and Robbins (2019).
[142] Ideally decentralized.

techniques (Nuñez 2018) might resolve many of the most difficult challenges to its implementation.

5.2.3 The voluntary transition to self-sovereign identities

Centralized authorities that have entrenched themselves as society's informational infrastructure through the Web 2.0 period now hold much of the data about us[143]. It has been compellingly argued that "forces of intelligent persuasion" "undermine the integrity of the human will" (Williams 2018), and that data should be treated as labor rather than capital (Ibarra 2017).

As the legal custodians of the personal information of billions of humans, these centralized data custodians should voluntarily lead the transition of data custody to their self-sovereign users. Acknowledging the "irrationality"[144] and possible legal challenges in doing so, it is, in this author's view, clearly the right and sustainable thing to do.

[143] This dissertation is being composed on a Google Doc. Does that mean it belongs to Google? Should Google be allowed to read it as it is being written?
[144] Economic irrationality. Rationality exists on many spectrums and varies with the scope of perspective.

6 Conclusion

This research was motivated by my curiosity about the world and guided by a dawning understanding of the magnitude of the opportunity and risk we face in the coming decades. Information is a fundamental component of our reality[145], and yet we have only begun to grasp its nature. In this moment - an apparent inflection point - I am compelled to leverage my privilege, skills and knowledge to promote the ideals I hold.

Computers are in many ways the crowning technological achievement of the modern world, providing an intensely useful tool for us humans. As a result, our globalizing society is weaving them into the fabric of our lives, meaning the physical and informational infrastructure we rely on to sustain civilization increasingly depends on these devices and the insights they yield.

As with any tool, computers are ambivalent to their application[146]: it is the users who decide if they will be used morally or not. And like every innovation, their intrinsic unfamiliarity has left us uncertain of how to properly govern them, and of the effects they have on human dignity and planetary well-being when adopted at scale.

These three qualities - the usefulness, ambivalence and unfamiliarity of computing technologies - explains my interest in pursuing this research. To invoke Greta once again: “We are failing but we have not yet failed” (Thunberg 2019 pp 44). Mistakes are to be expected, missteps and misjudgments forgiven. What is no longer acceptable, however, is to move fast, break things, then fail to learn, adapt, and evolve our systems and laws and norms and cultures and behavior in response to our discernment of harm and inequity - even if it is difficult to do so.

Much as cooling water reaches a point at which its molecules become ordered and aligned, a similar phenomenon might occur as we approach saturation of

[145] Both individual and shared - perhaps another interesting aspect of information and other conceptual objects is that they are all that comprises each of our individual realities. Physical reality is shared - it is through interpretation we find incompatibility.
[146] Although - as informational entities - do computers have the potential to care?

information transfer between sentient entities. As understanding improves, entities might adapt to be less likely to act in a way that would disrupt another's intentions. Deconfliction of behavior may be possible[147].

Toward a theory of conceptual reality

Reflecting on the breadth of technical, scientific, ethical, political and economic reading and learning I have done in this research effort[148], I am struck by a missing link, one that I am coming to realize is fundamental to a complete and coherent worldview inclusive of the informational domain.

Communication is the transfer of meaning between informational entities. Shannon (1948) formalized a mathematical theory of communication, based on digital[149] data. However, not all communication is mathematical in nature[150]. While Shannon's definition of a "message" is useful in his context, it is obviously only partial. Further, it seems to be a profound simplification to reduce the model of communication to a one dimensional series with only two radices - the two binary digits[151].

My recognition of this missing link is pointing to an interesting finding: information does not objectively exist. Information is subjectively perceived, and only exists within the awarenesses of sentient[152] agents. Furthermore, data seems to *only* objectively exist - it is physical. Data can be stored on a physical object as matter[153] or be transmitted as physical energy through space. Data is governed by the laws of physics[154]. For the information contained within data to be discerned, it must be sensed and interpreted by another informational entity[155].

---

[147] These ideas on another scale entirely: could entangled (Simonsen 2018) blockchain networks offer a solution to the dark forest theory of deterrence (Liu 2008)? 🌀
[148] Incomplete though it is ...
[149] In his case, binary.
[150] Perhaps none of it is? Communication is the transfer of meaning *through space and time*. Mathematics exists in the conceptual space - in which neither time nor space exists.
[151] See Appendix 7.
[152] If "sentient" means able "to perceive one's environment" (State of Victoria 2017).
[153] And thereby persist through time.
[154] Albeit quantum physics, which is less well understood than classical mechanics.
[155] i.e. in the awareness of the informee.

This distinction may resolve many of the questions left inadequately addressed by the literature reviewed and the contemporary paradigm. A complete investigation into its validity and implications is well beyond the scope of this dissertation, but initial ideas and observations are outlined in Appendix 7.

Of particular interest is the difference between physical reality and conceptual reality entailed by this distinction. Physical objects exist in the physical space, and adhere to the physical laws - these are well understood. Conceptual objects[156], however, do not appear to exist in a physical space[157], but rather in conceptual space, only manifesting within the awareness of a perceiving[158] entity. As conceptual objects do not exist in physical space, they do not adhere to the physical laws[159].

So - what are the laws governing conceptual reality?

[156] Concepts? Are concepts the same as conceptual objects?
[157] Or at least do not primarily manifest there. Initial enquiry suggests that each conceptual object is necessarily grounded in some physical manifestation - see Appendix 7.
[158] Also, informational; discerning of meaning.
[159] See Appendix 6 for an exploration of an example conceptual object.

Appendices

Appendix 1: Submodels

On model instantiation a single instance of the `Blockchain` class is created. The specified number of `Sensor` objects are created and the model `Blockchain` object is added to each as an instance variable[160]. `Sensor` instances were then added to the model scheduler - but the `Blockchain` object was not.

Each tick a number of operations were executed.

In an order randomized each step, `Sensor` agents were activated by invoking their `Sensor.step()` methods. If the `Sensor` was active[161], this process began with the invocation of the `Sensor.record()` method. If the sensor was determined to attempt a transmission that tick[162], then a `Sensor.transmit()` method simulated the preparation and signing of a transaction, transmitted to the blockchain network [163].

[160] For these model runs the `Blockchain` object could have been assigned the `Sensor` objects as a class variable. However, this design pattern was chosen to leave room for a model design that includes the simulation of multiple blockchain networks, with different sensors connecting to different chains - or even the same sensor submitting transactions to different chains based on some logic executed on the device. In this way, each sensor attained the ability to invoke multiple `Blockchain` object methods.
[161] In order to enable modeling of edge sensors constrained by limited energy supplies, a `Sensor.mortal` boolean was included. If the `Sensor.battery_life` instance variable was depleted below zero, the `Sensor` instance would not perform on-device actions such as recording and data transmission. These sensors were not simply removed from the scheduler because they might still have unvalidated transactions pending in the mempool, and needed to be accessed via model scheduler for the invocation of `Sensor.confirm_tx()` method by the `Blockchain` object upon validation. This would ensure complete recording of costs. For the simulations analyzed here, however, edge devices were not mortal - i.e. they were connected to a reliable power source, rather than a limited battery.
[162] Based on comparison of its `transmit_freq` variable with either a randomly-generated float value (probabilistic) or the current block number (deterministic, interval-based).
[163] A `Sensor.compute()` method to simulate data reduction by analyzing data on the edge device was defined, but not utilized in the batch runs performed. This code would enable modeling of energy usage at the edge, and would have implications for each of the dependent variables measured in model runs: mining times, financial costs and informational currency. This method could simulate the generation of a hash (perhaps an IPFS Content ID) on board the device, to be registered on chain. This could help system users ensure that edge data integrity has been maintained. Alternatively, summary statistics (or

Within the `Sensor.transmit()` method execution the sensor invoked the instance's `Blockchain.add_to_mempool(tx)` method, simulating the submission of the valid, signed transaction to the blockchain miners to add to the queue of unvalidated transactions.

Within the model `SensorBlockchainNetwork.step()` method, after randomly activating each `Sensor` agent, the `Blockchain.mine_block()` method was called [164], in which a subset of unvalidated transactions was selected from the mempool[165]. Upon transaction validation, the `Blockchain.chain` dataframe was updated such that the appropriate sensor's column reflected the current state of on-chain data availability. Additionally, the appropriate sensor's `Sensor.confirm_tx()` method was invoked, simulating the deduction of the gwei spent for transaction validation from the agent's account.

---

some other information reduction technique) could enable a balance to be struck between edge compute resource usage and wireless bandwidth usage.

[164] This aspect of model design would need adaptation to simulate the operation of multiple blockchain networks. If the activation of all `Sensor` agents prior to the `Blockchain` agents mining action is as important as assumed, the `RandomActivationByBreed` subclass defined in Mesa's `wolf_sheep/schedule.py` example may achieve this (Core Mesa Team 2018).

[165] Transactions were selected for inclusion in a block based on transaction value, to simulate the miners' incentive to mine highest-value transactions first, then block submission time, to prevent transactions arbitrarily remaining unvalidated (getting "stuck"). It is worth noting that transactions can get stuck in the Ethereum mainnet mempool, remaining unvalidated (usually because the gas price is so low that miners do not select it for validation) (McDonald 2017). It was decided to remove this feature to simplify the modeling of network informational currency; it is unknown how its inclusion might affect model behavior. Furthermore, although gas prices were fixed in this research, the modeling of variable gas prices could shed light on the costs of increased informational currency in heterogeneous agent networks. This approach to transaction selection should capably handle such an adaptation.

Appendix 2: Additional Results

Network Sizes

Mining Dynamics

A positive relationship between network size and mean transaction mining times per transaction was observed: larger networks tended to validate transactions more slowly.

