The background of the entire image is a circular collage composed of several different photographs. At the top, there are several sailboats on a body of water under a cloudy sky. To the right, a large, colorful sunset or sunrise dominates the upper right quadrant. Below the boats, a city skyline with numerous skyscrapers is visible. In the bottom left, a close-up view of a rocky beach with small pebbles is shown. The overall composition is a dense, multi-layered visual.

Money in Metaverses

Bitcoin in global social immersive
mixed reality systems

**John O'Hare, Allen Fairchild,
and Umran Ali**

[No Copyright](#)

2022 John O'Hare & Allen Fairchild

PUBLISHED BY J.OHARE@SALFORD.AC.UK FOR THE CYBER FOUNDRY PROGRAMME
AT THE UNIVERSITY OF SALFORD

[RAW GITHUB HYPERLINK](#)

The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See Other Information below.

This license is acceptable for Free Cultural Works. Other Information

In no way are the patent or trademark rights of any person affected by CC0, nor are the rights that other persons may have in the work or in how the work is used, such as publicity or privacy rights.

Unless expressly stated otherwise, the person who associated a work with this deed makes no warranties about the work, and disclaims liability for all uses of the work, to the fullest extent permitted by applicable law.

When using or citing the work, you should not imply endorsement by the author or the affirmer.

First printing, March 2022



Contents

I

State of the art and proposal

0.1	Conflict of interest statements	13
1	Introduction	14
1.1	Abstract	14
1.2	Introduction	14
2	Web3 / decentralised web	17
2.1	Semantic web	17
2.2	Spatial web	19
2.3	Web3	20
2.3.1	Emerging consensus	20
2.4	Example applications	22
2.4.1	Podcasting2.0	22
2.4.2	Crowd funding	22
2.4.3	Distributed exchanges	22
2.4.4	NFT marketplaces	23
2.4.5	Slashtags / Hypercore	23
2.4.6	Distributed DNS applications	23
2.4.7	Impervious browser	24
2.5	The common thread	24
3	DLT, Blockchain, and Bitcoin	26
3.1	Ethereum	28
3.1.1	Mining and Gas	29
3.1.2	Upgrade roadmap	31

3.2	Bitcoin	32
3.2.1	The Bitcoin Network Software	34
3.2.2	What are ECDSA, addresses and UTXO's	34
3.2.3	Mining and Energy concerns	36
3.2.4	Bitcoin nodes	38
3.2.5	Upgrade roadmap	40
3.3	Extending the BTC ecosystem	40
3.3.1	Block & SpiralBTC	40
3.3.2	BTCPayServer	41
3.4	Lightning (Layer 2)	41
3.4.1	Micropayments	42
3.4.2	BOLT12 and recurring payments	42
3.4.3	LNBits	42
3.4.4	Etleneum	45
3.4.5	Message passing	45
3.5	Liquid federation (layer 2)	45
3.6	Bitcoin Layer 3	46
3.6.1	LNP/BP and RGB	46
3.6.2	Synonym & Omnibolt	48
3.6.3	DLCFD	48
3.7	Bitcoin adjacent chains	48
3.7.1	Stacks and STX	49
3.7.2	Sovryn and RSK	49
3.7.3	Sidechains	49
3.7.4	Spacechains	49
3.7.5	Drivechain	50
3.7.6	Softchains	50
3.7.7	Statechains	50
3.8	Other chains and networks	50
3.8.1	Layer 1 chains	50
3.8.2	Crosschains	52
3.8.3	Decentralised data	52
3.9	Risks and mitigations	52
3.9.1	Digital assets	52
3.9.2	Bitcoin specifically	53
4	Money in the real world	54
4.1	Defining money	54
4.1.1	Global categories of capital	55
4.2	International money transfer networks	55
4.2.1	Swift, ISO 20022, and correspondence banking	55
4.2.2	VISA etc	56
4.2.3	Money transfer operators	56
4.2.4	Digital disruptive fintech	56

4.2.5	Stablecoins	56
4.3	Central bank digital currencies	57
4.4	Bitcoin as a money	60
4.5	Risks	62
4.5.1	Hyperbitcoinization	62
4.6	Does DeFi matter to SMEs	62
4.7	Do DAOs matter to SMEs	63
4.7.1	Bisq DAO	64
4.7.2	Risks	64
5	Distributed Self Sovereign Identity	65
5.1	Applications of DID/SSI	65
5.2	Classic DID/SSI	66
5.3	Newer Technologies	67
5.3.1	Slashtags	67
5.3.2	LNURL-Auth	67
5.3.3	Sovrin	67
5.3.4	Keri	67
5.3.5	Atala Prism	67
5.3.6	Microsoft ION	67
5.3.7	Atala Prism (ADA Ecosystem)	68
5.4	Risks & Challenges?	68
6	Non Fungible Tokens	69
6.1	Digital Art	69
6.2	Other uses	72
6.3	Is any of this useful?	72
6.4	User stories / behaviours	72
6.5	Comparing the technologies	72
6.6	User stories / behaviours	73
6.7	Comparing the technologies	73
6.8	What are the options	74
6.9	Risks	74
6.9.1	Wage slavery / pyramid schemes	74
6.9.2	Energy concerns	74
6.10	Why choose bitcoin again	75
7	Metaverses	76
7.1	History and market need	76
7.2	Current telecollaboration	76
7.2.1	Video conferencing	76

7.3	What's important for human communication	79
7.3.1	Vocal	79
7.3.2	Nonverbal	80
7.4	Psychology of Technology-Mediated Interaction	83
7.4.1	Proxemics	83
7.4.2	Attention	84
7.4.3	Behaviour	85
7.4.4	Dyadic	86
7.4.5	Triadic and Small Group	87
7.4.6	Presence, Co-presence, and Social Presence	88
7.4.7	Other Systems to Support Business	89
7.4.8	Mutual Gaze in Telepresence	90
7.4.9	Other Systems to Support Business	90
7.4.10	Beyond 2D Screens	91
7.4.11	Network Issues	101
7.5	Support for Less Formal?	101
7.5.1	Support for Dynamic Meetings	103
7.6	Summary	103
7.6.1	Comparing Key Systems and their Affordances	104
7.7	Theoretical Framework	105
7.7.1	Problem Statement	105
7.7.2	Core Assumptions	105
7.7.3	Peripheral Assumptions	107
7.8	Post 'Meta' metaverse	107
7.8.1	Mixed reality as a metaverse	108
7.9	Digital Land Metaverses	108
7.9.1	Legacy Web2	108
7.9.2	The new stuff	108
7.10	Global enterprise perspective	109
7.11	NFT as metaverse narrative	109
7.12	MMORG games and NFTs	109
7.13	Crypto metaverses	111
7.13.1	Immersive and third person XR	111
7.13.2	WebXR/WebGL/WebGPU	113
7.13.3	Integration with web and game engines	114
7.14	User stories	114
7.15	Recommendations for value transfer	114
7.15.1	Streaming micropayments	114
7.16	Bitcoin market gap	114
7.16.1	Money	114
7.16.2	Identity proofs	114
7.16.3	Digital object tracking	114
7.16.4	Object transfer and trading	114

7.17 Risks	114
8 Hardware and software choices to support this	115
 8.1 Networking layer	115
8.1.1 Proxmox	115
8.1.2 VyOS firewall	115
8.1.3 Privacy aspects and using TOR etc	115
 8.2 Bitcoin value management stack	115
8.2.1 Bitcoin Core on NixOS	115
8.2.2 Electrum server	115
8.2.3 C-Lightning and CLBoss	116
8.2.4 LDK based Sensei	116
8.2.5 LNBits with RGB and metaverse plugin	116
8.2.6 Backup & Watchtower	116
8.2.7 Object and media tracking: RGB	116
 8.3 Identity	116
8.3.1 nostr	116
 8.4 Metaverse	116
8.4.1 Vircadia	116
8.4.2 Wolvic	116
8.4.3 WebM	116
 8.5 Messengers and boards	116
8.5.1 Nostr	116
8.5.2 Cyberpost	116
8.5.3 Matrix	117
 8.6 Addressing identified risks	117
9 Current example deployment	119
 9.1 GitHub	119
10 Our proposition	121
 10.1 Potential applications	121
10.1.1 Global cybersec course delivery	122
11 Conclusions	125
12 Future work	126

II

Guides for deploying the software

12.1 Lab	128
12.1.1 Overview	128
12.1.2 Prerequisites	128
12.1.3 Network details	128

12.1.4	Server configuration	128
12.1.5	Proxmox VE	128
12.1.6	Setup an internal only network in Proxmox VE	130
12.1.7	Install and configure Internet gateway server virtual machine	130
12.1.8	Install and configure a Debian virtual machine	132
12.1.9	Deploying the nix-bitcoin node	132

III

Appendix

12.2	Acknowledgements and thanks	136
12.3	Author Biographies	136



List of Figures

1.1	Web 3, Metaverse, and Bitcoin are inter-sectional technologies.	16
2.1	Semantic Web Stack (CC0 image)	18
2.2	Deloitte Spatial Web Overview Reused with permission.	19
2.3	Edelman 2020 trust barometer (rights requested)	20
2.4	A meme showing differing approached to logging in on a website.	21
2.5	ARK slide on Web3. Rights requested	25
3.1	Dan Held: Bitcoin prehistory used with permission.	27
3.2	Intersecting disciplines. Reused with permission Dhruv Bansal	28
3.3	Ethereum is thought to look like a speculative bubble. Rights requested .	30
3.4	The rate of token generation has changed unpredictably over time. Rights requested	31
3.5	Growth in settlement value on the Bitcoin network.	33
3.6	Bitpaint: Contributions to the Bitcoin ecosystem. Reused with permission.	34
3.7	Given a start point on the curve and a number of reflection operations it's trivial to find a number at the end point, but impossible to find the number of hops from the two end points alone.	35
3.8	Intimate tie between energy and money, Henry Ford	37
3.9	Hash rate suddenly migrates from China (Reuse rights requested)	37
3.10	Goldman suggest growth opportunity and potential demonetisation of gold?	39
3.11	Arcane research lightning adoption overview	43
3.12	Two of the many prebuilt and kit options for Lightning 'point of sale' . .	44
3.13	Allocations given at the beginning of public blockchain, by Messari. . .	50
4.1	Potential market exposure to Bitcoin as a money	61
5.1	Part of the DID SSI specs	66
7.1	Elon Musk agrees with this on Twitter	77

7.2 Eye tracked eye gaze awareness in VR. Murray et al. used immersive and semi immersive systems alongside eye trackers to examine the ability of two avatars to detect the gaze awareness of a similarly immersed collaborator.	84
7.3 The Venn diagram shows areas of research which have been identified in blue. These interlock and overlap as shown. The most relevant identified researchers from the literature are shown in black close to the fields of study which they represent. This diagram is a view of the core assumptions for the research, with the most important fields at the centre.	106
7.4 Connected concepts revealed during the review, and roughly how they relate to one another. The web moves top left to bottom right as the attention focuses. The left hand column shows the broad flow from interpersonal communication in the real (top) to home working (bottom).	107
8.1 Proposed deployment of the software within the VMs on a single hardware system.	118
9.1 Current diagram of the proxmox as seen on the github.	120
10.1 Functional elements for infrastructure.	123
10.2 Client server C4 diagrams.	124

List of Tables

6.1 This table is basically broken and out of date and will be sorted out soon! 74

State of the art and proposal

1	Introduction	14
2	Web3 / decentralised web	17
3	DLT, Blockchain, and Bitcoin	26
4	Money in the real world	54
5	Distributed Self Sovereign Identity	65
6	Non Fungible Tokens	69
7	Metaverses	76
8	Hardware and software choices to support this	115
9	Current example deployment	119
10	Our proposition	121
11	Conclusions	125
12	Future work	126

0.1 **Conflict of interest statements**

The authors own small numbers of the various tokens referenced in the text for experimentation and/or investment purposes. This includes Solana, Ethereum, and Bitcoin locked on the Lightning network. No NFTs are owned at this time. There are no financial stakes in the development of any of these ecosystems.



1. Introduction

1.1 Abstract

We present a state of the art and positioning book, about Web3, Bitcoin, and ‘Metaverse’; describing the intersections and synergies. A high level overview of Web3 technologies leads to a description of blockchain, and the Bitcoin network is specifically selected for detailed examination. Suitable components of the extended Bitcoin ecosystem are described in more depth. Other mechanisms for native digital value transfer are described, with a focus on ‘money’. Metaverse technology is over-viewed, primarily from the perspective of Bitcoin and extended reality.

Bitcoin is selected as the best contender for value transfer in metaverses because of it’s free and open source nature, and network effect. Challenges and risks of this approach are identified. A cloud deployable virtual machine based technology stack deployment guide with a focus on cybersecurity best practice can be downloaded from GitHub to experiment with the technologies. This deployable lab is designed to inform development of secure value transaction, for small and medium sized companies. Next steps are touched upon in the conclusion.

1.2 Introduction

This document presents a high level view of technologies and their potential within the emergent Web3 and metaverse narrative, focusing around the transmission of value within and across such global networks, with a further focus on the Bitcoin monetary network. It was written to organise the thoughts of the authors, who were unfamiliar with Bitcoin technologies until recently.

As adoption of these technologies increases it will be necessary for people, and AI actors, to pass economic value between themselves. These ‘goods and services’ interactions, within the virtual social spaces, should be underpinned by a trust system, which scales globally, and presents low friction. Current secure international payment rails are poorly suited to such interactions; indeed it is likely with legacy systems, that parties would be forced to leave the metaverse application, and instead navigate their banking applications to exchange value with overseas entities in a secure fashion. This might conceivably take

several days.