Table 7: Summary statistics, Mean mining times per transaction across network size parameter sweep

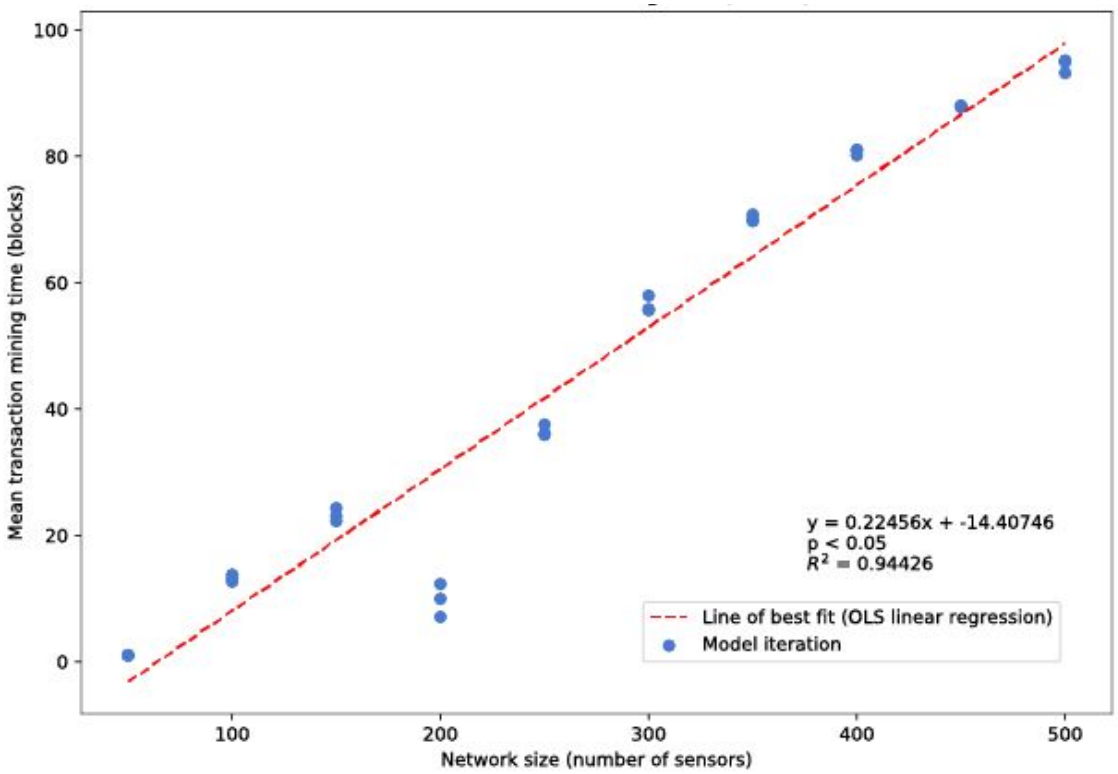
n iterations		30
Transaction mining time	Mean	47.34600
	σ	33.75534
	Minimum	1.00404
	Median	46.59344
	Maximum	95.18947

The observed mean mining times exhibited a positive correlation (Pearson's correlation coefficient r_{xy} = 0.971728). It stands to reason that this correlation is due to the block gas limit, which fixes a limit on the amount of data the network can write each tick. Transactions each specified the same gas price, and data capture volumes and transmit frequencies were homogeneous[166] across sensors, so network size directly correlated with the volume of transactions being submitted to the blockchain network for validation. Because blocks have a fixed limit to the gas that could be consumed - here, the amount of data to be written[167] - in larger networks generating more data, transactions will take longer to be validated.

An ordinary least squares simple linear regression analysis was conducted to produce a line of best fit modeling the relationship between network size and mean mining times.

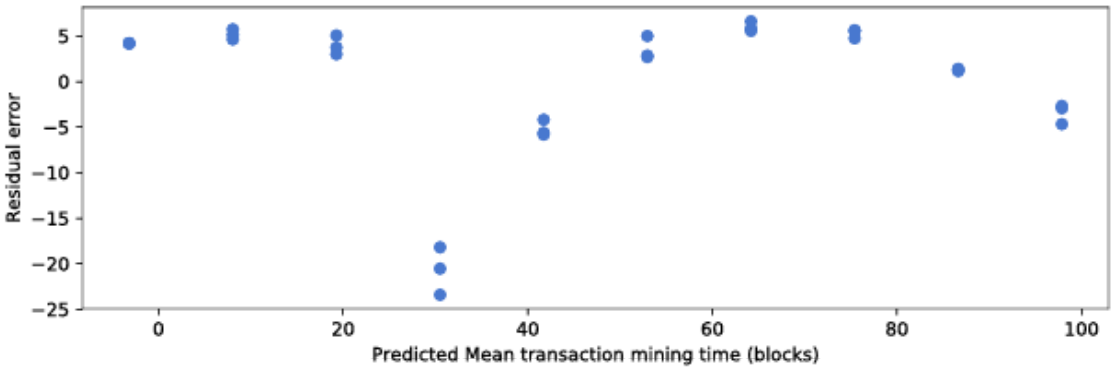
[166] Though stochasticity did cause slight variation.
[167] Since on-chain compute operations were excluded in this model.

Figure 6: The effect of network size on mean mining time per transaction



While the correlation coefficient indicates a strong positive correlation between the variables considered, the residual errors do not appear to be normally distributed (Figure 7), calling into question the validity of the line of best fit calculated using OLS simple linear regression.

Figure 7: Residual errors versus fitted values, OLS simple linear regression on network size versus mean transaction mining times



Curiously, the observed values at a network size of 200 breaks with the trend established in smaller network sizes; mean mining times were, on average, less than 50% of the values observed in sensor networks of 150 sensors. It is unclear if this break in the otherwise consistent positive relationship between the two variables is due to some quirk of the model or true emergent behavior of these systems interacting in reality; the former seems likelier, but without a dataset to validate observations from the simulation this is difficult to assess.

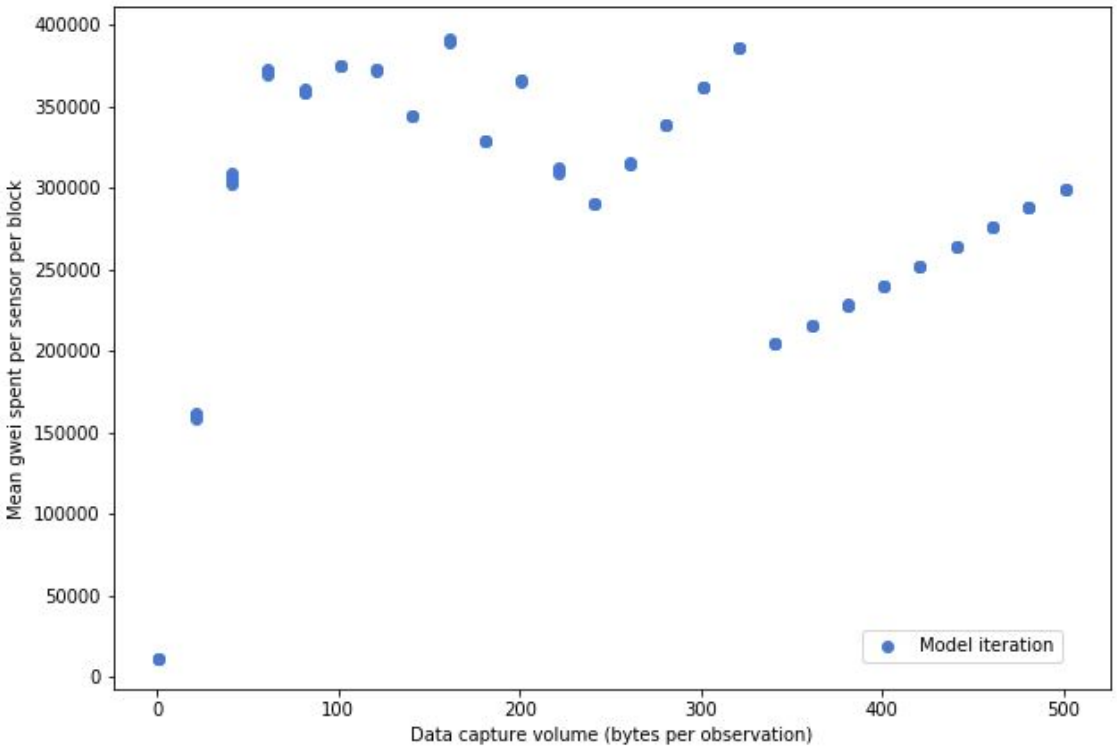
Visual inspection of the scatterplot suggests the possible subsequent establishment of a new, convex positive non-linear relationship for networks of 200 or greater sensors. Further investigation of the nonlinear effects of transaction volumes on mining dynamics is warranted, but beyond the scope of this investigation.

Recording Volumes

Gwei spent

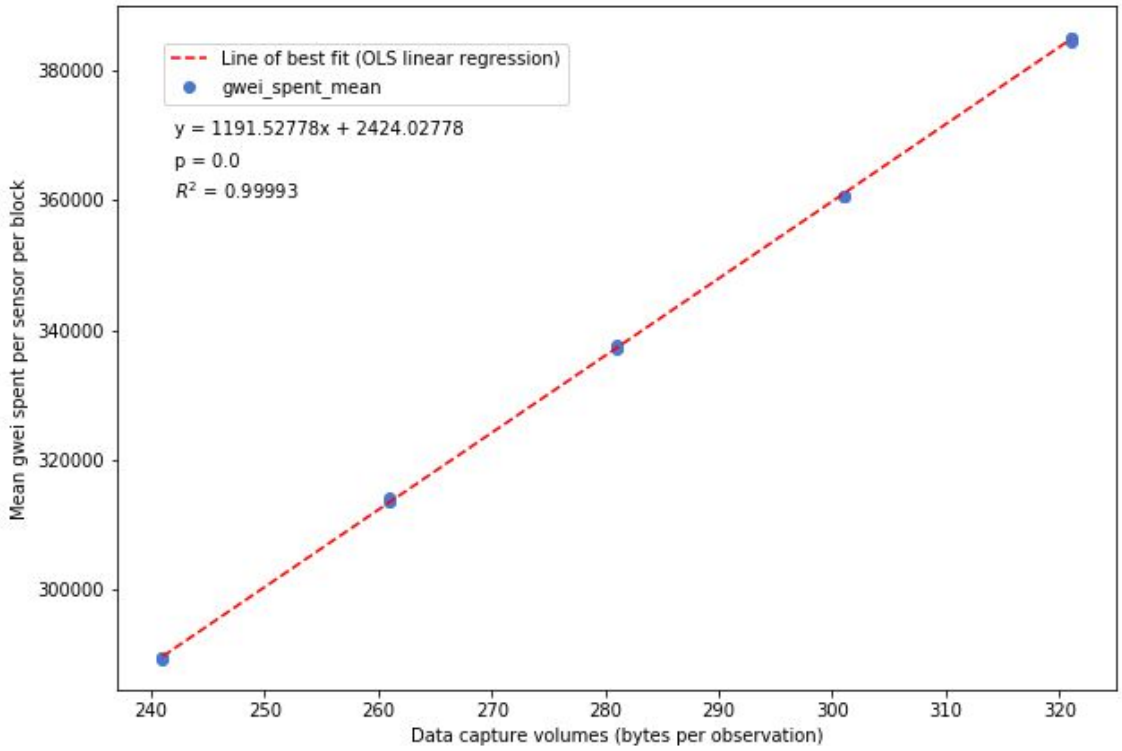
As Figure 24 shows, recording volumes had an unusual effect on gwei spent per sensor. At smaller record volumes, the financial expenditures rose rapidly with the independent variable, quickly leveling off between 350000 - 400000 gwei spent per sensor. These mean recordings appeared to begin to decrease as record volumes approached 241 bytes, at which point two highly uniform segments of data are seen.

Figure 24: Data capture volumes against mean gwei spent per sensor



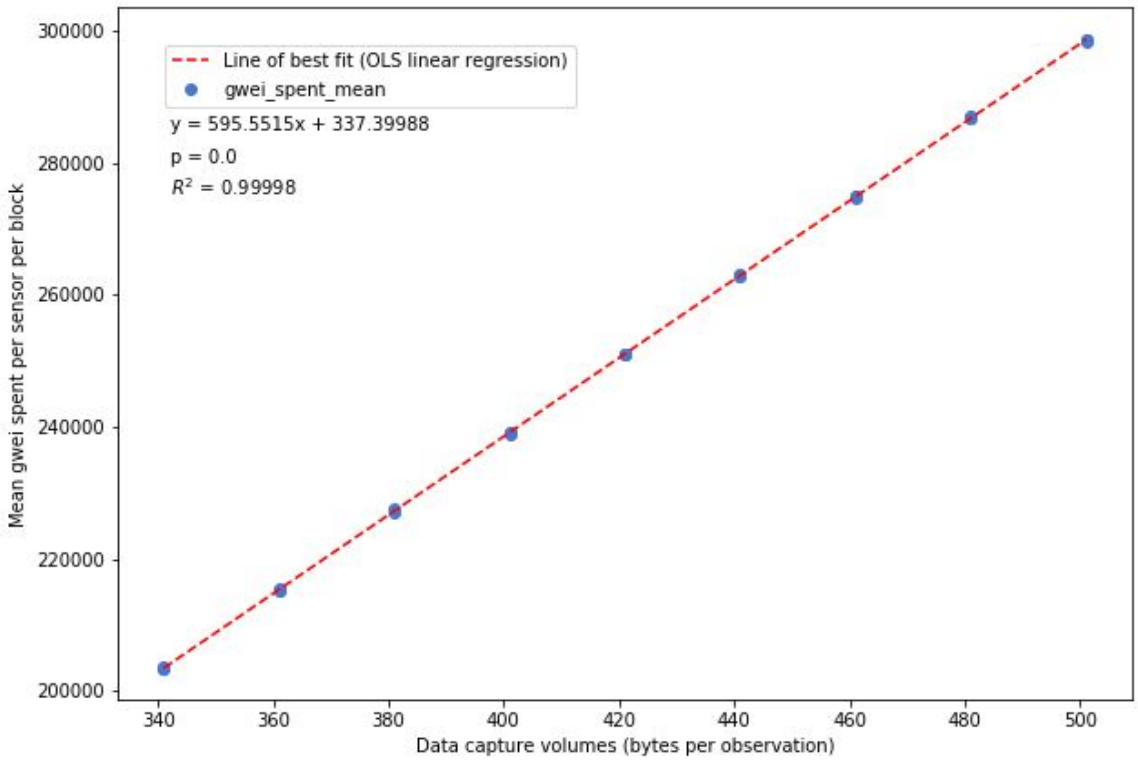
First, in model runs simulating edge recording volumes of 241 - 321 bytes, a very strong positive linear correlation is observed (r_{xy} = 0.99997). Figure 25 depicts this segment of the sample, including the OLS line of best fit - note the R^2 value of 0.99993, indicating that 99.99% of variance is explained by the independent variable.

Figure 25: Data capture volumes (241 <= bytes <= 321) against mean gwei spent per sensor



Between 321 and 341 bytes per record, average gwei expenditures drop dramatically and begin another, less steep upward trend, also a positive linear correlation (Pearson's correlation coefficient r_{xy} = 0.99999) which continues to the end of the parameter sweep. This segment of the sample is shown in Figure 26; an R^2 value of 0.99998 indicates that this positive linear relationship is very strong.

Figure 26: Data capture volumes (341 <= bytes <= 501) against mean gwei spent per sensor



While these results are interesting, it seems unlikely that these near-perfectly correlated relationships, as well as the mid-sweep discontinuity, are due to something other than an unintended aspect of model design. As such, these results are determined to be not useful to this investigation into the dynamics of a scaling blockchain network.

Recording Frequencies

Gwei spent

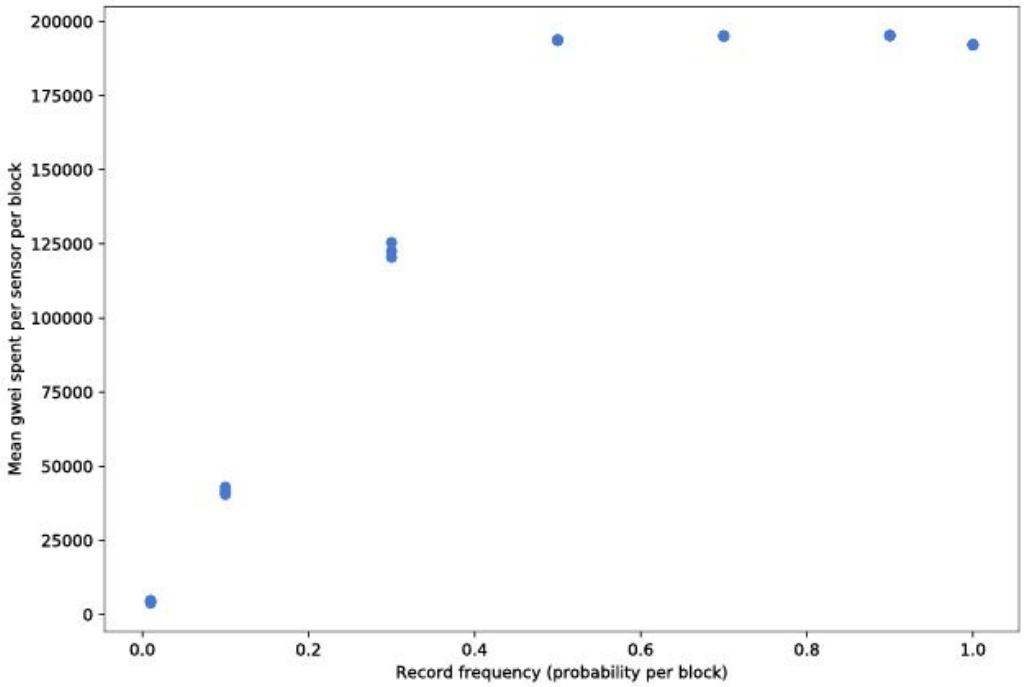
Figure 27 shows that mean gwei spent per sensor per block exhibited a positive linear correlation with record frequencies until a limit was reached, at an approximate record frequency of 0.5, after which the level remained roughly constant. This limit occurred at just below 200000 gwei per sensor per block. (Note the small difference between the median, third quartile and maximum values in Table 12). In a 40 sensor network, this is almost certainly caused by the block gas

limit: higher gwei expenditures would have exceeded the block gas limit, an impossibility according to the blockchain protocol. (Rather than exceed this limit, unvalidated transactions are simply left in the mempool to be validated in a future block.)

Table 12: Summary statistics, Mean gwei spent across record frequency parameter sweep

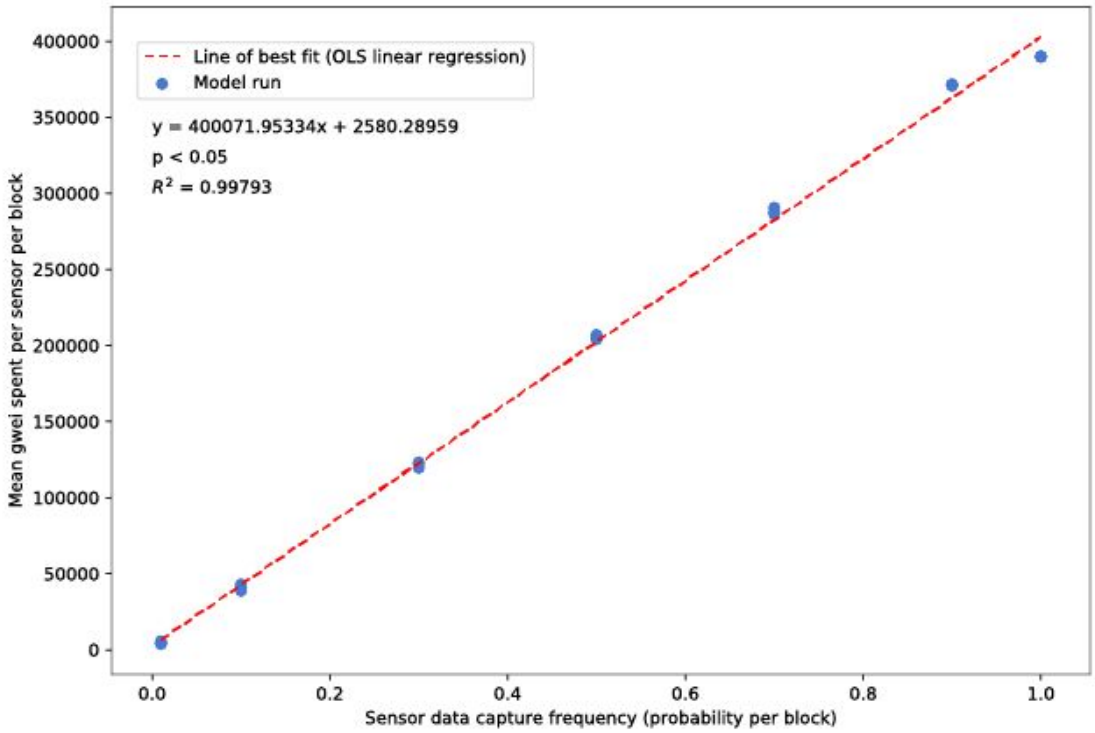
n		21
Mean gwei spent per sensor per block	Mean	135006.0
	σ	77359.1
	Minimum	3818.8
	Quartile 1	42969.8
	Median	192168.8
	Quartile 3	194968.8
	Maximum	195385.4

Figure 27: Record frequencies versus mean gwei spent per sensor per block



An OLS simple linear regression was performed to model the linear relationship between the variables at record frequency values less than 0.5; this plot, including the line of best fit, are depicted in Figure 28. An R^2 value of 0.99924 corroborates the visual assessment of a near-perfect linear relationship.

Figure 28: Record frequency (< 0.5) versus mean gwei spent per sensor per block



Appendix 3: Allen's 10 Principles of Self-Sovereign Identity

Reprinted verbatim from *The Path to Self-Sovereign Identity* (Allen 2016)

1. **Existence.** *Users must have an independent existence.* Any self-sovereign identity is ultimately based on the ineffable "I" that's at the heart of identity. It can never exist wholly in digital form. This must be the kernel of self that is upheld and supported. A self-sovereign identity simply makes public and accessible some limited aspects of the "I" that already exists.

2. **Control.** *Users must control their identities.* Subject to well-understood and secure algorithms that ensure the continued validity of an identity and its claims, the user is the ultimate authority on their identity. They should always be able to refer to it, update it, or even hide it. They must be able to choose celebrity or privacy as they prefer. This doesn't mean that a user controls all of the claims on their identity: other users may make claims about a user, but they should not be central to the identity itself.

3. **Access.** *Users must have access to their own data.* A user must always be able to easily retrieve all the claims and other data within his identity. There must be no hidden data and no gatekeepers. This does not mean that a user can necessarily modify all the claims associated with his identity, but it does mean they should be aware of them. It also does not mean that users have equal access to others' data, only to their own.

4. **Transparency**. *Systems and algorithms must be transparent.* The systems used to administer and operate a network of identities must be open, both in how they function and in how they are managed and updated. The algorithms should be free, open-source, well-known, and as independent as possible of any particular architecture; anyone should be able to examine how they work.