Fortunately, the whole landscape of money and [value transfer is changing](#). Huge global financial players are entering the space. HSBC have [just bought](#) metaverse ‘land’ in The Sandbox, JP Morgan have [opened a ‘lounge’](#) in another, and the world’s largest pension fund manager Blackrock is adding these asset to their management engine which manages tens of trillions of dollars. Of their recent investments KPMG global said:

“We’ve invested in a strong cryptoassets practice and we will continue to enhance and build on our capabilities across Decentralized Finance (DeFi), Non-Fungible Tokens (NFTs) and the Metaverse, to name a few”.

It’s actually likely that for such organisations it makes better business sense to take a punt on hype bubbles like this, than to do a proper due diligence with a team of internal staff who understand their business. These endorsements should be taken with a large pinch of salt. As [Alex Jones says](#):

“At some point in the future, it’s possible that the digital worlds being built today will have aggregated sufficient user attention and engagement that financial services companies will need to invest in the metaverse as an acquisition and customer service channel. But we’re not there yet. Until the metaverse is a little less empty, resist the temptation to colonize it with branches and billboards.”

Meanwhile, Google have [recently blogged](#): “*Web3 also opens up new opportunities for creators. We believe new technologies like blockchain and NFTs can allow creators to build deeper relationships with their fans. Together, they’ll be able to collaborate on new projects and make money in ways not previously possible. For example, giving a verifiable way for fans to own unique videos, photos, art, and even experiences from their favourite creators could be a compelling prospect for creators and their audiences. There’s a lot to consider in making sure we approach these new technologies responsibly, but we think there’s incredible potential as well. Finally, we couldn’t have a piece about innovation without touching on the metaverse! We’re thinking big about how to make viewing more immersive.*”

It’s already the case that the recent bubble of [hype is dwindling](#), but there enormous investment into teams and startups will potentially bear fruit in this coming period, and this perhaps implications for (Small and medium-sized enterprises). It is timely then to explore the potential of recent technologies, which can address metaverse interactions in *business to business* (B2B), *business to customer* (B2C), and the newer C2C (social commerce; *creator to consumer, customer to customer, consumer to consumer*[[1](#)]). Figure 1.1 demonstrates how some of these domains intersect.

This book seeks to overview and explain the available open source technologies. It supports an open source [github repository](#) which enables SMEs to access these emergent platforms and ecosystems. It aims to build toward a minimum viable product for trust minimised transfer of value within a social immersive space.

Referencing is in two styles; academic works are numeric, while opinion pieces, grey statistics, and pertinent news articles are hyperlinked in blue from the text. As the book develops the veracity of the citations should improve. It is likely that this document is incomplete as you are reading it as it’s designed to be an open ended and open source resource. It is an experiment in a living document and repository.

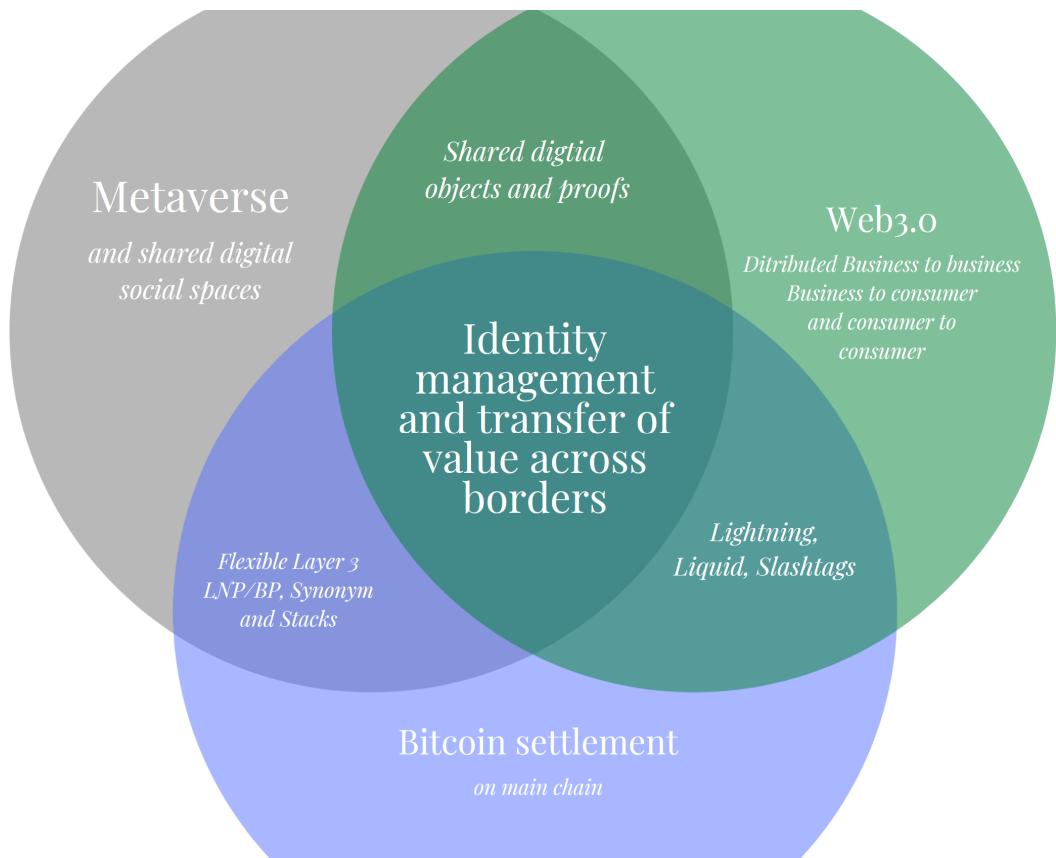


Figure 1.1: Web 3, Metaverse, and Bitcoin are inter-sectional technologies.



2. Web3 / decentralised web

Web3 is a rapidly evolving set of technologies and specifications which are drifting further from their origin. Decentralised web is perhaps a more useful name, but focus in this section will be on the evolution of the popularised term Web3.

2.1 Semantic web

The “semantic web” definition of Web3.0 has been somewhat overhauled by other innovations in decentralised internet technologies, now evolving toward the slightly different Web3 moniker. Tim Berners Lee (of WWW fame) first mentioned the semantic web in 1999 [2].

“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A “Semantic Web”, which makes this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The “intelligent agents” people have touted for ages will finally materialize.”

Attention developed around three core themes, ubiquitous availability and searchability of data, intelligent search assistants, and highly available end points such as phones, and ‘internet of things’ devices. This is certainly manifesting in home devices, but few people think of this as a Web3 revolution. The framework can be seen in Figure 2.1.

Since ratification of the standards by the [World Wide Web \(W3C\) consortium](#) it seems that their imperative toward decentralisation has become lost. Instead, it can be seen that Facebook, Amazon, Google, and Apple have a harmful oligopoly on users data [3]. This is as odds with Berners-Lee’s vision.

It is worth taking a look at his software implementation called [Solid](#), which is far more mindful of the sovereignty of user data.

“Solid is an exciting new project led by Prof. Tim Berners-Lee, inventor of the World Wide Web, taking place at MIT. The project aims to radically change the way Web applications work today, resulting in true data ownership as well as improved privacy. Solid (derived from “social linked data”) is a proposed set of conventions and tools for building decentralized social applications based on Linked Data principles. Solid is modular and



Figure 2.1: Semantic Web Stack [CC0 image]

Three tiers of IT infrastructure and building the Spatial Web

As the technologies and capabilities that compose and connect IT architecture converge, the Spatial Web will mature. The figure below shows how key enabling technologies drive their respective computing eras.



*Note: Date ranges are approximate and meant for directional purposes only.

Source: Deloitte analysis adapted from Gabriel René and Dan Mapes, *The Spatial Web: How Web 3.0 Will Connect Humans, Machines, and AI to Transform the World* (Amazon, 2019).

Deloitte Insights | deloitte.com/insights

Figure 2.2: [Deloitte Spatial Web Overview](#) Reused with permission.

extensible and it relies as much as possible on existing W3C standards and protocols.”

Excitement around this kind of differentiated trust model, hinted at in ubiquitous availability of data (and implemented explicitly in Solid), has led to exploration of different paths by cryptographers, and this will be described later. For instance, one of the main developers of Solid, Melvin Carvelho, is now a lead developer at Nostr, another very interesting option which will be described later. This technology space is prolific, but still comparatively young and small.

2.2 Spatial web

“The Spatial Web”, a blurring of the boundaries between digital and geospatial physical objects, seems to have developed from the strands in the original W3C scope around devices in the real world. It has been concentrating around AR and VR but is being marketed and amplified with the same references to availability of data (See Figure 2.2 from a Deloitte accounting report). This too is finding little traction in practice, though obviously the component technologies continue to enjoy rapid development. Nonetheless, this interpretation of Web3 becomes valuable when examining Metaverse later.



Figure 2.3: Edelman 2020 trust barometer [rights requested]

2.3 Web3

More recently Web3 is [being touted](#) as a way to connect content creators directly to content consumers, without centralised companies acting as gatekeepers of the data. It implies that all users have a cryptographic key management system, to which they attach metadata, that they make requirements of peers with whom they communicate, and that they maintain trust ‘scores’ with peers.

It seems likely that this new model is less driven by a market need, and more by the high availability of tools which allow this to happen (the ecosystems described later). Add to this a social response to the [collapse in trust of companies such as Facebook](#) (Figure 2.3), a wish by consumers to pass more of the economic incentive to content creators, without the ‘rent seeking’ layer afforded by businesses, and a healthy dose of mania driven market speculation.

2.3.1 Emerging consensus

It’s possible to frame Web3 as a hugely complex and inefficient digital rights management system (DRM). DRM is something that users of the internet are increasingly familiar and comfortable with. It’s somewhat debatable whether decentralising this is worthwhile. The thesis of the developers of the technology seems to be that without it, control of ‘value’ will

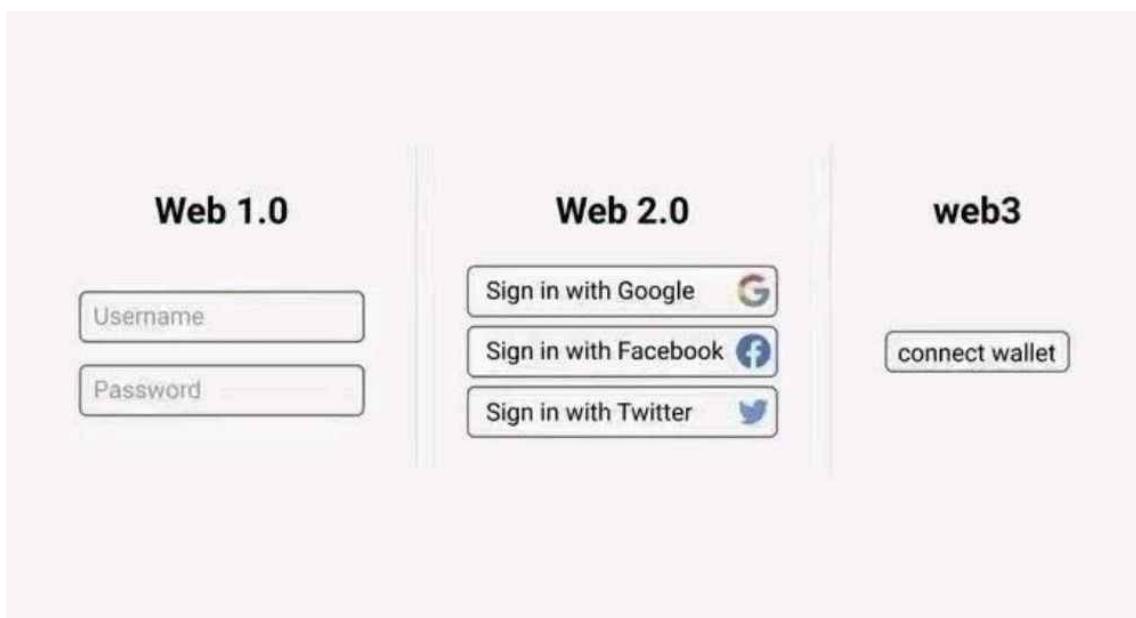


Figure 2.4: A meme showing differing approached to logging in on a website.

accrete over time, to one or more hegemonic controlling entities. It's a strong argument, but there is a [substantial counter argument](#) emerging that users just don't want this stuff. The nervousness of legislators in the USA to the attempt by Facebook/Meta to enter this peer-to-peer value transmission space is telling in terms of the perception of who is driving Web3.

At the end of 2021 and the beginning of 2022 there is much furore on the internet over what Web3 might be, and who it 'serves'. Enthusiasts feel that products such as [Sign-In with Ethereum](#) (EIP-4361) will give users choice over their data sovereignty, and a meme to this effect is seen in Figure 2.4. In practice though users are expecting to use badly written, buggy, economically vulnerable 'crypto' wallets to log into websites. It doesn't seem to make much sense yet on the face of it. There are in fact examples of the technology completely failing at censorship resistance. Popular 'Web3' browser extension Metamask and NFT platform Opensea have both [recently banned countries](#) in response to global sanction pressure. This failure to meaningfully decentralise will be explored further in the distributed identity section.