5. **Persistence.** *Identities must be long-lived.* Preferably, identities should last forever, or at least for as long as the user wishes. Though private keys might need to be rotated and data might need to be changed, the identity remains. In the fast-moving world of the Internet, this goal may not be entirely reasonable, so at the least identities should last until they've been outdated by newer identity systems. This must not contradict a "right to be forgotten"; a user should be able to dispose of an identity if he wishes and claims should be modified or removed as appropriate over time. To do this requires a firm separation between an identity and its claims: they can't be tied forever.

6. **Portability.** *Information and services about identity must be transportable.* Identities must not be held by a singular third-party entity, even if it's a trusted entity that is expected to work in the best interest of the user. The problem is that entities can disappear — and on the Internet, most

eventually do. Regimes may change, users may move to different jurisdictions. Transportable identities ensure that the user remains in control of his identity no matter what, and can also improve an identity's persistence over time.

7. **Interoperability.** *Identities should be as widely usable as possible.* Identities are of little value if they only work in limited niches. The goal of a 21st-century digital identity system is to make identity information widely available, crossing international boundaries to create global identities, without losing user control. Thanks to persistence and autonomy these widely available identities can then become continually available.

8. **Consent.** *Users must agree to the use of their identity.* Any identity system is built around sharing that identity and its claims, and an interoperable system increases the amount of sharing that occurs. However, sharing of data must only occur with the consent of the user. Though other users such as an employer, a credit bureau, or a friend might present claims, the user must still offer consent for them to become valid. Note that this consent might not be interactive, but it must still be deliberate and well-understood.

9. **Minimalization.** *Disclosure of claims must be minimized.* When data is disclosed, that disclosure should involve the minimum amount of data necessary to accomplish the task at hand. For example, if only a minimum age is called for, then the exact age should not be disclosed, and if only an age is requested, then the more precise date of birth should not be disclosed. This principle can be supported with selective disclosure, range proofs, and other zero-knowledge techniques, but non-correlatibility is still a very hard (perhaps impossible) task; the best we can do is to use minimalization to support privacy as best as possible.

10. **Protection.** *The rights of users must be protected.* When there is a conflict between the needs of the identity network and the rights of individual users, then the network should err on the side of preserving the freedoms and rights of the individuals over the needs of the network. To ensure this, identity authentication must occur through independent algorithms that are censorship-resistant and force-resilient and that are run in a decentralized manner.

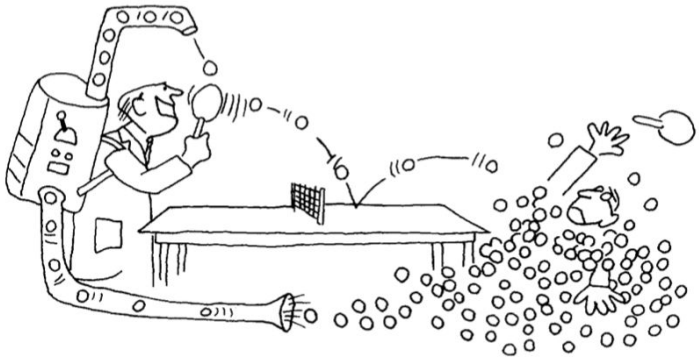
Appendix 4: The Three Tragedies of the Informational Commons

1. The failure to share information with someone who would benefit from it.

2. The abuse of a public channel's openness, rendering it unusable for others.

   Distributed Denial of Service attacks are an example of this (Beal 2019).

   Figure 29: The misuse of a public channel



. . . filibustering destroys communication.

Reprinted from Licklider (1968 pp 35)

3. Not admitting to a security breach from embarrassment or due to reputational cost.

   "Many firms treat hacks like gonorrhoea, an embarrassing affliction no one wants to admit even if speaking about it would stop its spread. Some call it a tragedy of the cyber-commons." (The Economist 2019 Schumpeter).

Appendix 5: History of Computing

A5.1: Analytical machines

Building on the work of Gottfried Wilhelm Leibniz's work on a mechanical calculator of "the four fundamental operations of arithmetic"[168] (Martin 1992 pp 39), Charles Babbage first proposed an automatic computer in 1822, later outlining an even more sophisticated analytical engine, a "general purpose computer" (CrashCourse 2017). Based on Babbage's proposed mechanical computers (Babbage 1822), Ada Lovelace foresaw the potential of computing[169], going so far as to develop software programs for Babbage's still-unbuilt machine (Lovelace 1843, Fuegi 2003, Essinger 2018). Technology was moving toward the reliable mechanization of logical operations.

In the early 20th century, mathematicians, researchers and engineers worked to incorporate the use of digital electronics in computing systems, based on advancements of understandings in the properties of electricity and conductive materials, as well as Akira Nakashima's early work on switching theory, based on two-valued Boolean algebra (Wynn-Williams 1931, Stankovic 2008). These technologies built toward Alan Turing's seminal work.

A5.2: Transmitting information prior to Shannon

As innovations in computing machines were progressing, so were techniques for transmitting information over distances. Signals had been sent on optical and auditory channels for millennia (Gleick 2011); by the early 1800s the semaphore telegraph[170] was in wide use (Burns 2004). The encoding of symbols into electrical

[168] Addition, subtraction, multiplication and division.
[169] It could be argued that Lovelace foresaw such cutting-edge computational techniques as the application of generative adversarial networks to music composition (Engel 2019): "Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent." -Ada Lovelace (Toole 1998, pp 694)
[170] By which messages are transmitted via line-of-sight relay stations on a visual channel.

currents had been discussed for some time (Fahie 1884) prior to “the first working electrostatic telegraph” being built in 1816. Such a system conferred substantial benefits over the optical telegraphs in use at the time[171] (Norman 2019).

These telegraphy systems shared the attribute of providing a communication channel upon which messages could travel rapidly and accurately. They necessarily relied on the establishment of some system of encoding information on the channel - without the ability to interpret the symbols encoded, the data would be meaningless to the informee. Enter Claude Shannon.

A5.3: Early innovation in internetworking

The vision for networked computers was put forward with striking clarity in the late 1950s and early 1960s by computer scientist J.C.R. Licklider (Living Internet 2019): a “galactic network” of interconnected computers (Leiner 1997). Licklider and his contemporaries investigated the concept, including the development of a theory of packet switching, which enabled a channel to be used by multiple traffic sources (Kleinrock 1961). These developments led to the first wide-area network connection being established over a telephone line connected California and Massachusetts in 1965 by Merrill and Roberts (Leiner 1997).

In 1966, the Defense Advanced Research Projects Agency began funding the development of ARPANET, which accelerated the refinement of “the overall structure and specifications” for such a network of computers (Leiner 1997). By the early 1970s, the technical capability for networked computing was established.

However, many of the networks relied on protocols implemented within organizations: because “these protocols have addressed only the problem of

[171] According to inventor Francis Ronalds, the system offered "a mode of conveying telegraphic intelligence with great rapidity, accuracy, and certainty, in all states of the atmosphere, either at night or in the day, and at small expense." Privacy was also greater, as the optical channels semaphores used were public; accessing the electrical signal carried on an electrical conduit would require additional specialized equipment and knowledge. Expense was reduced not least due to the reduction in the required line-of-sight infrastructure and personnel required to relay an optical message.

communication on the same network" (Cerf 1974 pp 1), internetworking remained difficult. In *A Protocol for Packet Network Intercommunication*, Cerf and Kahn (1974) established the Transport Communication Protocol, which "provides reliable, ordered, and error-checked delivery of a stream of octets (bytes) between applications running on hosts communicating via an [Internet Protocol] network" (Wikipedia 2019d). It could be argued that this marked the birth of the Internet, as this protocol is crucial to its functioning.

Appendix 6: Money as a conceptual object

A relevant example of a conceptual object: a monetary unit. It seems that our understanding of money is that it adheres to the physical laws - like physical objects, that there can only exist one of any one instance of a class (i.e., dollar). These units are scarce; it is part of what makes money money. This derives from our worldview regarding the nature of physical objects, and from the origins of money, when units were physical.

However, multiple instances of the same unit of a currency can exist, because they are conceptual objects, not physical ones. This is the basis, in this author's highly limited understanding, of fractional reserve banking: banks lend out a second version of the dollars they are given by depositors. It appears that this is enabled by the fungibility of money (Hodgson 2012) - the "sameness" of its individual units (which is not a physical property, but a conceptual one, see Appendix 7) - depositors do not care if they receive a dollar different to the one they deposited when they go to withdraw funds. If they did - if money were non-fungible[172] - banks could not promise that depositors could withdraw whenever they want and only keep a fraction of deposits on hand (Investopedia 2019).

It seems that this violation of the conceptual integrity of monetary units[173] does not occur in the context of cryptocurrencies such as bitcoin and ether. Though they exist almost entirely within the informational space, there can only be one of each instance[174], represented by the private key required to transfer the unit to another. This is the first time money has had this attribute of conceptual integrity since at least the abandonment of the gold standard, before which monetary units represented physical objects: pieces of gold of a standard weight - though gold is considered to be fungible; bitcoin may be the first conceptual object with complete integrity ever created. Based on this, Bitcoin should not be thought of as a bank where money is stored; it *is* the money ...

[172] Also meaning heterogeneous - as all physical objects are (Appendix 7).
[173] Violating integrity because they are intended to be unreplicable.
[174] Based on the existence of a single authoritative ledger representing ownership, the blockchain.

While this author is skeptical of the sustainability of the current configuration of fractional reserve banking, its implementation in reality should not be confused with it in the ideal. It clearly offers some value, or it wouldn't have been adopted.

Here the potential of the locked coins comes to mind. Of the 17,899,562 BTC in circulation at the time of writing (coinmarketcap 12:27:00 UTC+1, 27 August 2019, Block 591965) - coins that exist, that were mined, are stored at a valid wallet address [175] - it is estimated that 36% of the Bitcoin minted is held in wallets where the private keys represented have been lost by their custodians (Kelso 2018). Based on the current bitcoin price of approximately $10,150, this represents approximately $65,400,000,000 in monetary value that has been captured then lost, as without these private keys to sign a transaction transferring (i.e. spending) these coins, according to the Bitcoin protocol, they can never be used. Furthermore, many account holders store their bitcoin for long periods, rarely accessing the funds.

Could that value, captured in the blockchain, somehow be put to use? As long as holders' coins are always, of course, available for the key holder to spend, as that is a requirement of the system[176] - could that value, lying fallow in the system, generate more value for users, for humanity? It seems clear that this would need to exist at the protocol level, but could leverage a critical difference between physical and conceptual objects, especially the ability for conceptual objects to be in two places at once.

[175] Which means private keys exist in the keyspace that could sign a transaction specifying their transfer (Quora 2014).
[176] Of course the problem is what if suddenly all "lost" keys submitted transactions at once?

Appendix 7: Initial observations[177] on conceptual reality

Observations on conceptual reality

The unmanifest conceptual space is ever present, global[178] and infinite.

The meaning contained within the universe is infinite at every point through time and space[179].

Without sentience conceptual reality does not manifest. Within a sentient entity's private awareness, it does - to the depth that the awareness perceives[180].

Manifestations within conceptual reality are caused by the perception of some data. They do not manifest without this physical cause[181].

Every object has a subject and every subject has an object[182]. Without either the entity would not exist[183].

[177] These are intuitions and beliefs, stated because they provide a paradigm that resolves some previously unresolved issues in this author's worldview. Inspired in no small part by Wilber (2001).
[178] i.e. universal.
[179] Put another way, the informational potential of the unmanifest conceptual space is infinite.
[180] In this author's definition, digital computers are sentient and possess an awareness.
[181] So it appears ... See Footnote 193 for treatment of conception vs perception.
[182] Consider: my body is an object - matter and energy in space. My subject is the meaning about that object - my name, my experiences, my ideals, my aspirations and so on - my self. To complicate matters, it appears that subjects may also be objects in their own right, with subjects of their own ...
[183] This may be because object classes only exist in the conceptual space; sameness only exists within the conceptual space. This claim derives from the recognition that physical reality is entirely heterogeneous: every point in the universe, at every instant, is perfectly unique. Only in interpretation is comparison possible and similarity ascribed. Since discernment of the boundaries required for the identification of discrete objects is necessarily subjective, it must take place within the perceiving entity's subjective awareness (i.e. conceptual space).

Data is tangible and manifests in physical space; information is intangible and manifests in conceptual space.

Data[184] can be stored in matter or transmitted as energy[185].

Data carries informational potential[186] about the qualities of its source when it was emitted, reflected or refracted.

[184] Note: here, "data" refers to analogue data. Analogue data exists on a continuous spectrum; digital information exists on a discrete spectrum. It is only within awareness that the boundaries exist necessary for discrete entities to arise; thus, digital information only exists within an informational entity's awareness. By this reasoning, it is possible that "digital data" is a misnomer. This digital information is, as described in the Literature Review, encoded onto an analogue signal. Investigation of the differences between analogue and digital holds promise for explaining the nature of conceptual reality. Taking light as an example, a key distinction between analogue and digital data is that, in analogue, meaning is typically conveyed by the two dimensional spatial configuration of individual photons with varying qualities changing over time. Our retinas detect these relative variations - dark and light, or colors - and images are discerned. This is what was meant by "parallel channel", Footnote 16. Digital information can be carried on a one-dimensional analogue channel: a stream of single photons, as referenced in Footnote 151.

[185] A clear distinction is necessary between static data - matter, often intentionally inscribed with symbols - and dynamic data, traveling through space borne on a carrier wave. To come into contact with a sensory organ it seems it must be transmitted on such a wave. The ink markings on a book page serve as an example of static data. To be read (perceived), they must be carried to the reader's eye by light - dynamic data. It appears that dynamic data is very closely related to - or perhaps just is - energy. It is emitted or reflected by a source and travels through some "'ether' … the hypothetical invisible medium that permeates the universe" (Buterin 2014b). This "ether" includes whatever it is light travels through; air, as the auditory channel carries data; as well as conductive materials, which can carry data on an electrical signal. Data can also be conveyed through two objects of mass physically contacting each other - Footnote 188.

[186] Contingent upon sensing, perception and interpretation within the awareness of an informee.

1. Data must come into contact with the sensory organ[187] of a perceiving entity in order to be sensed[188].

2. If sensed, by focusing its attention[189] on the incoming data the informational potential carried can become information and interpreted by the informee as meaning[190]. This is qualia[191].

3. Within the conceptual space in the informational entity's awareness[192], if its attention is placed upon the signal carrying the data into its organism[193], perception can occur[194].

4. As data is perceived, information[195] is created, and can be discerned, interpreted and acted upon.

[187] In biological organisms: eye, ear, tongue, skin, cell membrane; in synthetic informational organisms, antenna or conductive connection with the signal origin, or empirical sensor.
[188] A sleeping person is still an informational entity: it can detect a signal - say, a stick poking them or a loud noise - and awaken.
[189] Probably a key to understanding all of this ...
[190] It appears that information only exists within the awareness of a perceiving entity. Physical data holds the *potential* for meaning (information), but it is only upon sensing and perception that this information comes into existence. This trait - that information only exists within the manifest conceptual space (i.e. the awareness) of informational entities - means that it is a necessarily subjective phenomenon.
[191] Generally, qualia is the process of data entering a sentient entity's awareness and becoming information - the objective-subjective interface. Because the physical universe is perfectly unique, each instant of qualia is unique (and infinite), a moment in your lived experience where the universe pours into you. It is by conceptualizing these qualia that we discern information and ascribe meaning to our perceived reality.
[192] Which is local, private and manifest.
[193] Here, focused on perception of external events. It seems likely that the conception of internal events must arise from the detection of internal data entering awareness.
[194] Section 3.2.1 refers to the process of machine qualia in which analogue data is perceived as binary information through the machine's empirical sensor.
[195] A type of conceptual object?

References

@santisiri. (2019). *santi.* Twitter [online]. Available at https://twitter.com/santisiri/status/1163779311221923840 [Accessed 21 August 2019].

@santisiri. (2019b). *santi.* Twitter [online]. Available at https://twitter.com/santisiri/status/1155139202318602240 [Accessed 27 August 2019].

@Snowden. (2019). *Edward Snowden*. Twitter. [online] Available at https://twitter.com/Snowden/status/1165096076799619073

@VitalikButerin. (2018). *Vitalik Non-giver of Ether on Twitter*. Twitter [online]. Available at https://twitter.com/VitalikButerin/status/1051160932699770882 [Accessed 19 August 2019].

Allen, C. (2016). *The Path to Self-Sovereign Identity*. Life with Alacrity [online]. Available at http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html [Accessed 10 May 2019].

Allen, C. et al. (2015). *Decentralized Public Key Infrastructure*. Github [online]. Available at https://github.com/WebOfTrustInfo/rwot1-sf/blob/master/draft-documents/Decentralized-Public-Key-Infrastructure-CURRENT.md [Accessed 7 May 2019].

Anderson, Ross. (2003). *'Trusted Computing' Frequently Asked Questions*. [online]. Available at https://www.cl.cam.ac.uk/~rja14/tcpa-faq.html [Accessed 24 August 2019].

Antonopoulos, A. (2017). *Mastering Bitcoin*. Github [online]. Available at https://github.com/bitcoinbook/bitcoinbook [Accessed 12 August 2019].

Antonopoulos, A. (2017b). *Money as a System-of-Control*. Available at https://www.youtube.com/watch?v=FyK4P7ZdOK8&vl=en [Accessed 20 April 2019].

Antonopoulos, A. (2018). *Mastering Ethereum.* Github [online]. Available at https://github.com/ethereumbook/ethereumbook [Accessed 10 October 2019].

Antonopoulos, A. (2019). *Coinscrum presents :: Andreas M Antonopoulos :: 13/06/19*. Youtube [online]. Available at https://www.youtube.com/watch?v=LlvuKoXRrsU [Accessed 6 June 2019].

Aragon. (2019). *Aragon Network*. Github [online]. Available at https://github.com/aragon/whitepaper [Accessed 20 June 2019].

Babbage, C. (1822). *A Note Respecting the Application of Machinery to the Calculation of Astronomical Tables*. [online]. Available at http://cyn.io/charles-babbage-a-note-respecting-the-application-of-machinery-to-the-calculation-of-astronomical-tables/ [Accessed 18 August 2019].

BBC. (2017). *CryptoKitties craze slows down transactions on Ethereum.* BBC News [online]. Available at https://www.bbc.com/news/technology-42237162 [Accessed 22 August 2019].

Beal, V. (2019). *DDoS attack - Distributed Denial of Service*. [online]. Available at https://www.webopedia.com/TERM/D/DDoS_attack.html [Accessed 25 August 2019].

Becha, H. (2019). *Smart Containers: Real-time Smart Container data for supply chain excellence*. UNECE - UN / CEFACT [online]. Available at https://www.unece.org/fileadmin/DAM/cefact/GuidanceMaterials/WhitePapers/WP-SmartContainers_Eng.pdf [Accessed 20 June 2019].

Benet, Juan. (2014). *IPFS - Content-Addressed, Versioned, P2P File System*. Github [online]. Available at https://github.com/ipfs/papers/raw/master/ipfs-cap2pfs/ipfs-p2p-file-system.pdf [Accessed 21 March 2019].

Berners-Lee, T. (1990). *WorldWideWeb: Proposal for a HyperText Project*. W3.org [online]. Available at https://www.w3.org/Proposal.html [Accessed 21 August 2019].

Bevir, M. (2012). *Governance: A Very Short Introduction*. Oxford: Oxford University Press.

Bilbow, Angela. (2019). *High Court tackles Bitcoin 'property' first*. Available at https://www.cdr-news.com/categories/litigation/10003-high-court-tackles-bitcoin-property-first [Accessed 24 August 2019].

Bitvalley. (2018). *IBISA: Disrupting Agriculture Insurance*. Bitvalley [online]. Available at https://www.ibisa.network/ [Accessed 13 March 2019].

Boyd-Rice, J. (2018). *New A.I. application can write its own code*. Futurity [online]. Available at https://www.futurity.org/artificial-intelligence-bayou-coding-1740702/ [Accessed 23 August 2019].

Braendgaard, P. (2016). *Simple Convention for Human Readable Terms for Smart Contracts*. Stake Ventures [online]. Available at https://blog.stakeventures.com/articles/smart-contract-terms [Accessed 26 August 2019].

Burns, R. (2004). *Communications: An International History of the Formative Years.* The Institution of Electrical Engineers [online]. Available at https://books.google.co.uk/books?id=7eUUy8-VvwoC&pg=PA29&redir_esc=y#v=onepage&q&f=false [Accessed on 14 August 2019].

Buterin, V., Weyl, G. (2018). *Liberation Through Radical Decentralization*. Medium [online]. Available at https://medium.com/@VitalikButerin/liberation-through-radical-decentralization-22fc4bedc2ac [Accessed 26 August 2019].

Buterin, V. (2013). *A Next-Generation Smart Contract and Decentralized Application Platform*. Ethereum [online]. Available at https://github.com/ethereum/wiki/wiki/White-Paper [Accessed 10 October 2018].

Buterin, V. (2014b). *Ethereum Community Forum*. [online]. Available at https://forum.ethereum.org/discussion/comment/3389/#Comment_3389 [Accessed 22 August 2019].

Buterin, V. (2016). *A Proof of Stake Design Philosophy*. Medium. [Accessed at https://medium.com/@VitalikButerin/a-proof-of-stake-design-philosophy-506585978d51]

Buterin, V. (2017). *The Meaning of Decentralization*. Medium [online]. Available at https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274 [Accessed 14 August 2019].

Carney, M. (2019). *The Growing Challenges for Monetary Policy in the current International Monetary and Financial System*. Bank of England [online]. Available at https://www.bankofengland.co.uk/speech/2019/mark-carney-speech-at-jackson-hole-economic-symposium-wyoming [Accessed 27 August 2019].

Cerf, V.G., Kahn, R.E. (1974). *A Protocol for Packet Network Intercommunication 13.* IEEE Trans on Comms, Vol Com-22 No 5 [online]. Available at https://www.cs.princeton.edu/courses/archive/fall06/cos561/papers/cerf74.pdf. [Accessed 20 August 2019].

Chen, J. (2019). *Agency Problem.* Investopedia [online]. Available at https://www.investopedia.com/terms/a/agencyproblem.asp [Accessed 21 August 2019].