The current hype cycle is ignoring the legacy definitions described above and instead focusing almost exclusively on Ethereum based peer to peer projects (more on these later). It can be seen that the description is somewhat in the eye of the beholder.

This new hyped push for Web3 is being driven by enormous venture capital investment. A16Z are a [major player](#) in this new landscape and have released their [ten principles](#) for emergent Web3.

- Establish a clear vision to foster decentralized digital infrastructure
- Embrace multi-stakeholder approaches to governance and regulation
- Create targeted, risk-calibrated oversight regimes for different web3 activities
- Foster innovation with composability, open source code, and the power of open communities
- Broaden access to the economic benefits of the innovation economy
- Unlock the potential of DAOs

- Deploy web3 to further sustainability goals
- Embrace the role of well-regulated stablecoins in financial inclusion and innovation
- Collaborate with other nations to harmonize standards and regulatory frameworks
- Provide clear, fair tax rules for the reporting of digital assets, and leverage technical solutions for tax compliance

Jeremy Allaire of ‘Circle’ venture capital talks about the recent legislative order in the USA as follows:

“this is a watershed moment for crypto, digital assets, and Web 3, akin to the 1996/1997 whole of government wakeup to the commercial internet. The U.S. seems to be taking on the reality that digital assets represent one of the most significant technologies and infrastructures for the 21st century; it’s rewarding to see this from the WH after so many of us have been making the case for 9+ years.”

Many of the mentioned components here will be described later in the book. This list seems targeted toward the coming regulatory landscape, and could be considered at odds with the original tenants of an organically emergent, decentralised internet.

Dante Disparte, chief strategy officer of Circle, said in testimony to a US senate hearing; that Web 1 was ‘read’, Web 2 was ‘read write’, and that Web 3 will ‘read write own’. The important takeaway here is not so much this oft quoted elevator pitch for Web3, but the fact that legislative bodies now consider this technology a force which they need to be aware of and [potentially contend with](#). The suggestion is that this is happening regardless of a decent use case or definition.

It’s a complex evolving narrative, and clearly contradictions are common. Into this confusion this book advances a narrow take, and toolset, which might extract some value from the technologies, while maintaining a low barrier to entry.

2.4 Example applications

2.4.1 Podcasting2.0

[Podcasting 2.0](#) leverages RSS (the original dissemination system for podcasts) and the Bitcoin Lightning network, to enable so-called ‘value for value’ broadcasting. Subscribers use one of a variety of apps to stream micro-transactions of Bitcoin directly to the content creators as they listen to the podcast. No intermediate business takes a cut. Some variation on this model exists, such as John Carvalho’s crowd funded podcast “The Biz” which progressively unlocks more minutes for everyone based on [crowd funded donations](#).

2.4.2 Crowd funding

At time of writing a [crowd funding initiative](#) based around a digital decentralised construct called a DAO (explained later in detail) [managed to raise \\$46 million dollars to bid for a copy of the US constitution](#) at Southerbys auction house. The attempt narrowly failed, but the press [heralded this new era of “Web3” economic might](#). This model might be the only use for DAOs and is likely just a way to avoid regulatory scrutiny. There is more detail on DAOs later.

2.4.3 Distributed exchanges

There are dozens of decentralised exchanges deployed on various blockchains. These platforms allow users to trade back and forth between various tokens (including ‘normal

money’ stablecoins) and charge a fee for doing so. They operate within the logic of the smart contracts, within the distributed blockchains. This makes them extremely hard to ban, and as a result they operate in a legal grey area. At the extreme end of this is “distributed apps” (dApps) and “Decentralised Finance” (DeFi) which allows users access to complex financial instruments without legal or privacy constraints. DeFi will be touched on briefly later.

This is a huge area, and of only limited interest to the topics expanded in this book. It’s perhaps worth noting [BitcoinDEX](#), which runs in JavaScript in a web browser. It is effectively uncensorable, [auditable by the user](#), and has no counter party risk since it operates entirely in the Bitcoin network. It is clearly an early prototype but manages this complex feature without the more expressive logic of more ‘modern’ public blockchains.

2.4.4 NFT marketplaces

NFT markets are far more centralised services which match ‘owners’ of digital assets with potential buyers. The concept is a staple of the more recent interpretation of Web3, even though in practice these seem to be centralised concerns. [Opensea](#) claims to be the largest decentralised NFT marketplace, but they have the ability to [remove listings](#) in response to legal challenges. This seems to fly in the face of Web3 principles. NFTs are currently a [deeply flawed](#) technology but seem likely to persist and will be covered later.

2.4.5 Slashtags / Hypercore

Slashtags is a web of trust model decentralised peer-to-peer model which assigns metadata and trust scores to ‘any’ data and connection, with a security model rooted in the Bitcoin cryptographic ‘keys’ but crucially not the bitcoin network. This makes it interoperable with bitcoin but not reliant upon it. In principle this allows users to build complex networks of inherited trust bi-directionally with their networks over time. Every connection to a peer can be a new schema, with individual metadata managed by the user. It is new and has low adoption at this time. The user controls the source of the data and can allow them to be used by centralised services. This flips the authentication and data management paradigm of Web2 around, putting the user in charge of their data. This is a familiar concept to the DID/SSI communities (described later) but with significant investment. As Slashtags use keys as endpoints they act as a web of naming and routing, bypassing the existing web infrastructure of DNS. It is likely very complex to use in practice and will be revisited later. It is being paired with the [Hypercore protocol](#) for peer-to-peer data sharing, more specifically the ‘hole punching’ capability of the hypercore system which ensures connections through firewalls^[4]. The first application by the affiliated Hyperdivision team is an open source peer-to-peer live video streaming app called [Dazaar](#). Once again, it’s not clear yet who wants or needs this bit-torrent style service.

2.4.6 Distributed DNS applications

There are many perceived problems with having centralised authorities for overseeing the database which translates between human readable internet names and the underlying machine-readable address notation. The databases which manage this globally are already somewhat distributed, and this distributed trust model is managed through a cryptographic chain of trust called DNSSEC which is capped by a somewhat [bizarre key ceremony](#). The authority around naming is centralised in ICANN. There has been talk for many years

about ‘properly’ distributing this database using decentralised/blockchain technologies[5]. The nature of this problem means that it either moves from control by ICANN, or it does not, and so far it has not, but there are many attempted and somewhat mature attempts at this difficult problem. Of these [Namecoin](#) is the most prominent, and is a fork of Bitcoin. The ubiquity of Bitcoin in such systems is perhaps becoming apparent.

2.4.7 Impervious browser

It might be that the future of Web3 comes in the guise of integrated suites such as the proposed [Impervious web browser](#). They say that “without centralized intermediaries” it features:

- Zoom, without Zoom.
- Google Docs, without Google.
- Medium, without Medium.
- WhatsApp, without WhatsApp.
- Payments, without banks.
- Identity, without the state.

This is obviously leading marketing hype, but what they’re talking about here is an integration of the components mentioned in this book. If they can get critical mass around this browser then perhaps the Web3 market can be kickstarted.

2.5 The common thread

One feature which persists throughout all of these interpretations of Web3 is the need for decentralised trust. If there is to be no central controlling party(s) as in the Web 2 model then nothing can happen without a cryptographically secure underpinning, which favours no party beyond the terms of their collectively agreed rights and privileges. From this base layer we also get the potential for secure and trust minimised identity management. This nascent field of distributed identity management is explained later. From distributed trust models we can see ‘trustless’ transmission of economic value. The ability to send value from one person to another person or service without a third party. This whole area is ‘crypto’, which is increasingly seeping into the human consciousness, and saw an astonishing \$13B of [capital investment in 2021](#) alone. At time of writing the industry is a [over 2 trillion](#) dollar market. According to [Nathaniel Whittemore](#), a journalist for Coindesk, “The Web3 moniker positions this industry in opposition to big tech”. In practice it seems that this is just another avenue for existing players to experiment with new [models of control and monetisation](#), as in [this announcement](#) from Stripe, the worlds biggest payment processor. Overall, the space is hype, and [rife with scams](#).

Of their 2022 ‘[Big Ideas](#)’ report, ARK investment LLC (who manage a \$50B tech investment) [said the following](#) (Figure 2.5), which connects some of the dots already mentioned, and leads us into the next section which is Blockchain and Bitcoin:

“While many (with heavily vested interests) want to define all things blockchain as web3 we believe that web3 is best understood as just 1 of 3 revolutions that the innovation of bitcoin has catalyzed.

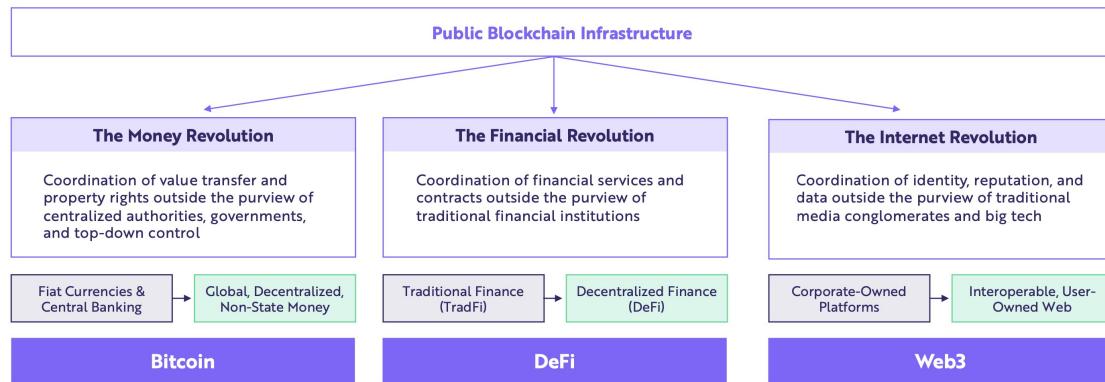
- *The Money Revolution*
- *The Financial Revolution*
- *The Internet Revolution”*

All the new ‘crypto’ technologies circling the Web3 narrative are bound tightly together,



Public Blockchains Are Stirring Several Revolutions

In our view, the Bitcoin protocol created the most profound application of public blockchain infrastructure. In addition to the Money Revolution, public blockchains also have catalyzed Financial and Internet Revolutions.



Forecasts are inherently limited and cannot be relied upon. | For informational purposes only and should not be considered investment advice, or a recommendation to buy, sell or hold any particular security/cryptocurrency.
Source: ARK Investment Management LLC, 2021

Figure 2.5: ARK slide on Web3. Rights requested

but there is currently very little meaningful value to be seen.

The rest of this book will focus on the trust and value transfer elements of this shift in internet technologies, and attempt to build a case for its use in decentralised, open source, metaverse applications.



3. DLT, Blockchain, and Bitcoin

Distributed ledger technology (DLT) is a data structure distributed across multiple managing stakeholders. A subset of DLT is blockchain, which is a less efficient, immutable data structure with a slightly different trust model. Rauchs et al. of the Cambridge Centre for Alternative Finance provide a detailed taxonomy and conceptual framework [6]. It can be seen in their paper that the definitions are somewhat unclear in literature.

DLT, and especially blockchain, are rapidly gaining ground in the public imagination, within financial technology companies (FinTech), and in the broader corporate world.

The technology and the global legislative response are somewhat immature, and misapplications of both technologies are commonplace.

Distributed trust models emerged from cryptography research in the 1970s when Merkle, Diffie, and Hellman at Stanford worked out how to [send messages online](#) without a trusted third party [7, 8].

Soon after the 1980s saw the emergence of the cypherpunk activist movement, as a reaction to the emerging surveillance state [9, 10]. These early computer scientists in the USA saw the emerging intersectionality between information, computation, economics, and personal freedom [11]. Online discussion in the early nineties foresaw the emergence of trans-national digital markets, what would become the WWW [12, 13]. The issues of privacy and the exchange of digital value (digital / ecash) were of foremost importance within these discussions and while privacy was within reach thanks to “[public/private key pairs](#)”, ecash proved to be a more difficult problem.

Adam Back’s 1997 ‘hashcash’ [14] paved the way for later work by introducing the concept of ‘proof of work’. This was built upon by Dai [15], Szabo [16], Finney [17], and Nakamoto amongst others. In all it took 16 years of collaboration on the mailing lists to attack the problem of trust-minimised, distributed, digital cash. The culmination of these attempts were Bitcoin [18]. This is illustrated by Dan Held in Figure 3.1.

This is now a wider ecosystem of technologies.

There is enormous complexity and scope, as seen in Figure 3.2, and yet genuinely useful products are elusive. It can be argued that the whole concept of a distributed cryptographic blockchain is somewhat flawed, as the vast majority of the technology offerings are not properly distributed, and “there are many scenarios where traditional databases should be used instead”[19].

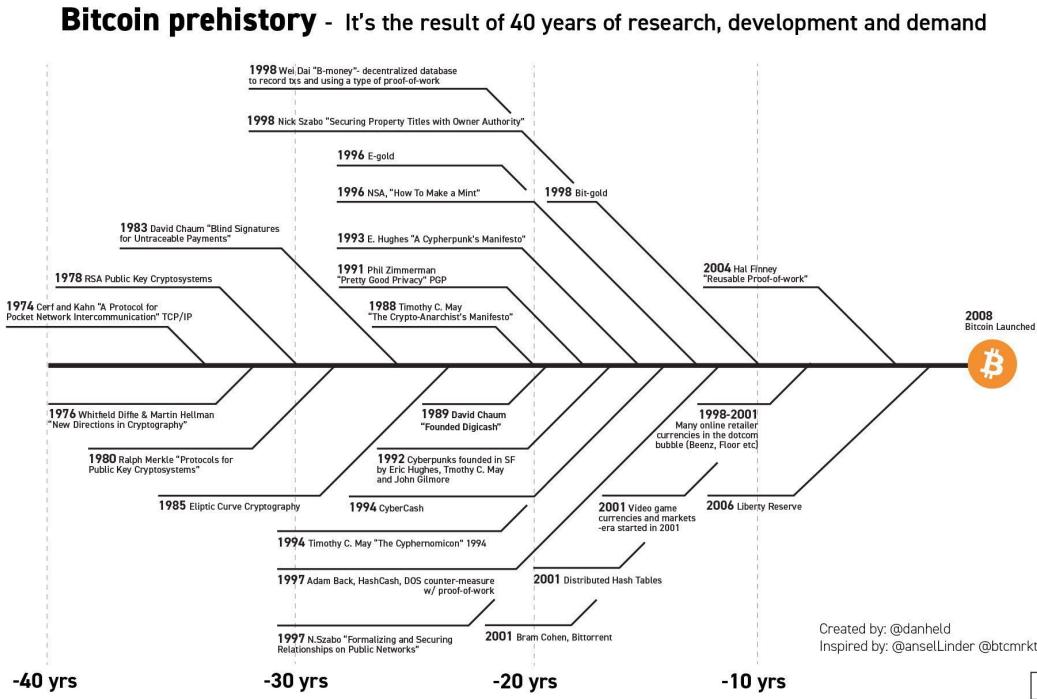


Figure 3.1: Dan Held: [Bitcoin prehistory](#) used with permission.

It is surprisingly hard to pin down a simple explanation for the features which define a blockchain. These “key takeaway” from [Investopedia](#) are a neat summary however.

- Blockchain is a specific type of database.
- It differs from a typical database in the way it stores information; blockchains store data in blocks that are then chained together.
- As new data comes in it is entered into a fresh block. Once the block is [filled with data](#) it is chained onto the previous block, which makes the data chained together in chronological order.
- Different types of information can be stored on a blockchain but the most common use so far has been as a ledger for transactions.
- In Bitcoin’s case, blockchain is used in a decentralized way so that no single person or group has control—rather, all users collectively retain control.
- Decentralized blockchains are “append only”. In effect this means that the data entered becomes irreversible over time. For Bitcoin, this means that simple economic transactions are permanently recorded and viewable to anyone.

In principle blockchains provide a differentiated trust model. With a properly distributed system a blockchain can be considered “trust-minimised”. This is important for some, but not all people. In an era when data breaches and corporate financial insolvency intersect with a collapse in trust of institutions, it is perhaps useful to have an alternative model for storage of data, and value. Thanks to a natural fit with strong encryption, and innate resistance to censorship by external parties, these systems lend themselves well to ‘borderless’ applications. Finally, a host of well engineered open source code repositories makes the cost of adoption relatively low.

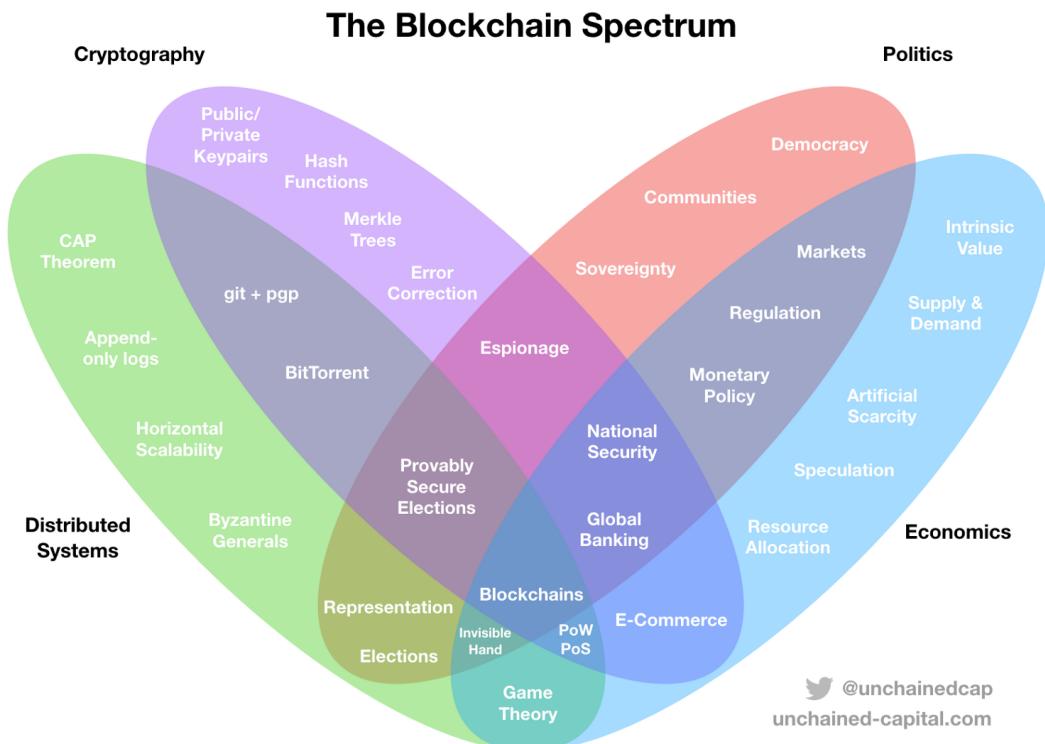


Figure 3.2: Intersecting disciplines. Reused with permission [Dhruv Bansal](#)

Within DLT/blockchain there seem to be as many opinions on the value of the technology as there are implementations. There are thousands of different ‘chains’ and many more tokens which represent value on them. A majority of these are code forks of earlier projects. Most are defunct yet still have some residual ‘value’ locked up in them as a function of their ‘distributed’ tokens.

Because the space is comparatively new, subject to [scant regulation](#), and often open source, it is possible to clone a github, change a few lines of code, and front it with a website in order to create ‘scams’, and this happens frequently [20].

The following sections give an overview of the major strands of the technology. First is Ethereum, mainly to discount it’s use and get to more accessible offerings.

3.1 Ethereum

Ethereum [21] is the second most secure public blockchain ([by about 50%](#)), and second most valuable by [market capitalisation](#) (though this comparison is somewhat strained). It is the natural connection from Web3 to the rest of the paper, so it will be considered first.

It is touted as ‘programmable money’. It, unlike bitcoin, is Turing complete, able to run a [virtual machine](#) within the distributed network (albeit slowly), and can therefore process complex transactional contracts in the settlement of value. This has given rise to the new field of ‘distributed finance’, or DeFi (described later), alongside many interesting trust-minimised immutable ledger public database ideas.

There are trade-offs and problems with Ethereum (Eth/Ether) which currently increase

the 'participation floor' and make the network far less suitable for entry level business-to-business use. The ledger itself being a computational engine, with write only properties, is enormous. Specialist cloud hardware is required to run a full node (copy of the ledger), and partial nodes are the norm. Even partial nodes are run chiefly by one specialist cloud provider ([Infura](#)), which has recently been forced to [exclude Venezuela](#) from the network. Moreover the network is centrally controlled by its creator and the 'miners'. There is a strong case to answer that Eth is [neither distributed](#), nor trustless, and in fact therefore fails to be differentiated from a DLT, undermining some of it's claims.

With that said there are many talented developers doing interesting work on the platform, and innovation is fast paced. It is entirely normal for technology projects to launch their distributed ledger idea on and within the Ethereum network. These generate tradable 'ERC20' tokens, which can accrue value or demonstrate smart contract utility. Because the value locked and generated in the Ethereum platform comes not just from the ETH token, but all the ERC technologies built upon it, there are hundreds of billions of pounds 'within' the network. Most of this money is pure market speculation (as is the case across blockchains). Many analysts cannot see this as anything but a speculative bubble, with all the predictable crash yet to come. This can be seen in the context of other bubbles in Figure 3.3. It seems that most of the projects in crypto more generally, but certainly with ETH and the NFTs within it are a new kind of social gambling, where online communities can reinforce groupthink around their speculative choices.

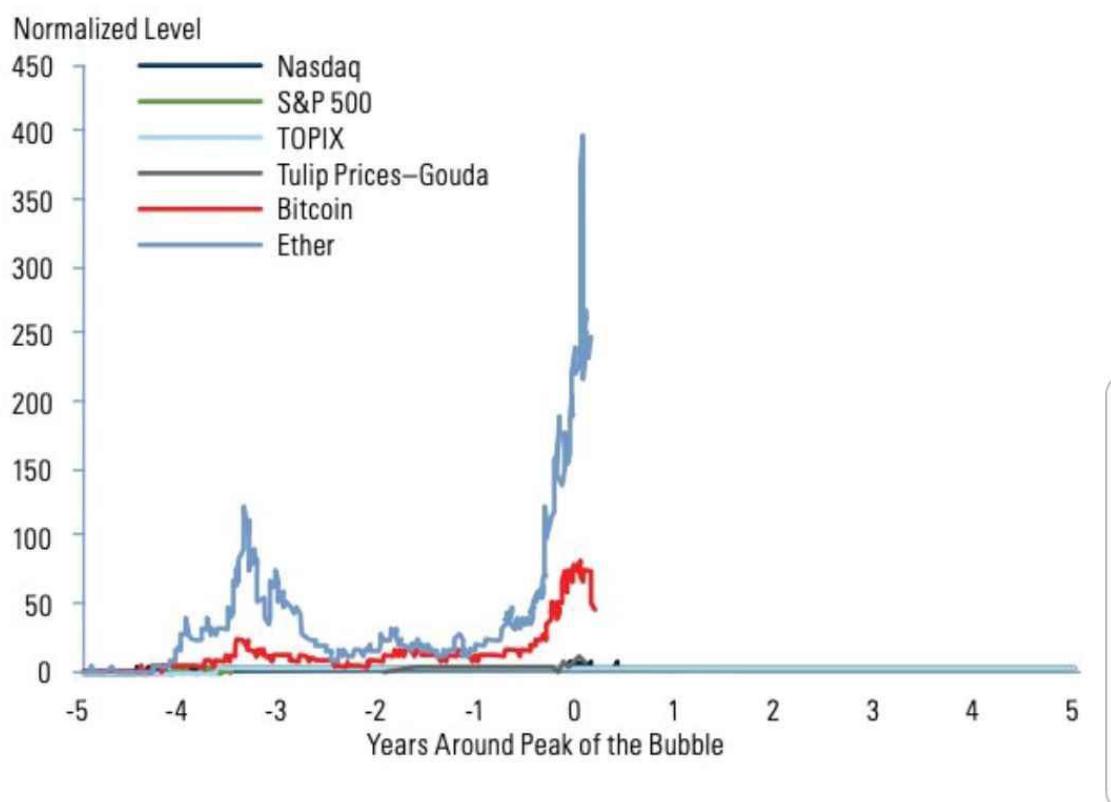
Such is the level of nefarious activity on these networks that they have a poor reputation, and are difficult to audit, launch, and maintain. The overriding problem of using a blockchain for utility applications is that people can and will simply lie for criminal purpose when entering data into the ledger. It is far more likely that Ethereum is simply a speculative bubble than any of the claims for utility being born out. Add to that [Morgan Stanleys recent assertion](#) that Ethereum is itself threatened by newer contender chains and it's future becomes unclear. The report correctly identifies that "High transaction fees create scalability problems and threaten user demand. High costs make Ethereum too expensive for small-value transactions."

3.1.1 Mining and Gas

Ethereum has a significant barrier to entry because of high fees to use the network. The system is Turing complete; able to programmatically replicate any other computational system. This includes endless loops in code, so it is trivial to lock up the computational bandwidth of the whole system, in a smart contract commitment, through a web wallet. To mitigate this existential 'denial of service attack' the 'gas' system demands that users spend some of their locked up value to operate on the network. In this way a transaction loop would quickly erode the available gas and stop looping. As the popularity of the system has grown, so too have the gas fees. It can sometimes cost hundreds of dollars to do a single transaction, though it is typically a few tens of dollars. This is a huge problem for potential uses of the network. It is currently a proof of work system like Bitcoin (this is described in the next section), and has a [huge energy footprint](#) to secure the network. It also ties up global supply of PC graphics cards used for it's mining model, making them far more expensive.

Ether in the Context of Equity Market Bubbles, Bitcoin and Tulips

The equity, tulip and Bitcoin bubbles are all dwarfed by the price moves in Ether.



Data through May 31, 2021.

Source: Investment Strategy Group, Bloomberg.

Figure 3.3: Ethereum is thought to look like a speculative bubble. Rights requested



Figure 3.4: The rate of token generation has changed unpredictably over time. Rights requested

3.1.2 Upgrade roadmap

Part of the challenge Ethereum faces is wrapped up with its complex token emission schedule. This is the rate at which tokens are generated and ‘burnt’ or destroyed in the network. The total supply of tokens is uncertain, and both emission and burn schedules are regularly tinkered with by the project. The changes to the rate at which ETH are generated can be seen in Figure 3.4. In addition, a recent upgrade (EIP-1559) results in tokens now being burnt at a higher rate than they are produced, deliberately leading to a diminishing supply. In theory this increases the value of each ETH on the network at around 3% per year. It’s very complex, with impacts on transaction Fees, waiting time, and consensus security, as examined by Liu et al. [22]. Additionally, there is now talk (by Butlerin, the creator of Ethereum) of extending this burn mechanism [further into the network](#).