Chen, R., Ramsundar, B., Robbins, R. (2019). *Fair value and decentralized governance of data*. Computable. [online] Available at https://github.com/computablelabs/computable/blob/master/whitepaper/computable_whitepaper.pdf [Accessed 26 August 2019].

Choudhury, N. (2014). World Wide Web and its Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies*, Vol. 5 (6) , 2014, 8096-8100.

Clack, C. D., Bakshi, V., Braine, L. (2016). Smart Contract Templates: foundations, design landscape and research directions. *Barclays Bank PLC* [online]. Available at https://arxiv.org/pdf/1608.00771.pdf [Accessed 26 August 2019].

Cohen, B. (2003). *Incentives build robustness in bittorrent*. In Workshop on Economics of Peer-to-Peer systems, volume 6, pages 68–72.

Corda. (2019). *Corda.* [online]. Available at https://www.corda.net/ [Accessed 10 July 2019].

Core Mesa Team. (2019). *Mesa: Agent-based modeling in Python 3*+. Available at https://github.com/projectmesa/mesa [Accessed 10 March 2019].

CrashCourse. (2017). *Early Computing: Crash Course Computer Science #1*. Youtube [online]. Available at https://www.youtube.com/watch?v=O5nskjZ_GoI [Accessed 10 July 2019].

CrashCourse. (2017b). Electronic Computing: Crash Course Computer Science #2. Youtube [online]. Available at https://www.youtube.com/watch?v=LN0ucKNX0hc [Accessed 10 July 2019].

Daemen, J., n.d. (1999). *The Rijndael Block Cipher 47.* Available at csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf#page=1 [Accessed 19 August 2019].

DAOstack. (2019). *An Operating System for Collective Intelligence*. DAOstack [online]. Available at https://daostack.io/wp/DAOstack-White-Paper-en.pdf [Accessed 20 February 2019].

DiNucci, D. (1999). *Fragmented Future*. Print [online]. Available at https://web.archive.org/web/20111110143942/http://darcyd.com/fragmented_future.pdf [Accessed 19 August 2019].

Dume, B. (2012). *Photon shape could be used to encode quantum information*. Ultrafast Science [online]. Available at https://physicsworld.com/a/photon-shape-could-be-used-to-encode-quantum-information/ [Accessed 17 August 2019].

Ehrsam, F. (2017). *Blockchain Governance: Programming Out Future*. Medium [online]. Available at https://medium.com/@FEhrsam/blockchain-governance-programming-our-future-c3bfe30f2d74 [Accessed 10 December 2017].

Engel, J. (2019). *GANSynth: Making music with GANs*. magenta [online]. Available at https://magenta.tensorflow.org/gansynth [Accessed 10 April 2019].

Essinger, J. (2018). *Ada Lovelace: a visionary of computing*. History Extra [online]. Available at https://www.historyextra.com/period/modern/ada-lovelace-stem-women-computing-science-facts-life/ [Accessed 15 August 2019].

Ethereum. (2019b). *On sharding blockchains.* Github [online]. Available at https://github.com/ethereum/wiki/wiki/Sharding-FAQ [Accessed 26 August 2019].

Ethereum Foundation. (2019). *Design Rationale*. Github. [Accessed at https://github.com/ethereum/wiki/wiki/Design-Rationale]

Etherscan.io. (2019). *Ethereum Block Time History*. Etherscan. [online] Available at https://etherscan.io/chart/blocktime [Accessed 20 August 2019].

Fahie, J. J. (1884). *A History of Electric Telegraphy to the Year 1837.* E. & F. N. Spon: London [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=2ahUKEwjn4JHy-qPkAhVkSxUIHUKzAYQQFjADegQIARAC&url=https%3A%2F%2Fwww.princeton.edu%2Fssp%2Fjoseph-henry-project%2Ftelegraph%2FA_history_of_electric_telegraphy_to_the.pdf&usg=AOvVaw20l3iQsyg_Avz0JrT9l2uv [Accessed 14 August 2019].

Fecke, M. (2018). *The Problem of Blockchain Oracles - Interview with Alexander Egberts*. Legal Tech Blog [online]. Available at

https://legal-tech-blog.de/the-problem-of-blockchain-oracles-interview-with-alexander-egberts [Accessed 21 August 2019].

Fernandez, M., Sanger, D. E., Martinez, M. T. (2019). *Ransomware Attacks Are Testing Resolve of Cities Across America*. The New York Times [online]. Available at https://www.nytimes.com/2019/08/22/us/ransomware-attacks-hacking.html [Accessed 23 August 2019].

Floridi, L. (2010). *Information: A Very Short Introduction*. Oxford: Oxford University Press.

Fuegi, J., Francis, J. (2003). *Lovelace & Babbage and the Creation of the 1843 'Notes'.* IEEE Annals of the History of Computing [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiPgu34_aPkAhWltXEKHeLDB6AQFjAAegQIABAC&url=https%3A%2F%2Fwww.scss.tcd.ie%2Fcoghlan%2Frepository%2FJ_Byrne%2FA_Lovelace%2FJ_Fuegi_%26_J_Francis_2003.pdf&usg=AOvVaw1H1JvrMxvmSNoGxeoMJ7-t [Accessed 15 August 2019].

Galperin, E., Hassine, W. B. (2015). *Changes to Facebook's "Real Names" Policy Still Don't Fix the Problem*. EFF [online]. Available at https://www.eff.org/deeplinks/2015/12/changes-facebooks-real-names-policy-still-dont-fix-problem [Accessed 23 August 2019].

Gentry, C. (2010). *Computing Arbitrary Functions of Encrypted Data*. [online]. Available at https://dl.acm.org/citation.cfm?id=1666444 [Accessed 21 August 2019].

Gilbert, N. (2008). *Agent-based models.* London: Sage Publications.

git-scm.com. (2019). *Git.* [online]. Available at: https://git-scm.com/. [Accessed 22 August 2019].

Glaser, A., Barak, B., Goldston, R. (2014). *A zero-knowledge protocol for nuclear warhead verification*. Nature [online]. Available at https://www.nature.com/articles/nature13457 [Accessed 24 August 2019].

Gleick , J. (2011). *The Information: A History, a Theory, a Flood*. Fourth Estate: London.

Grigg, I. (2004). The ricardian contract. In *Proceedings of the First IEEE International Workshop on Electronic Contracting*, pages 25–31. IEEE. http://iang.org/papers/ricardian_contract.html.

Gupta, V., Knight, R., Buchanan, A., Wray, C., Grigg, I., Kuhlman, C., Cimpoesu, M., Mainelli, M., Freedman, C. (2019). *Mattereum Working Paper*. Mattereum [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiZlObctp7kAhX8RxUIHcj7C0QQFjAAegQIARAC&url=https%3A%2F%2Fmattereum.com%2Fupload%2Fiblock%2Faf8%2Fmattereum_workingpaper.pdf&usg=AOvVaw1ZdsGE1x8Ks65p8Y_QEmXR [Accessed 2 February 2019].

Gutierrez, D. (2016). *Why Time-Value of Data Matters*. insideBIGDATA [online]. Available at https://insidebigdata.com/2016/04/08/why-time-value-of-data-matters/ [Accessed 26 August 2019].

gwei.io. (2019). *GWEI.IO*. [online]. Available at https://gwei.io/ [Accessed 18 August 2019].

Harris, M. (2019). *Pentagon testing mass surveillance balloons across the US*. The Guardian [online]. Available at https://www.theguardian.com/us-news/2019/aug/02/pentagon-balloons-surveillance-midwest [Accessed 23 August 2019].

Haselton, T. (2017). *Credit reporting firm Equifax says data breach could potentially affect 143 million US consumers.* CNBC [online]. Available at https://www.cnbc.com/2017/09/07/credit-reporting-firm-equifax-says-cybersecurity-incident-could-potentially-affect-143-million-us-consumers.html [Accessed 23 August 2019].

Heller, J. (1961). *Catch-22*. London: Vintage.

Hodgson, G. (2012). *A revisionist critique of fractional reserve banking*. Positive Money. [online] Available at https://positivemoney.org/2012/01/revisionist-critique-fractional-reserve-banking/ [Accessed 27 August 2019].

Hoopes, J. (2019). *Reimagining "global": Programmable incentivization and implications for personal governance*. Github [online]. Available at https://github.com/robisoniv/rwot9-prague/blob/master/topics-and-advance-readings/reimagining-global-rwot9.md [Accessed 19 August 2019].

Hopkins, N. (2017). *Deloitte hit by cyber-attack revealing clients' secret emails.* The Guardian [online]. Available at https://www.theguardian.com/business/2017/sep/25/deloitte-hit-by-cyber-attack-revealing-clients-secret-emails. [Accessed 23 August 2019].

Huang, A. (2017). *The Origins of Computing: The Sumerians,, The Abacus & Binary Code*. Medium [online]. Available at https://medium.com/@Ashley_Huang/the-origins-of-computing-the-sumerians-the-abacus-binary-code-15289c3ced16 [Accessed 1 August 2019].

Hunter, J. D. (2007). *Matplotlib: A 2D Graphics Environment.*, Computing in Science & Engineering, 9, 90-95, DOI:10.1109/MCSE.2007.55

Ibarra, I. A., Goff, L., Hernández, D. J., Lanier, J., Weyl, E. G. (2018). *Should We Treat Data as Labor? Moving Beyond "Free"*. Data as Labor Papers and Proceedings. [Available at https://ssrn.com/abstract=3093683]

Investopedia. (2019). *Fractional Reserve Banking*. Investopedia [online]. Available at https://www.investopedia.com/terms/f/fractionalreservebanking.asp [Accessed 24 August 2019].

jnnk. (2015). *What is Gas Limit in Ethereum?* Stackexchange.com [online]. Available at https://bitcoin.stackexchange.com/questions/39132/what-is-gas-limit-in-ethereum [Accessed 21 August 2019].

Kelso, C. E. (2018). *BTC: 36% in Circulation Lost, 23% Held by Speculators, US Tax Authority Monitoring*. Bitcoin.com. [online] Available at

https://news.bitcoin.com/btc-36-in-circulation-lost-23-held-by-speculators-us-tax-authority-monitoring/ [Accessed 27 August 2019].

Ker, Andrew D. (2014). *Computer Security*. Oxford: Oxford University Department of Computer Science, [online]. Available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjQ-ZH636DkAhXZTxUIHfVBD_EQFjAAegQIBBAC&url=http%3A%2F%2Fwww.cs.ox.ac.uk%2Fandrew.ker%2Fdocs%2Fcomputersecurity-lecture-notes-mt2014.pdf&usg=AOvVaw2reitUz7E6ktrlCORJ2vj5 [Accessed 20 October 2018].

Kerckhoffs, A. (1883). *La Cryptographie Militaire (première partie) 37*. Available at https://www.petitcolas.net/kerckhoffs/crypto_militaire_1_b.pdf; translation at https://www.petitcolas.net/kerckhoffs/index.html [Accessed 17 August 2019].