Ethereum was designed from the beginning to move to a ‘proof of stake’ model where token holders underpin network consensus through complex automated voting systems based upon their token holding. This is now called [Ethereum Consensus Layer](#). Like much of the rest of ‘crypto’ the proposed changes will concentrate decisions and economic rewards in the hands of major players, early investors, and incumbents. This is a far cry from the stated aims of the technology. The move to proof of stake has recently earned it the [MIT breakthrough technology award](#), despite not being complete. It’s clearly a technology which is designed to innovate at the expense of predictability. This might work out very well for the platform, but right now the barrier to participation (in gas fees) is so high that we do not intend for Ethereum to be in scope as a method for value transfer within metaverses.

3.2 Bitcoin

The first practical blockchain was the Bitcoin network [18], some two decades after Haber et al. first described the idea [23]. It can be considered a triple entry book keeping system [24, 25], the first of it's kind, integrating a ‘provable’ timestamp with a transaction ledger. Some see this as the first major innovation in ledger technology since double entry was codified in Venice in fourteen seventy five[26].

It was created pseudonomously by an entity or group calling themselves ‘Satoshi Nakamoto’ in 2009, as a direct response to the perceived mishandling of the 2008 global financial crisis, with the stated aim of challenging the status quo, with an **uncensorable** technology, to create a money which could not be debased by inflation policy. There will only ever be (just short) of 21 million bitcoins issued, of which around 19 million have already been minted, and around 4 million lost forever. This ‘hard money’ absolute scarcity is a strong component of the Bitcoin meme landscape. These are basically arbitrary figures though; a combination of part of the issuance schedule, and a finger in the air guess by Nakamoto.

“I wanted to pick something that would make prices similar to existing currencies, but without knowing the future, that’s very hard. I ended up picking something in the middle. If Bitcoin remains a small niche, it’ll be worth less per unit than existing currencies. If you imagine it being used for some fraction of world commerce, then there’s only going to be 21 million coins for the whole world, so it would be worth much more per unit”

In theory there is no **barrier to access**, and **equality of opportunity** to accumulate and save over long periods. Each bitcoin can be divided into 100 million satoshis (sats), so anyone buying into bitcoin can buy a thousandth of a pound, assuming they can find someone willing to transact that with them.

The IMF has recently conceded that the Bitcoin **poses a risk** to the traditional financial systems, so it could be argued that it is succeeding in this original aim. The “**genesis block**” which was hard coded at the beginning of the ‘chain’ contains text from The Times newspaper detailing the second bank bailout.

Satoshi Nakamoto (the name of the publishing entity) **disappeared from the forums** forever in 2010.

Although there were some earlier experiments (hashcash, b-money etc), Bitcoin is the first viably decentralised ‘cryptocurrency’; the network is used to **store economic value** because it is judged to be secure and trusted. It is a singular event in that it became established at scale, such that it could be seen to be a fully distributed system, without a controlling entity. This is the differentiated trust model previously mentioned. This relative security is the specific unique selling point of the network. It is many times more secure than all the networks which came after based on a like for like comparison of **transaction ‘confirmations’**. This network effect of Bitcoin is a compounding feature, attracting value through the security of the system. It is deliberately more conservative and feature poor, preferring instead to **add to its feature set** slowly, preserving the integrity of the value invested in it over the last decade. At time of writing it is a **top quartile** largest global currency and has settled over \$12 trillion Dollars in 2021, though Makarov et al. contest this, citing network overheads, and speculation [27]. Institution grade ‘exchange tradable funds’ which allow investment in Bitcoin are available throughout the world, and the native asset can be bought by the public easily through apps in all but a handful of countries as seen in Figure 3.5.



Figure 3.5: [Growth in settlement](#) value on the Bitcoin network.

Only around 7 transactions per second can be settled on Bitcoin. The native protocol does not scale well, and this is an inherent trade-off as described by Croman et al. in their positioning paper on public blockchains [28]. Over time, competition for the limited transaction bandwidth drives up the price to use the network. This effectively prices out small transactions, even locking up some value below what is termed the '[dust limit](#)' of unspent transactions too small to ever move again [29].

Bitcoin has developed quickly, with a [faster adoption](#) than even the internet itself. It is now a mature ecosystem, and is seeing adoption as a [corporate treasury asset](#).

Adoption by civil authorities is increasing, and legislators the world over are being forced to [adopt a position](#). Many city treasuries have added it to their balance sheet. The Swiss city of Lugano is launching a [huge initiative](#) alongside Tether. It is already legal tender in the country of El Salvador[30]. This will be explored more later. Global asset manager "Fidelity" wrote the following in their [2021 trends report](#).

"We also think there is very high stakes game theory at play here, whereby if bitcoin adoption increases, the countries that secure some bitcoin today will be better off competitively than their peers. Therefore, even if other countries do not believe in the investment thesis or adoption of bitcoin, they will be forced to acquire some as a form of insurance. In other words, a small cost can be paid today as a hedge compared to a potentially much larger cost years in the future. We therefore wouldn't be surprised to see other sovereign nation states acquire bitcoin in 2022 and perhaps even see a central bank make an acquisition."

3.2.1 The Bitcoin Network Software

There isn't a single GitHub which can be considered the final arbiter of the development direction, because it is a distributed community effort with some [400 developers](#) out of a wider 'crypto' pool of around 9000 contributors. [Development and innovation continues](#) but there is an emphasis on careful iteration to avoid damage to the network. Visualisation of code commitments to the various open source software repositories can be seen at [Bitpaint youtube channel](#) and in Figure 3.6.

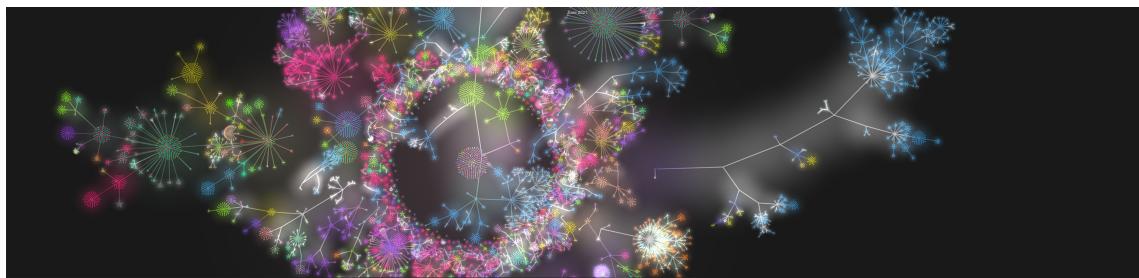


Figure 3.6: [Bitpaint](#): Contributions to the Bitcoin ecosystem. Reused with permission.

[Bitcoin core](#) is the main historical effort, but there are alternatives ([LibBitcoin in C++](#), [BTCD in Go](#), and [BitcoinJ in Java](#)).

3.2.2 What are ECDSA, addresses and UTXO's

All these technologies use the same underpinning elliptic curve cryptography, but it makes sense to unpack this here just once, and only in the context of Bitcoin, as this will be the main focus of our attention. To that end we will not consider solutions which use the "integer factorisation problem".

In Bitcoin the ECDSA algorithm is used on the `secp256k1` elliptic function to create a trapdoor. This (essentially one way mathematical operation) was originally the "discrete log problem" and part of the research in cryptography by Diffie and Hellman [7]. In this mathematical construct a modulus operator creates an infinite number of possible variations on operations which multiply enormous exponential numbers together, in different orders, to create key pairs. In order to reverse back through the 'trapdoor' a probably impossible number of guesses would have to be applied.

Latterly, elliptic curves such as the `secp256k1` curve used in Bitcoin have substantially simplified the computation problems. Rather than exponentials used by Diffie Hellman instead a repeated operation is applied to an elliptic curve function, and this itself creates a discrete log problem trapdoor in maths, far more efficiently. Figure 3.7 suggests how this works.

This makes it easier, faster, and cheaper to provide secure key pairs on basic computational resources. Elliptic curve solutions are not 'provably' secure in the same way as the Diffie-Hellman approach, and the security of this system is very sensitive to the randomness which is applied to the operation. ECDSA has already been replaced by the more efficient Schnorr signature method, but this will take so time for organic adoption, and ECDSA will never be deprecated.

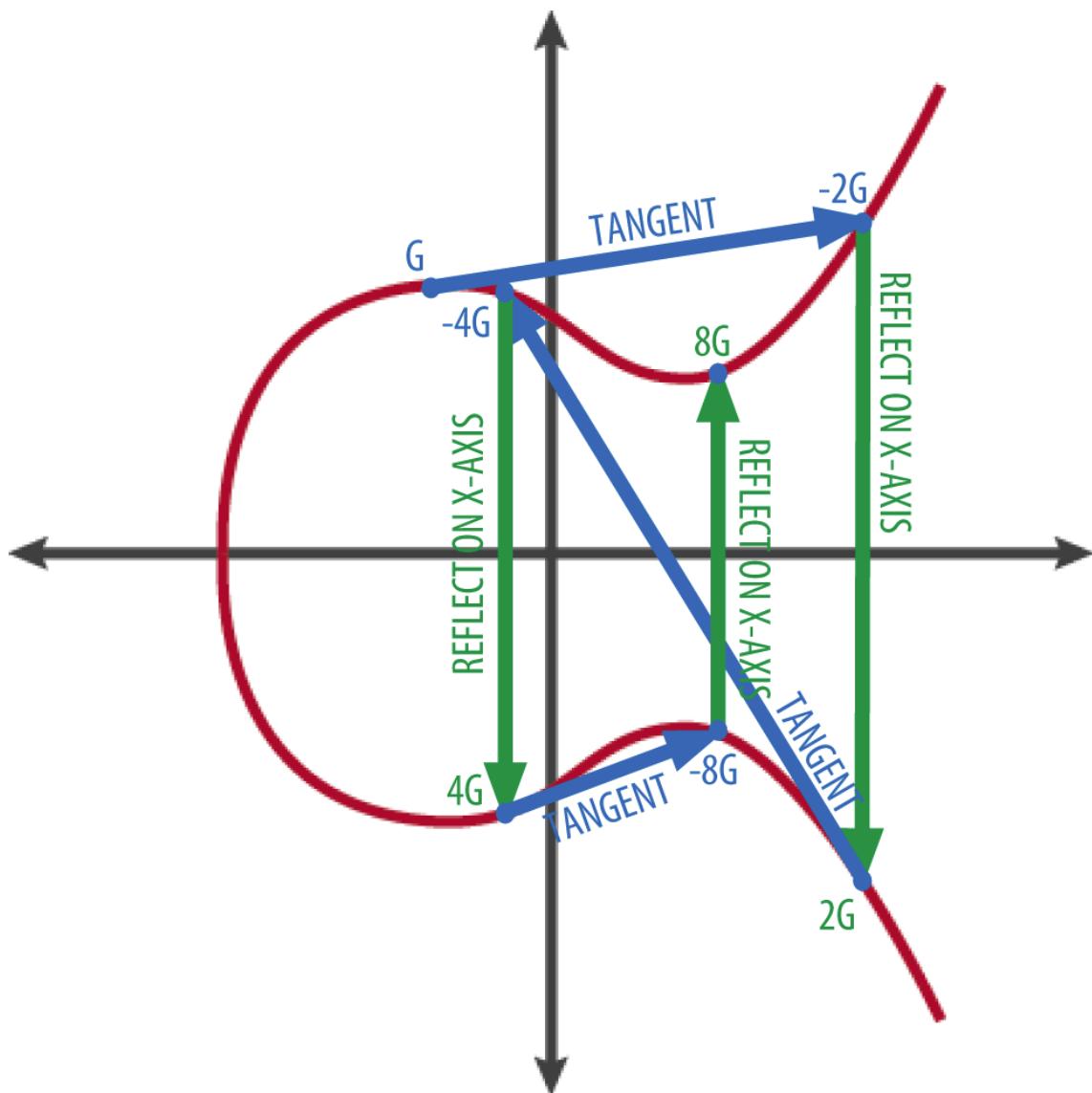


Figure 3.7: Given a start point on the curve and a number of reflection operations it's trivial to find a number at the end point, but impossible to find the number of hops from the two end points alone.

3.2.3 Mining and Energy concerns

Bitcoin mining is the process of adding public transactions into the ledger, in return for two economic rewards, paid in Bitcoin. These are the mining fee, and the block reward. The transactions which are added into the next 'block' of the chain are selected preferentially based on the fee they offer, which is up to the user trying to get their transaction into the chain. This can be within the next 10 minutes (next block), or a gamble out toward 'never' depending how competitive the network is at any time. Miners try to find a sufficiently low "magicnumber" resulting from a hash function, and upon finding it, they can take their pre-prepared 'block' of transactions sourced from their local queue (mempool), and add it into the chain, for confirmation by other miners. In return they take all the fees within that mined block, and whatever the block reward is at the time. When the network started the block reward was 50 Bitcoin, but has [halved](#) repeatedly every 210,000 blocks (four years) and now stands at 6.25 BTC. The rate of mining is kept roughly at one block every 10 minutes, by a difficulty adjustment every 2016 blocks (2 weeks). This is a complex interdependent mechanism and is explained very well in [this article](#).

Bitcoin uses a staggering amount of energy to secure the blockchain, and this [has climate repercussions](#). It is an industrial scale global business with 'mining companies' investing [hundreds of millions of pounds](#) at a time in specialist ASIC mining hardware and facilities. The latest purpose designed Intel chip [touts](#) both Web3 and metaverse applications. This is Adam Back's "proof of work", and is essential to the technology. [The Cambridge Bitcoin Energy Consumption Index](#) monitors this energy usage.

Such businesses can mine a Bitcoin for around \$5k-\$10k per coin, so the profit margins [are considerable](#) (based on 30-40 Joule/terahash and power rate less than 5 cents/kilowatt hour and excluding hardware costs). This is not to say that all mining is, or should be, so concentrated. Anyone running the hashing algorithm can [get lucky](#) and claim the block reward. PoW ties the value of the 'money' component of Bitcoin directly to energy production. This is not a new idea as can be seen in Figure 3.8. Henry Ford proposed an intimate tie between energy and money to create a separation of powers from government.

The potential ecological footprint of the network has always been a concern; Hal Finney himself was [thinking about this issue](#) with a mature Bitcoin network as early as 2009, and a [debate](#) on the Bitcoin mailing lists called the mining process "thermodynamically perverse"

Proponents of the technology say that the balance shifted dramatically in 2021 with China outright banning the technology; this has forced the bulk of the energy use away from 'dirty coal' as seen in Figure 3.9. Some analysts [propose mitigations](#), or even suggest that '[ending financialisation](#)' through use of Bitcoin may be net positive for the environment. Indeed it may [provide a route](#) to support electrifying everything through local subsidy initiatives [31].

Recently for example Baur and Oll found that "*Bitcoin investments can be less carbon intensive than standard equity investments and thus reduce the total carbon footprint of a portfolio.*"[32]. Perhaps of note for the near future is that KPMG whose investment was mentioned in the introduction also matched their position in the space with equivalent carbon offsets. This may provide an investment and growth model for others.

The power commitment to the network is variously projected [to increase](#), or [level off over time](#), but certainly not decrease. The industry now argues that economic pressures mean that most of the 'hashrate' is [generated by renewable energy](#)[33]. Certainly, there is growing interest and adoption of so called "stranded energy mining" which cannot be



Figure 3.8: Intimate tie between energy and money, Henry Ford

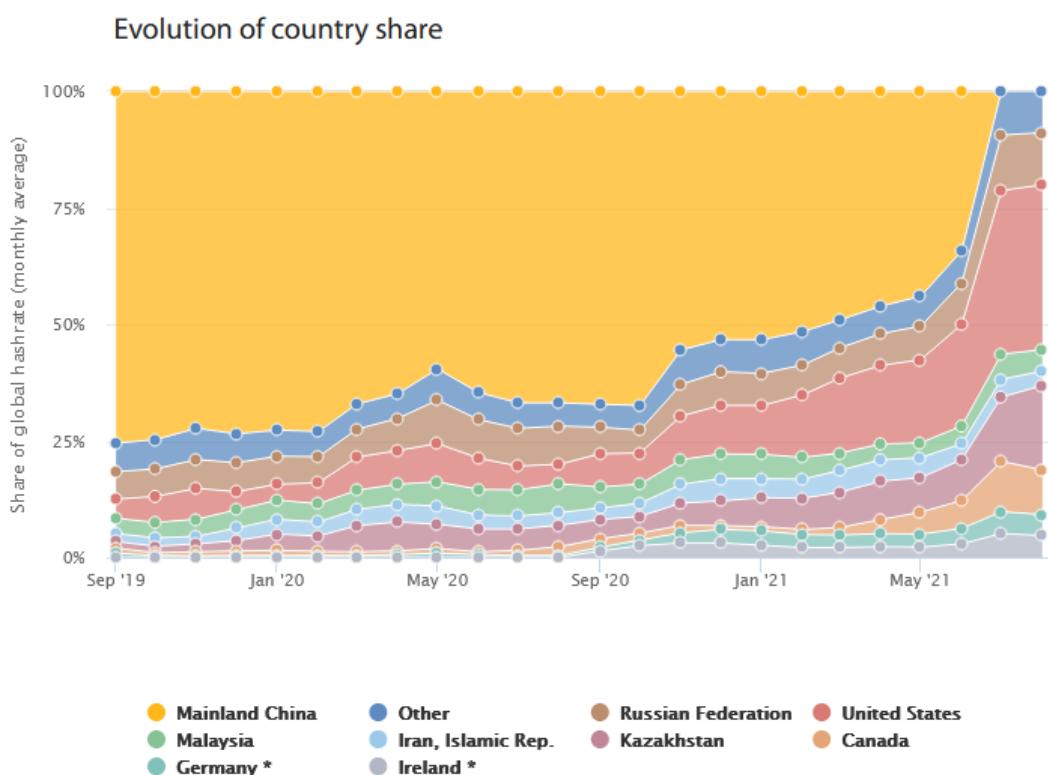


Figure 3.9: Hash rate suddenly migrates from China [Reuse rights requested]

effectively transmitted to consumers, and is thereby sold at a huge discount while also developing power capacity [34]. The most cited example of this is El Salvador's 'volcano mining' which is supporting their national power infrastructure plans. A more poignant example is the [Mechanicville hydro plant in the USA](#). The refurbishment of this 123 year old power plant is being funded by Bitcoin mining. This is the "buyer of last resort" model first advanced by [Square Inc](#). Critics highlight the potential impact of mining on local energy prices[35].

The debate whether this consumption is 'worth' it is complex and rapidly evolving. Is a trillion dollar asset which potentially replaces the money utility of gold, but doesn't need to be stored under guard in vaults (Figure 3.10), worth the equivalent power consumption of clothes dryers in North America? Probably not with the current level of adoption, but this is an experiment in replacing global money.

Legislators globally are starting to codify their positions on proof of work as a technology (including Bitcoin). The USA is variously supporting or constricting the technology, according to [state legislatures](#). Notably New York has submitted a bill to ban mining for 3 years, while rust and farm belt states with energy build-out problems are providing incentives.

The EU has just voted to add the whole of 'crypto', including PoW, to the EU taxonomy for sustainable activities. This EU wide classification system provides investors with guidance as to the sustainability of a given technology, and can have a meaningful impact on the flows of investment. With that said the report and addition of PoW is not slated until 2025, and it is by no means clear what the analysis will be by that point.

3.2.4 Bitcoin nodes

The Bitcoin network can be considered a triumvirate of economic actors, each with different incentives. These are:

- Holders and users of the tokens, including exchanges and market makers, who make money speculating, arbitraging, and providing liquidity into the network.
- Miners, who profit from creation of new UTXOs, and receive payments for adding transactions to the chain. In return they secure the network.
- Node operators, who enforce the consensus ruleset which the miners must abide by in order to propagate new transaction into the network. In return node operators optimise their trust minimisation, and help protect the network from changes which might undermine their speculation and use of the tokens.

There are currently around [15,000 bitcoin nodes](#) distributed across the world. Since IT engineer [Stadicus](#) released his [Raspibolt guide](#) in 2017 there has been an explosion of small scale Bitcoin and Lightning node operators. Around thirty thousand Raspberry Pi Lightning nodes (which are also by definition bitcoin nodes) run one of the following open source distributions, with the most noteworthy explained alongside their standout featureset:

- [Raspiblitz](#) offers fully opensource lightning focused functionality with a touchscreen display
- [Mynode](#) focuses on easy of use through a web interface and has many modules which users can try out.
- [Umbrel](#) is a more user friendly multi purpose node allowing access to a suite of Bitcoin and self sovereign individual tools.
- [RoninDojo](#) is designed for use alongside the privacy focused [Samourai mobile wallet](#).

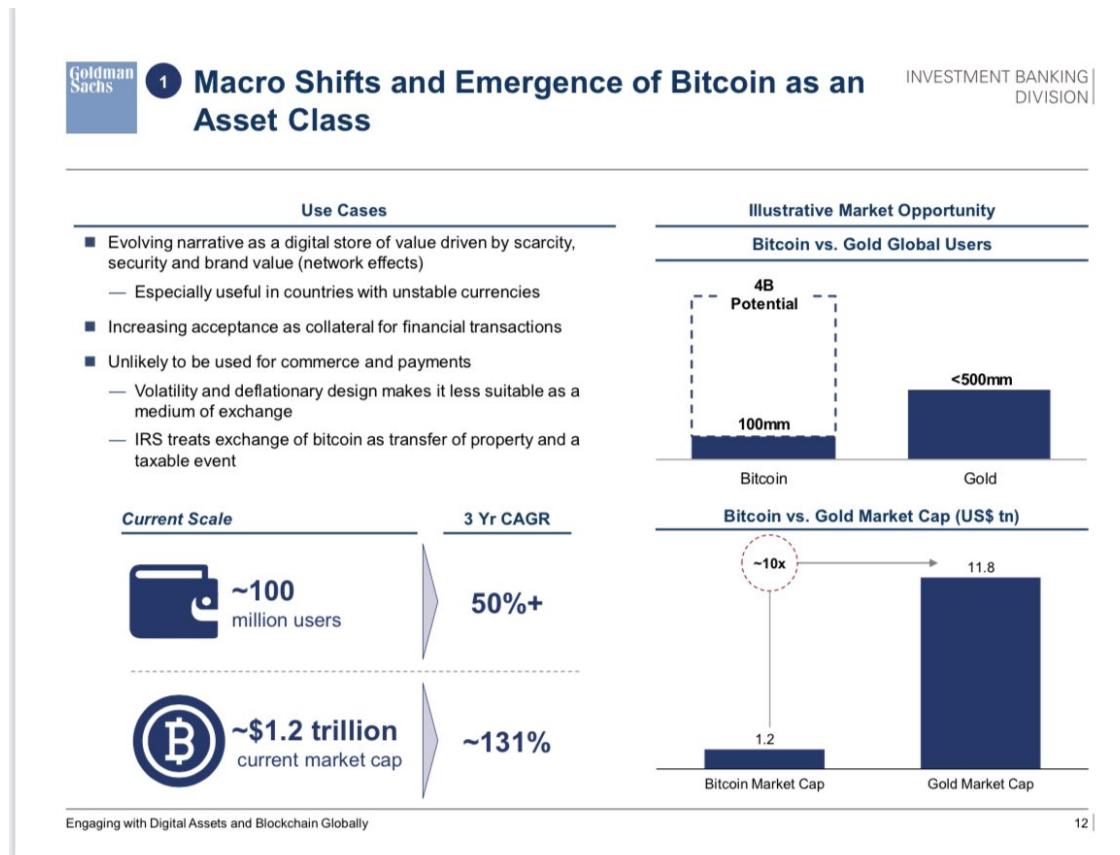


Figure 3.10: Goldman suggest growth opportunity and potential demonetisation of gold?

- [nix-bitcoin](#) focuses on security of the underlying operating system by building on NixOS.
- [NODL](#) is a premium prebuilt node focusing on security of the more performant hardware, and underlying operating system. It offers additional privacy tools.
- [Start9 Embassy](#) is a small form factor prebuilt unit at a lower price. It is a venture capital funded project with a more restrictive license but offers a suite of easy to use self sovereignty tools including Bitcoin.

We plan to support a new metaverse focused suite of tools that utilises the nix-bitcoin distribution for the Bitcoin node. [Consider mining the NYDIG primer for information](#)

3.2.5 Upgrade roadmap

3.2.5.1 Taproot

‘Taproot’ is the most recent upgrade to the Bitcoin network. It was first [described in 2018](#) on bitcoin-dev mailing list, and become [BIP-0341](#) in 2019. It brings improved scripting, smart contract capability, privacy, and schnorr signatures, which are a maximally efficient signature verification method. The network will always support older address types. It is rare to get such a large update to the network, and deployment and upgrade was carefully managed over several months under BIP-0008. Uptake will be slow as wallet manufacturers and exchanges add the feature. It can be considered an [upgrade in progress \(0.3%\)](#). Aaron van Wirdum, a journalist and educator in the space describes Taproot in detail in [an article](#).

3.2.5.2 AnyPrevOut

[BIP-0118](#), is a “soft-fork that allows a transaction to be signed without reference to any specific previous output”. It enables “Eltoo, a protocol that fulfils Satoshi’s vision for nSequence”

This is Lightning Network upgrade technology in the main. The Eltoo [whitepaper](#) or this more [readable explanation](#) from developer fiatjaf go into detail.

3.2.5.3 CheckTemplateVerify

[BIP-0119](#) is “a simple proposal to power the next wave of Bitcoin adoption and applications. The underlying technology is carefully engineered to be simple to understand, easy to use, and safe to deploy.”

3.3 Extending the BTC ecosystem

The following section are by no means an exhaustive view of development on the Bitcoin network, but it does highlight some potentially useful ideas for supporting metaverse interactions in a useful timeframe.

3.3.1 Block & SpiralBTC

Block (formally the payment processor “Square” is now an umbrella company for several smaller ‘building block’ companies, all of which are major players in the space.

SpiralBTC, formally ‘Square Crypto’ (a subsidiary of Square) is funding development in Bitcoin and Lightning. Their main internal product is the [Lightning Development Kit \(LDK\)](#). This promising open source library and API will allow developer to add lightning

functionality to apps and wallets. It is a useful contender for our metaverse applications. They also fund external open source development.

3.3.2 BTCPayServer

BTCPayServer is one of the recipients of a Spiral grant. It is a self hosted Bitcoin and Lightning payment processor system which allows merchants, online, and physical stores and businesses to integrate Bitcoin into their accounting systems.