Kim, C. (2019). *Code For Ethereum's Proof-of-Stake Blockchain to Be Finalized Next Month*. Coindesk [online]. Available at https://www.coindesk.com/code-for-ethereums-proof-of-stake-blockchain-to-be-finalized-next-month [Accessed on 20 August 2019].

Kleinrock, Leonard. (1961). *Information Flow in Large Communication Nets*. Massachusetts Institute of Technology Research Laboratory of Electronics [online]. Available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiE3LaF4JbkAhWoTxUIHTGDDrwQFjAAegQIAhAC&url=https%3A%2F%2Fwww.lk.cs.ucla.edu%2Fdata%2Ffiles%2FKleinrock%2FInformation%2520Flow%2520in%2520Large%2520Communication%2520Nets.pdf&usg=AOvVaw11sqIM4l6SD1VAaJbM3ODH [Accessed 20 August 2019].

Krebs, Brian. (2019). *Facebook Stored Hundreds of Millions of User Passwords in Plain Text for Years.* Krebs on Security [online]. Available at https://krebsonsecurity.com/2019/03/facebook-stored-hundreds-of-millions-of-user-passwords-in-plain-text-for-years/ [Accessed 22 August 2019].

Lee, D. (2018). *Facebook security breach: Up to 50m accounts attacked*. BBC, [online]. Available at https://www.bbc.com/news/technology-45686890 [Accessed 26 August 2019].

Leiner, B., Cerf, V., Clark, D. D., Kahn, R. E., Kleinrock, L, Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. (1997). *Brief History of the Internet.* Internet Society [online]. Available at https://www.internetsociety.org/wp-content/uploads/2017/09/ISOC-History-of-the-Internet_1997.pdf [Accessed 20 August 2019].

Licklider, J. C. R. (1968). *The Computer as a Communication Device*. Science and Technology [online]. Available at http://memex.org/licklider.pdf [Accessed 20 August 2019].

Liu, C. (2008). *The Dark Forest*. Tor, A Tom Doherty Associates Book.

Living Internet. (2019). *J. C. R. Licklider And The Universal Network.* Living History [online]. Available at https://www.livinginternet.com/i/ii_licklider.htm [Accessed 14 August 2019].

Lloyd's. (2019). *Triggering innovation: How smart contracts bring policies to life*. Queen Mary Centre for Commercial Law Studies [online]. Available at https://www.lloyds.com/news-and-risk-insight/risk-reports/library/technology/triggering-innovation [Accessed 11 August 2019].

Long, C. (2019). *Bitcoin, The Dollar And Facebook's Cryptocurrency: Price Volatility Versus Systemic Volatility*. Forbes [online]. Available at https://www.forbes.com/sites/caitlinlong/2019/06/29/bitcoin-the-dollar-and-facebooks-cryptocurrency-price-volatility-versus-systemic-volatility/#1aa3f14d88b8 [Accessed 10 July 2019].

Lovelace, A. (1843). *Notes on L. Menabrea's 'Sketch of the Analytical Engine Invented by Charles Babbage, Esq.'*. Taylor's Scientific Memoirs, *3*(1843), p.1843.

Malan, D. (2014). *Lecture 0: Introduction to Computer Science I*. Youtube [online]. Available at: https://www.youtube.com/watch?v=z-OxzIC6pic [Accessed 13 July 2016].

Malik, N. (2008). *The Tradeoff Between Security And Privacy: How do Terrorists Use Encryption?* Forbes [online]. Available at https://www.forbes.com/sites/nikitamalik/2018/11/07/the-tradeoff-between-security-and-privacy-how-do-terrorists-use-encryption/#368d68862d8c [Accessed 17 August 2019].

Martin, Ernst. (1992). *The Calculating Machines: Their History and Development*. The MIT Press [online]. Available at www.rechenmaschinen-illustrated.com/Martins_book/Ernst%20Martin%20-%20Rechen%20Machinen%20OCR%204.pdf [Accessed 17 August 2019].

McCoy, M. (2015). *6 Notorious Cases of Data Loss All Hosting Providers Can Learn From*. R1 Blog [online]. Available at https://www.r1soft.com/blog/6-notorious-cases-of-data-loss-all-hosting-providers-can-learn-from [Accessed 23 August 2019].

McDonald, J. (2017). *Releasing Stuck Ethereum Transactions*. Medium [online]. Available at https://medium.com/@jgm.orinoco/releasing-stuck-ethereum-transactions-1390149f297d [Accesseed 25 August 2019].

McKinney, Wes. (2010). *Data Structures for Statistical Computing in Python.* Proceedings of the 9th Python in Science Conference, 51-56.

Merkle, R. (2016). *DAOs, Democracy and Governance*. [online]. Available at merkle.com/papers/DAOdemocracyDraft.pdf [Accessed 28 August 2019].

Merritt, R. (2003). *New group aims to secure PCs, PDAs, cell phones*. [online]. Available at https://www.eetimes.com/document.asp?doc_id=1202119 [Accessed 20 August 2019].

Moriya, H. (2018). *How to get Ethereum Block Gas Limit*. Medium. [online] Available at https://medium.com/@piyopiyo/how-to-get-ethereum-block-gas-limit-eba2c8f32ce [Accessed 10 July 2019].

Mullins, R. (2012). *What is a Turing machine?* University of Cambridge Department of Computer Science and Technology [online]. Available at https://www.cl.cam.ac.uk/projects/raspberrypi/tutorials/turing-machine/one.html [Accessed 12 August 2019].

Nair, B. Somanathan. (2002). *Digital Electronics and Logic Design*. Prentice Hall [online]. Available at https://books.google.co.uk/books?id=WK45wLHL-ycC&pg=PA289&dq=%22digital+signals+are%22+%22digital+electronics%22&hl=sv&sa=X&redir_esc=y#v=onepage&q=%22digital%20signals%20are%22%20%22digital%20electronics%22&f=false [Accessed 16 August 2019].

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Bitcoin.org [online]. Available at https://bitcoin.org/bitcoin.pdf [Accessed 23 October 2018].

Nakamoto, S. (2009). *Bitcoin v0.1 released.* The Cryptography Mailing List [online]. Available at https://www.metzdowd.com/pipermail/cryptography/2009-January/014994.html. [Accessed 18 August 2019].

National Institute of Standards and Technology. (2001). *FIPS 197, Advanced Encryption Standard (AES)*, *51*. Available at nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf [Accessed 19 August 2019].

Node.js Foundation. (2018). *Node.js*. [online] Available at: https://nodejs.org/en/. [Accessed 28 January 2018].

Norman, J. (2019). *Francis Ronalds Builds the First Working Electric Telegraph*. Jeremy Norman's HistoryofInformation.com [online]. Available at http://historyofinformation.com/detail.php?entryid=519 [Accessed 15 August 2019].

Nuñez, D. (2018). *Umbral: A Threshold Proxy Re-encryption Scheme*. NICS Lab, University of Malaga, Spain [online]. Available at https://github.com/nucypher/umbral-doc/blob/master/umbral-doc.pdf [Accessed 1 March 2019].

Oliphant, T. E. (2006). *A guide to NumPy.* USA: Trelgol Publishing. Available at https://www.numpy.org/

Orcutt, M. (2017). *A Mind-Bending Cryptographic Trick Promises to Take Blockchains Mainstream.* MIT Technology Review [online]. Available at https://www.technologyreview.com/s/609448/a-mind-bending-cryptographic-trick-promises-to-take-blockchains-mainstream/ [Accessed 26 August 2019].

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. ISBN 978-0-521-40599-7.

Oxford. (2019a). *Cipher*. Lexico [online]. Available at https://www.lexico.com/en/definition/cipher [Accessed 20 August 2019].

Oxford. (2019b). *Algorithm*. Lexico [online]. Available at https://www.lexico.com/en/definition/algorithm [Accessed 20 August 2019].

Paul, M. S. (2018). *Hyperledger — Chapter 3| When to use the Blockchain Technology*. The Startup [online]. Available at https://medium.com/swlh/hyperledger-chapter-3-when-to-use-the-blockchain-technology-a5c414221bdf [Accessed 27 August 2019].

Peaster, W. (2018). *FCoin GPM Listing Competition Acutely Clogs Ethereum*. Bitsonline [online]. Available at https://bitsonline.com/fcoin-gpm-listing-cloggage/ [Accessed 23 August 2019].

Pérez, Fernando, Granger, Brian E. (2007). *IPython: A System for Interactive Scientific Computing*. Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, doi:10.1109/MCSE.2007.53. Available at https://ipython.org.

phant0m. (2013). *What does the ^ (XOR) operator do?* Stackoverflow.com [online]. Available at https://stackoverflow.com/questions/14526584/what-does-the-xor-operator-do [Accessed 19 August 2019].

Prince, M. (2019). *Terminating Service for 8chan*. Cloudflare [online]. Available at https://blog.cloudflare.com/terminating-service-for-8chan/ [Accessed 23 August 2019].

Protocol Labs. (2019b). Filecoin FAQs. [online] Available at https://filecoin.io/faqs/ [Accessed 26 August 2019].

Protocol Labs. (2019). *Content Identifiers (CIDs).* Accessed at https://docs.ipfs.io/guides/concepts/cid/]

Quinn, J. (2008). *Greenspan admits mistakes in 'once in a century credit tsunami'*. The Telegraph [online]. Available at https://www.telegraph.co.uk/finance/financialcrisis/3248774/Greenspan-admits-mistakes-in-once-in-a-century-credit-tsunami.html [Accessed 20 August 2019].

Quora. (2014). *What would happen if I tried to send BTC to a nonexistent bitcoin address*. Quora. [online] Available at https://www.quora.com/What-would-happen-if-I-tried-to-send-BTC-to-a-nonexistent-bitcoin-address [Accessed 27 August 2019].

Quorum. (2019). *Quorum.* [online]. Available at https://www.goquorum.com/ [Accessed 26 August 2019].

Raymond, E. (1999). *Release Early, Release Often*. [online]. Available at http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/ar01s04.html [Accessed 14 March 2017].

Rea, A., Fischer, A., du Rose, J. (2019). *Colony Technical White Paper*. Colony [online]. Available at https://colony.io/whitepaper.pdf [Accessed 20 August 2019].

Rouse, M. (2019). *Access control.* [online]. Available at https://searchsecurity.techtarget.com/definition/access-control [Accessed 18 August 2019].

Ryan, D. (2017). *Calculating Costs in Ethereum Contracts*. Hackernoon [online]. Available at https://hackernoon.com/ether-purchase-power-df40a38c5a2f [Accessed 3 July 2019].

Sabanal, P. (2016). *Thingbots: The Future of Botnets in the Internet of Things*. SecurityIntelligence [online]. Available at https://securityintelligence.com/thingbots-the-future-of-botnets-in-the-internet-of-things/ [Accessed on 21 August 2019].

Safford, D. (2003). *Take Control of TCPA*. Linux Journal [online]. Available at https://www.linuxjournal.com/article/6633 [Accessed 23 August 2019].

Seabold, S., Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python*. Proceedings of the 9th Python in Science Conference.

Shannon, C. (1949). *Communication Theory of Secrecy Systems*. Bell System Technical Journal [online]. Available at https://archive.org/stream/bstj28-4-656#page/n5/mode/2up [Accessed 21 August 2019].

Shannon, C. E., (1938). *A symbolic analysis of relay and switching circuits*. Trans. Am. Inst. Electr. Eng. 57, 713–723. https://doi.org/10.1109/T-AIEE.1938.5057767

Shannon, C. E. (1948). *A Mathematical Theory of Communication*. The Bell System Technical Journal Vol XXVII No. 3 pp 379-423.

Shirer, M, MacGillivray, C. (2019). *The Growth in Connected IoT Devices Is Expected to Generate 79.4ZB of Data in 2025, According to a New IDC Forecast*. [Accessed at https://www.idc.com/getdoc.jsp?containerId=prUS45213219]

Simonsen, S. (2018). *Quantum Communication Just Took a Great Leap Forward*. Singularity Hub [online]. Available at https://singularityhub.com/2018/12/26/quantum-communication-just-took-a-great-leap-forward/ [Accessed 27 August 2019].

Stackoverflow. (2016). *Is there a limit for sha256 input?* Stackoverflow.com [online]. Available at https://stackoverflow.com/questions/17388177/is-there-a-limit-for-sha256-input [Accessed 27 August 2019].

Stankovic, R., Astola, J. (2008). *Reprints from the Early Days of Information Sciences: On the Contributions of Akira Nakashima to Switching Theory*. Tampere International Center for Signal Processing [online]. Available at http://ticsp.cs.tut.fi/index.php/Reprints_from_the_Early_Days_of_Information_Sciences.html [Accessed 16 August 2019].

State of Victoria. (2017). *What is sentience?* Victoria State Government [online]. Availabe at http://agriculture.vic.gov.au/pets/care-and-welfare/animals-and-people/what-is-sentience [Accessed 26 August 2019].

Szabo, N. (1994). *Smart Contracts*. Nick Szabo [online]. Available at http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html [Accessed on 17 June 2019].

Taleb, N. N. (2013). *Antifragile: Things That Gain From Disorder*. Penguin.

Techopedia. (2019). *Cryptographic Key*. technopedia [online]. Available at https://www.techopedia.com/definition/24749/cryptographic-key [Accessed 20 August 2019].

The Economist. (2017). *A clever way to transmit data on the cheap.* The Economist [online]. Available at https://www.economist.com/science-and-technology/2017/09/16/a-clever-way-to-transmit-data-on-the-cheap [Accessed 16 August 2019].

The Economist. (2019). *Schumpeter: The Exxon-Valdez of Cyberspace*. Available at https://www.economist.com/business/2019/08/08/the-exxon-valdez-of-cyberspace. Accessed 24 August 2019].

The Linux Foundation. (2019). *Hyperledger.* [online]. Available at https://www.hyperledger.org/ [Accessed 20 May 2019]

Thunberg, G. (2019). *No One is Too Small to Make a Difference*. UK: Penguin.

Toole, B. A. (1998). *Ada, The Enchantress Of Numbers: Prophet Of The Computer Age*. Strawberry Press [online]. Available at http://www.cs.yale.edu/homes/tap/Files/ada-lovelace-notes.html [Accessed 16 August 2019].

Torvalds, L. (2005). *Initial revision of "git", the information manager from hell.* Github [online]. Available at https://github.com/git/git/commit/e83c5163316f89bfbde7d9ab23ca2e25604af290 [Accessed 22 August 2019].

Turing, A. M., (1937). *On Computable Numbers, with an Application to the Entscheidungsproblem*. Proceedings of the London Mathematical Society s2-42, 230–265. https://doi.org/10.1112/plms/s2-42.1.230

van Rossum, G. (1995). *Python tutorial, Technical Report CS-R9526.* Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

Varian, H. R. (1998). *Markets for Information Goods*. University of California, Berkeley. Available at http://people.ischool.berkeley.edu/~hal/Papers/japan/ [Accessed 22 August 2019].

Veness, C. (2019). *SHA-256 Cryptographic Hash Algorithm*. Movable Type Scripts [online]. Available at https://www.movable-type.co.uk/scripts/sha256.html [Accessed 20 August 2019].

Vincent, J. (2019). *Bitcoin consumes more energy than Switzerland, according to new estimate*. The Verge [online]. Available at https://www.theverge.com/2019/7/4/20682109/bitcoin-energy-consumption-annual-calculation-cambridge-index-cbeci-country-comparison [Accessed 18 August 2019].

Vincent, S. (2018). *Willis Towers Watson launches reef insurance project*. insuranceday [online]. Available at https://insuranceday.maritimeintelligence.informa.com/ID1122617/Willis-Towers-Watson-launches-reef-insurance-project [Accessed 25 August 2019].

Vu, T. M. (2015). *Are the differences between Public Good and Common Pool Resource too blurred?* Researchgate [online]. Available at https://www.researchgate.net/post/Are_the_differences_between_Public_Good_and_Common_Pool_Resource_too_blurred [Accessed 25 August 2019].

Vyas, Tanvi. (2016). *No More Passwords over HTTP, Please!* Mozilla [online]. Available at https://blog.mozilla.org/tanvi/2016/01/28/no-more-passwords-over-http-please/ [Accessed 20 August 2019].

Waters, R. (2019). *Three ways that Big Tech could be broken up.* Financial Times, [online]. Available at: https://www.ft.com/content/cb8b707c-88ca-11e9-a028-86cea8523dc2 [Accessed 26 August 2019].

web3.js. (2019). *Callbacks Promises Events — web3.js 1.0.0 documentation*. [online]. Available at: https://web3js.readthedocs.io/en/v1.2.1/callbacks-promises-events.html. [Accessed 21 February 2019].

Weigert, M. (2015). *The surveillance state is inevitable*. Meshed Society [online]. Available at https://meshedsociety.com/the-surveillance-state-is-inevitable/ [Accessed 25 August 2019].

Wikipedia Contributors. (2019a). *Data type*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Data_type [Accessed 27 August 2019].

Wikipedia Contributors. (2019b). *Message*. Wikipedia [online]. Available at: https://en.wikipedia.org/wiki/Message [Accessed 12 August 2019].

Wikipedia Contributors. (2019c). *Homomorphic encryption*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Homomorphic_encryption [Accessed 10 May 2019].

Wikipedia Contributors. (2019d). *Transmission Control Protocol*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Transmission_Control_Protocol [Accessed 19 August 2019].

Wikipedia Contributors. (2019e). *History of free and open-source software*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/History_of_free_and_open-source_software [Accessed 19 August 2019].

Wikipedia Contributors. (2019f). *Zero-knowledge proof*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Zero-knowledge_proof [Accessed 26 August 2019].

Wikipedia Contributors. (2019g). *Government*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Government [Accessed 25 August 2019].

Wikipedia Contributors. (2019h). *Treaty*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Treaty [Accessed 23 August 2019].

Wikipedia Contributors. (2019j). *Checksum*. Wikipedia [online]. Available at https://en.wikipedia.org/wiki/Checksum [Accessed 24 August 2019].

Wilber, K. (2001). *A Brief History Of Everything*. Gill & Macmillan.

Williams, J. (2018). *Stand Out Of Our Light: Freedom and Resistance in the Attention Economy*. Cambridge University Press [online]. Available at https://www.cambridge.org/core/books/stand-out-of-our-light/3F8D7BA2C0FE3A7126A4D9B73A89415D [Accessed 14 January 2019].

Wood, G. (2014). ĐApps: What Web 3.0 Looks Like. Insights into a Modern World [online]. Available at http://gavwood.com/dappsweb3.html [Accessed 17 August 2019].

Wood, G. (2019). *Ethereum Yellow Paper.* Github [online]. Available at https://github.com/ethereum/yellowpaper [Accessed 15 October 2019].

Wu, H., Wang, F. (2014). A Survey of Noninteractive Zero Knowledge Proof System and Its Applications. *Scientific World Journal* [online]. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032740/ [Accessed 26 August 2019].

Wynn-Williams, C. E. (1931). *The use of thyratrons for high speed automatic counting physical phenomena*. The Royal Society Publishing [online]. Available at https://doi.org/10.1098/rspa.1931.0102. [Accessed 15 August 2019].

Xu, L.D., He, W., Li, S. (2014). *Internet of Things in Industries: A Survey*. IEEE Trans. Ind. Inf. 10, 2233–2243. https://doi.org/10.1109/TII.2014.2300753

Zamfir, V. (2017). *Against on-chain governance*. Medium [online]. Available at https://medium.com/@Vlad_Zamfir/against-on-chain-governance-a4ceacd040ca [Accessed 25 May 2019].

Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M. (2014). *Internet of Things for Smart Cities*. IEEE Internet Things J. 1, 22–32. https://doi.org/10.1109/JIOT.2014.2306328

Acknowledgements

This work would not have been possible without input and support over countless coffees and calls with an incredible group of experts, advisors, mentors and friends who shared their time and invaluable perspectives with me. An incomplete list:

From UCL: my supervisor Dr Duncan Wilson, Dr Sarah Wise, Dr Kira Kempinska, Dr Thomas Oléron-Evans, Kristian Lunow Nielsen, Katie Jamieson, Dagmar Ellefsen, Pili Mayora, Aude Vuilliomenet, Katherine Coates and Dr Abel Maciel.

From the blockchain community - through London Blockchain Labs: Yakko Majuri, Ilya Kleyner, Lorcan Delaney, Alex Terenda, Amber Rignell, Isaac Sultan, Alex Zakharov and Amanda Tan, as well as Jon Geater, Ben Bennett and James Worthington, Anthony Beaumont and Jean-Yves Rotté-Geoffroy. Lawrence Tilli and the BitBoysAndGirls crowd, and Vu Tien Khang.

From the peace, security and development community, Dr Curtis Bell, Dr Conor Seyle, Dr Sarah Glaser, Admiral Sir James Burnell-Nugent, Larry Sampler, Maisie Pigeon, Simon Williams, Dr Kuzi Charamba, John Filitz, Dr Clayton Besaw, Sean Duncan, Greg Clough, Melissa Hanham, Saskia Westhof and Camilo Casas, along with Dr Ian Stewart, Catherine Dill, Simon Ring, Mike Lewis, Adrian Wilkinson, Beth Clark, Mariyana Radeva Berket, and Janne and Andrew Kaiser-Tedesco.

From the London insurance world, Dr.rer.Pol. Magdalena Ramada Sarasola, Walid al Saqqaf, Aqua Sanfelice and Joe Mellen.

And from elsewhere - Jordan Bouley, Olivia Rhodes, Rollie Williams, Nicola Henderson, Zach Gorman, Chris Cote, Brandon Rattiner, Philipp Friemann, Ed Bayes, Lloyd Harrison, Caroline Clark, Lyle Barton, Jess Murray, Anne Milton, Helen Giles, Jen Carswell, Gemma Lyons and Michael Parratt.

And of course my family: Jocelyn, Jack, Emily, Jesse and Darlene. Thank you - 🙏.

‘And anything worth dying for,’ answered the sacrilegious old man, ‘is certainly worth living for.’

Joseph Heller
1961

We can possess nothing—neither thing nor thought—absent a border between self and State. Each individual right is derived from this line, which we call privacy.

@Snowden
24 August 2019