[check the roadmap](#)

3.4 Lightning (Layer 2)

Lightning was a 2016 proposal by Poon and Dryja [36], and is a community driven liquidity pool which enables scaling and speed improvements for the Bitcoin network. It is mainly ‘powered’ by [thousands of volunteers](#) who invest in hardware and lock up their Bitcoin in their nodes, to facilitate peer to peer transactions. Zebka et al. found that although the network is “fairly decentralised” it is more recently skewing to larger more established nodes [37]. Though this is a grassroots technology the nature of the design means it can likely be trusted for small scale commercial applications.

The following text is from [John Cantrell](#), an engineer who works on Lightning.

“The Lightning Network is a p2p network of payment channels. A payment channel is a contract between two people where they commit funds using a single onchain tx. Once the funds are committed they can make an unlimited amount of instant & free payments over the channel. You can think of it as a tab where each person tracks how much money they are owed. Each time a payment is made over the channel both parties update their record of how much money each person has. These updates all happen off-chain and only the parties involved know about them. When it’s time to settle up the two parties can take the final balances of the channel and create a channel closing transaction that will be broadcast on chain. This closing transaction sends each party the final amounts they are owed. This means for the cost of two on-chain transactions (the opening and closing of the channel) two parties can transact an unlimited number of times and the overall cost of each transaction approaches zero with every additional transaction they make over the channel. Payment channels are a great solution for two parties to transact quickly and cheaply but what if we want to be able to send money to anyone in the world quickly and cheaply? This is where the Lightning Network comes into play, it’s a p2p network of these payment channels. This means if Alice has a payment channel with Bob and Bob has a channel with Charlie that Alice can send a payment to Charlie with Bob’s help. This idea can be extended such that you can route a payment over an arbitrary number of channels until you can reach the entire world. Routing a payment over multiple channels uses a specific contract called a Hash Time Locked Contract (HTLC). It introduces the ability for Bob and any other nodes you route through to charge a small fee. These fees are typically orders of magnitude smaller than onchain fees. This all sounds great but what if someone tries to cheat? I thought the whole point of Bitcoin was that we no longer had to trust anyone and it sure sounds like there must be some trust in our channel partners to use the Lightning Network? The contracts used in Lightning are built to prevent fraud while requiring no trust. There is a built-in penalty mechanism where if someone tries to cheat and is caught then they lose all of their money. This does mean you need to be monitoring the chain for fraud attempts.”

Lightning is a key scaling innovation in the bitcoin network at this time. It is seeing rapid development and adoption (Figure 3.11). The popular payment app “Cash App” integrates the technology, and ‘Lightning Strike’ services the USA, El Salvador, and Argentina with zero exchange and transmission fees.

It allows for unbound scaling of transactions (millions of transactions per second compared for instance to around 45,000 TPS in the VISA settlement network). Transaction costs are incredibly low, and the transaction speed virtually instantaneous.

The main Lightning network git is the Daemon [here](#) but it’s worth knowing that Lightning itself really needs access to both a Bitcoin full node, and the Tor private network layer.

3.4.1 Micropayments

Possibly the most important affordance of the Lightning network is the concept of micropayments, and streaming micropayments. It is very simple to transfer even [one satoshi](#) on Lightning, which is one hundred millionth of a bitcoin, and a small fraction of a penny. This can be a single payment, for a very small goods or service, or a recurring payment on any cadence. This enables streaming payments for any service, or for remittance, or remuneration. These use cases likely have enormous consequences which are just beginning to be explored. Integration of this capability into metaverse applications will be explored later.

3.4.2 BOLT12 and recurring payments

[BOLT12](#) is a new and developing ’standard’ which simplifies and extends the capability of the network for recurring payments.

3.4.3 LNBits

LNBits is an open source, extensible, Lightning ‘source’ management suite. It is self hosted, and can connect to a variety of Lightning wallets, further abstracting the liquidity to provide additional functionality to network users. Remember that all of these tools run without a third party, on a £200 setup, hosted at home or within a business. The best way to explore this is to describe *some* of the plugins.

- “[Accounts System](#); Create multiple accounts/wallets. Run for yourself, friends/family, or the whole world!”
- [Events plugin](#) allows QR code tickets to be created for an event, and for payments to be taken for the tickets.
- [Jukebox](#) creates a Spotify based jukebox which can be deployed online or in physical locations.
- [Livestream](#) provides an interface for online live DJ sets to receive real-time Lightning tips, which can be split automatically in real-time with the music producer.
- [TPoS](#), [LNURLPoS](#) & [OfflineShop](#) support online and offline point of sale (Figure 3.12).
- [Paywall](#) creates web access control for content.
- [LightningTipBot](#) is a custodial Lightning wallet and tip handling bot within the popular on Telegram instant messenger service.

Together these plugins are incredibly useful primitives which are likely to be translatable to a multi party metaverse application. A proposal for building a more specific plugin



Figure 3.11: Arcane research lightning adoption overview.



Figure 3.12: Two of the many [prebuilt](#) and [kit](#) options for Lightning ‘point of sale’

along these lines is detailed later.

3.4.4 Etleneum

Etleneum is a centralised smart contract platform built around Lightning invoices. It is most notable as a sign of things to come. There are [many small contracts](#) available to try on the site, such as a [simple market](#) for moving value between lightning and Bitcoin layer 1, or this [simple auction](#). Contracts are able to operate on data drawn from the wider web, and automatically send and receive lightning payments based on conditional states. It should be viewed as an experiment which allows tinkering in smart contracts, and therefore potentially useful for the software proposed in the final section.

3.4.5 Message passing

It is possible to pass data alongside lightning payments, routing messages between parties across the global network. This means that a host of other applications can inherit the privacy and censorship resistance of the Lightning network. First amongst these has been simple message passing and group messenger clients such as Sphinx and Juggernaut. To be clear, this is considered by some to be a misappropriation of the function of the network. Once more developed use case has been demonstrated by the Impervious development team; they use the message passing capability to negotiate a virtual private network between two parties, using open source software. This allows a secure side channel between internet IP addresses to be opened without a trusted third party. This in itself is a much sought after function of privacy minded networking, and the basis for much of their Impervious browser feature set.

3.5 Liquid federation (layer 2)

Liquid is part of the open source development contribution of Blockstream, the company started by Adam Back (of proof of work fame). The Liquid side chain network, and it's own attendant Lightning layer 2, is a fork of bitcoin with different network parameters. In liquid the user of the network ‘pegs’ into the Bitcoin network, swapping tokens out from BTC to L-BTC (this can of course mean very small subunits of 1 Bitcoin). Once tokens have been ‘locked’ and swapped to Liquid the different network parameters used in the fork allow a different trust/performance trade-off. Liquid is fast on the L1 chain, cheaper to use at this time, and more private. The consensus achieved on this side chain network is faster because it is a far smaller group of node operators. The next block to be written to the side chain is chosen by a node operated by a member of a federation of dozens of major contributors to the Bitcoin technology space. These ‘trusted’ nodes all check one another’s security and network operations, meaning that the network is as secure as the aggregate of the trust placed in half of the membership at any one time. This is [still dozens](#) of major companies, development teams, and individual actors, with significant reputational investment.

“Federation members contribute to the Liquid Network’s security, gain voting rights in the board election and membership process, and provide valuable input on the development of new features. Members also benefit from the ability to perform a peg-out without a third party, allowing their users to convert between L-BTC and BTC seamlessly within their platform.”

Crucially for our purposes here Liquid allows tokenised asset transfer. Anyone [can](#)

[issue](#) an asset on Liquid. Such transfers of assets may be orders of magnitude cheaper than on chain Bitcoin transactions, but still potentially orders of magnitude more expensive than a simple Lightning transaction of value on the Bitcoin network.

Blockstream plan to add arbitrary (user generated) token support to their C-Lightning implementation at some point. This would be a very strong choice for specific use cases within an economically enabled metaverse application. When participants wish to ‘cash out’ of the Liquid network they must do this through one of the federation members who activate the other side of the ‘two-way peg’, dispensing the equivalent amount of Bitcoin. This is transparently handled through Blockstream’s “green wallet”.

All of this has the advantage of a far lower energy footprint compared to the main chain, but it’s not quite ready with a full suite of affordances.

The Liquid network is being used as the underlying asset for a novel new global financial product. El Salvador are working with Blockstream to issue a nation state backed bond.

[El Salvador volcano bonds stuff launching in March 20th](#)

3.6 Bitcoin Layer 3

Increasingly important features of modern blockchain implementations are programmability through smart contracts, and issuance of arbitrary tokens. Assigning a transaction to represent another thing like an economic unit, energy unit, or real world object, and operating on those abstractions within the chain logic. Chief among these use cases are stablecoins such as Tether, which are pegged to national currencies and described in the next section. Bitcoin has always supported very limited contracts called scripts, and stablecoin issuance has existed in Bitcoin since [Omni Layer](#). Omni was the first issuer of Tether, but more recently these important features have passed to other layer one chains. This year is likely to see the [resurgence of this capability](#) on Bitcoin, which of course benefits from a better security model. In order to properly understand the use of Bitcoin based technologies in metaverse applications it is necessary to examine what these newer ‘layer 3’ ideas bring.

3.6.1 LNP/BP and RGB

LNP/BP is a non profit standards organisation in Switzerland which contributes to open source development of Bitcoin layer 3 solutions into the Lightning protocol, and Bitcoin protocol (LNP/BP). One of the core product developments within their work is the ‘RGB’ protocol, which is somewhat of a meaningless name, evolved from “coloured coins” which were an early tokenised asset system on the Bitcoin network. RGB represents red, green, and blue. The proposal is built upon research by [Todd](#) and [Zucco](#). RGB is regarded as arcane Bitcoin technology, even within the already rarefied Bitcoin developer communities. Zucco provides the [following explanation](#):

“When I want to send you a bitcoin, I will sign the transaction, I will give the transaction only to you, you will be the only one verifying, and then we’ll take a commitment to this transaction and that I will give only the commitment to miners. Miners will basically build a blockchain of commitments, but without the actual validation part. That will be only left to you. And when you want to send the assets to somebody else, you will pass your signature, plus my signature, plus the previous signature, and so on.”

This is non-intuitive explanation of Todds ‘single-use-seals’, applied to Bitcoin, with the purpose of underpinning arbitrary asset transfer secured by the Bitcoin network. In this model the transacting parties are the exclusive holders of the information about what the object they are transferring actually represents. This primitive can (and has) been expanded by the LNP/BP group into a concept called ‘client side validation’. It’s appropriate to explain this concept several times from different perspectives, because this is potentially a profoundly useful technology for metaverse applications.

- A promise is made to spend a multi output transaction in the future. This establishes the RGB relationships between the parties.
- One of the pubkeys to be spent to is known by both parties.
- The second output is unknown and is a combination of the hash of the state, and schema, from the operation which has been performed.
- When the UTXO is spent the second spends pubkey can be processed against the shared data blob to validate the shared state in a two party consensus
- This is now tethered to the main chain. Some tokens from the issuance have gone to the recipient, and the remainder have gone back to the issuer. More tokens can be issued in the same way from this pool.
- A token schema in the blob will show the agreed issuance and the history back to the genesis for the token holder.
- The data blob contains the schema which is the key to RGB functions and the bulk of the work and innovation.
- Each issuance must be verified on chain by the receiving party.

This leverages the single-use-seal concept to add in smart contracts, and more advanced concepts to Bitcoin. Crucially, this is not conceptually the same as the highly expressive ‘layer one’ chains which offer this functionality within their chain logic. In those systems there is a globally available shared consensus of ‘state’. In the LNP/BP technologies the state data is owned, controlled, and stored by the transacting parties. Bitcoin provides the cryptographic external proof of a state change in the event of a proof being required. This is an elegant solution in that it takes up virtually no space on the blockchain, is private by design, and is extensible to layer 2 protocols like Lightning.

This expanding ecosystem of client side verified proposals is as follows:

- RGB smart contracts
- RGB assets are fungible tokens on Bitcoin L1 and L2, and non fungible Bitcoin L1 (and somewhat on L2).
- Bifrost is an [extension](#) to the Lightning protocol, with its own Rust based node implementation, and backwards compatibility with other nodes in the network. This means it can transparently participate in normal Lightning routing behaviour with other peers. Crucially however it can also negotiate passing the additional data for token transfer between two or more contiguous Bifrost enabled parties. This can be considered an additional network liquidity problem on top of Lightning, and is the essence of the “Layer 3” moniker associated with LNP/BP. It will require a great number of such nodes to successfully launch token transfer on Lightning. As a byproduct of its more ‘protocol’ minded design decisions Bifrost can also act as a generic peer to peer data network, enabling features like Storm file storage and Prometheus.
- [AluVM](#) is a RISC based virtual machine (programmable strictly in assembly) which can execute Turing complete complex logic, but only outputs a boolean result which

is compliant with the rest of the client side validation system. In this way a true or false can be returned into Bitcoin based logic, but be arbitrarily complex within the execution by the contract parties.

- Contractum is the proposed smart contract language which will compile the RGB20 contracts within AluVM (or other client side VMs) to provide accessible layer 3 smart contracts on Bitcoin. It is a very early proposal at this stage.
- Internet2: “Tor/noise-protocol Internet apps based on Lightning secure messaging
- Storm is a lightly specified escrow-based bitcoin data storage layer compliant with Lightning through Bifrost.
- Prometheus is a lightly specified multiparty high-load computing framework.

Really, any compute problem can be considered applicable to client side validation. In simplest terms a conventional computational problem is solved, and the cryptographically verifiable proof of this action, is made available to the stakeholders, on the Bitcoin ledger.

Less prosaically, at this stage of the project the more imminent proposed affordances of LNP/BP are described in ‘schema’ [on the project github](#). The most interesting to the technically minded layperson are:

- [RGB20](#) fungible assets. This could be stablecoins like dollar or pounds representation. This is a huge application area for Bitcoin, and similar to Omni, which will also be covered next.
- [RGB21](#) for nonfungible tokens and ownership rights. In principle BiFrost allows these to be transferred over a future version of the Lightning network, significantly lowering the barrier to entry for this whole technology. This is slated for release later in the year.
- [RGB22](#) may provide a route to identity proofs. This is covered in detail later.

3.6.2 Synonym & Omnidolt

Omnidolt [github](#)

3.6.3 DLCFD

Discrete log contracts are a form of externally arbitrated smart contract. Work is being done to extend this primitive to lightning.

Marty Bent “This particular contract allows (currently) one party to lock in a stable amount of USD value in BTC to avoid bitcoin price volatility while their counterpart takes a long position on BTC. As bitcoin’s price moves, the contract allocates sats to one party to make sure the individual who is engaged in the contract to lock in a USD value of bitcoin is doing just that. If the price of bitcoin goes up, sats are allocated to the individual who is taking the long position. If the price of bitcoin goes down, sats are allocated to the party looking to lock in a stable USD value. All of this happens on the Bitcoin blockchain and the Lightning Network. No obscure governance token, DAO, or central third party holding USD in a bank account necessary. All that is needed are sats and a willing counter-party.”

3.7 Bitcoin adjacent chains

In order of preference for functionality for Metaverse the following framework can be adopted. BitcoinL2>BitcoinL1>LiquidL2>Adjacentchains>Altchains

The reasons are:

- Free and open source network

- Network effect
- Security model
- Development community
- Transaction speed

This section is in early development.

3.7.1 Stacks and STX

“Stacks is an open-source network of decentralized apps and smart contracts built on Bitcoin.”

This novel approach saw the launch of a layer 1 blockchain token called STX, which is used in a similar way to gas in Ethereum, but claims settlement on the Bitcoin network. This is achieved through a novel bridging approach which they call Proof of Transfer (PoX).

Stacks users say this hybrid approach is a pragmatic solution which enables dApps, smart contracts, DeFi, NFTs etc without compromising security. In practice the speculative component of the STX tokens which underpin these operations clouds the issue somewhat. It is a potentially useful middle ground solution with a great deal of developer attention.

3.7.1.1 Citycoins

These are actually slightly interesting.

3.7.2 Sovryn and RSK

Ingamar | Top of the Block: Lightning <> Rootstock Bridge for lnBTC <> Stablecoin swaps

- [Exchange](#)
- [Repo 1](#)
- [Repo 2](#)
- [Repo 3](#)
- [Bridge aggregator](#)
- [Rest API docs](#)

3.7.3 Sidechains

3.7.4 Spacechains

Spacechains is a [proposal](#) by Ruben Somsen. It is a way to provide the functionality of any conceivable blockchain, by making it a sidechain to Bitcoin.

Like RGB described earlier it's a single use seal, but which can be closed by the highest bidder.

In a spacechain the Bitcoin tokens are destroyed in order to provably create the new spacechains tokens at a 1:1 value. These new tokens only have worth moving forward within the new chain ecosystem they represent, as they cannot be changed back. They nonetheless have the same security guarantees as the bitcoin main chain, though with a radically reduced ecological footprint (x1000?), and higher performance. Each ‘block’ in the new chain is a single bitcoin transaction.

- Outsource mining to BTC with only a single tx per block on the main chain.
- One way peg, Bitcoin in burnt to create spacechain tokens.
- Allows permissionless chain creation, without a speculative asset.

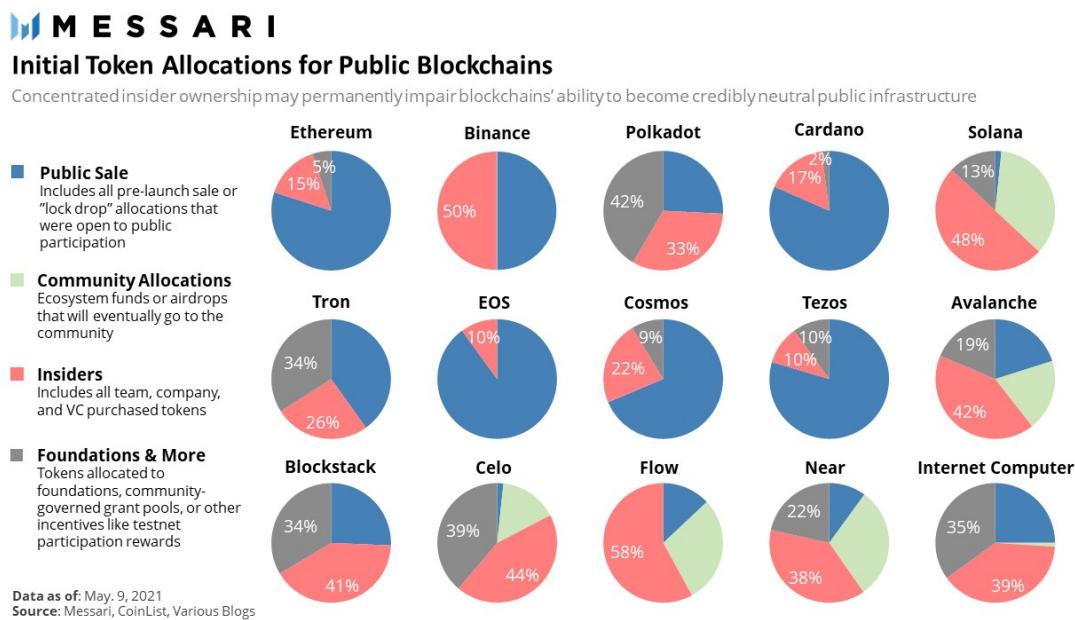


Figure 3.13: Allocations given at the beginning of public blockchain, by Messari.

- Fee bidding BMM is space efficient and incentive compatible. Miner just take the highest fees as normal.
- Paul Sztorc raised the idea, but it's best with a soft fork

3.7.5 Drivechain

3.7.6 Softchains

3.7.7 Statechains

There are many ??proposals for layer 2 scaling solutions for the bitcoin network. Ruben Somsen ??describes Softchains, Stateschains, and Spacechains, while [Drivechain is described](#) by the author Paul Sztorc on the project web pages. They are all hypothetical with the exception of sidechains.

3.8 Other chains and networks

It's useful to make some 'honourable mentions' of other options as this technology is moving so fast.

Figure 3.13 shows the allocations of proportions of value within different chains.

3.8.1 Layer 1 chains

- Solana is a far more centralised layer 1 proposition which uses a few hundred highly performant nodes to achieve high transaction throughput. The consensus algorithm is the novel “[proof of history](#)” system. Development of the technology has been funded and supported by huge venture capital investment, and even though the chain is quite unreliable it seems that the vested interests of the investors can keep interest going. It is cheaper, and more useful than Bitcoin and Ethereum, but lacks longevity and reliability .

- Polkadot There is much hype within the invested public blockchain community about “crosschain” protocols which can connect the logic of a smart contract on one chain to that of another. In practice, while it is possible that this is useful for distributed finance products, it seems that chains such as DOT might be promising more than the markets actually want or need. Governance of the token is a DAO like model where staking (locking up) the tokens theoretically controls the direction of the product.
- Luna is a relatively new [stablecoin offering](#) with DeFi built in. It is currently in ascendance and seems to have a stable and useful underpinning, but is far too new to judge with any certainty.
- Avalanche AVAX is a newer, ‘faster’ and more eco friendly DeFi ecosystem which promises returns within it’s own framework of permissionless money. It is one of the relative success stories of the DeFi narrative. It’s unclear what the value proposition, and sustainability of this token actually are.
- Tezos is an well established player with an early and somewhat battle tested proof of stake mechanism and distributed governance model. It has attracted many high profile partnerships and sponsors, but is primarily seeking to be a store of value token like bitcoin, which exposes the chain to the “winner takes all” landscape of digital money. There are some compelling NFT advocates of the technology, which is certainly ‘greener’ and cheaper to use, but the longevity in such an irrational market is uncertainly because it does not seem to have the network effect and growth velocity.
- Algorand’s ALGO token purports to be a more modern and useful proof of stake value transfer chain. It is fundamentally similar to Tezos.
- IOTA is noteworthy, interesting, and established concept, with an edge use case. It is the ‘distributed ledger of the internet of things’, the much hyped and clearly extant ecosystem of edge compute, sensors, smart devices etc. The marketing around IOTA correctly identifies it’s positioning and potential within this developing technology ecosystem, but it’s primary use case is too nascent and too niche to discuss in the same basket as the other ideas.
- VeChain is a long established platform with significant industry adoption which still doesn’t represent it’s market capitalisation, usefulness, or future success. It is the most exposed of all the chains to the assertion that immutable global ledgers of real work asset tracking somehow protect from fraudulent behaviour. It’s perhaps useful, but mainly in a highly automated industry 4.0 environment of minimal human interaction. it’s also unclear what is added
- Cardano foundation ADA is one of the more established players and has been developing methodically and slowly. They have made great strides in successfully enabling a provable proof of stake consensus structure. Proof of stake nonetheless has significant problems in that tokens and therefore control inevitably concentrates over time. There is no proposed solution to this. They have working products and partnerships, but perhaps not as many as the market cap of the ecosystem would suggest.
- HBAR claims “third generation” blockchain technology, with carbon positive, high speed, distributed applications. There are always tradeoffs bound by physical constraints within distributed computing, and Hedera HBAR has been accused of a cryptographic model which is inherently insecure.

- EOS is one of the early major successes of the ICO funding model in 2017, and they amassed an enormous war chest of bitcoin which they still hold. The onus is on them to deliver some kind of product, and with a multi billion dollar investment in their own promises they can likely deliver something.

3.8.2 Crosschains

RUNE? lots of hacks and unmanageable complex logic. Who is moving across chains anyway?

3.8.3 Decentralised data

Nick Grossman [blog piece](#) on adversarial open data might inform how we should specify the open data layer for the metaverse. Something about Filecoin? ARWeave? IPFS etc

3.9 Risks and mitigations

3.9.1 Digital assets

For digital assets more generally it is useful to look at the recent “[whole government executive order](#)” signed by President Biden. It is mainly framed in terms of “responsible innovation, and leadership” in the new space, is a product of multi agency collaboration, and has been long anticipated. It identifies high level risks, aspirations, and challenges, and strongly hints toward development of a “digital dollar” (CDBC, expanded later). The risks sections show how legislators are framing this, so it’s useful to break down here.

- Consumer and business protections. This is likely to pertain to custodians and is much needed. Miss-selling is rife. Security presents a challenge.
- Systemic risk, and market integrity are a concern. The legislators clearly worry about contagion risks from the sector.
- Illicit finance (criminality and sanction busting etc) are a concern, but not particularly front and centre. Criminality in 2021 was a mere 0.15% of transactions according to Chainalysis, but this number varies year to year. The US treasury department has recently published a National Risk Assessments for Money Laundering, Terrorist Financing, and Proliferation Financing. This is a comprehensive report and speaks to careful research across the space. It is broken into [three parts](#). Perhaps surprisingly, while they do see activity in these areas, they do not rate the risk as very significant. Cash remains the main problem for illicit funding. There is some talk that the nature of public blockchain analysis allows greater oversight of these tools and that this is to the advantage of government and civil enforcement agencies.
- Highlighting the need for international coordination suggests they are mindful of jurisdictional arbitrage. The partial regulatory capture of these technologies, where activity flows to globally more lenient legislative regimes, continues to be a concern. Many of the centralised exchanges for instance are located in tax havens such as Malta. As the world catches up with these products it is likely that this will be smoothed out.
- Climate goals, diversity, equality and inclusion are mentioned. It seems that the “environment” aspect of ESG is more important than “social” and “governance” at this time.
- Privacy and human rights are mentioned.

- Energy policy is highlighted, including grid management and reliability, energy efficiency incentives and standards, and sources of energy supply.

3.9.2 Bitcoin specifically

In addition it's useful for this document to focus more on the technical challenges to the Bitcoin network.

- The block reward is reduced every 4 years (epochs). This means a portion of the mining reward is trending to zero, and nobody knows what effect this will have on the incentives for securing the network through proof of work.
- Stablecoins are a vital transitional technology (described later) but do not meaningfully exist yet on the Bitcoin network. This may change.
- Bitcoin lacks privacy by design. All transactions are publicly viewable. This is a major drag to the concept of BTC as a money.
- The Lightning network (described later) has terrible UX design at this time.
- The basic ‘usability’ of the network is still poor in the main. Any problems which users experience demand a steep learning curve and risk loss of funds. There is obviously no technical support number people can call.
- Only around one billion unspent transactions can be generated a year on the network. This means that it might become impossible for everyone on the planet to have their own Bitcoin address (with its associated underpinning UTXO).
- Chip manufacture is concentrated in only a few companies and countries, as identified by Matthew Pines. He also identifies the [following points](#)
- Potential constraints on monetary policy flexibility.
- Future protocol changes.
- Unanticipated effects on the domestic and international energy system.
- Vulnerability to adversary attacks.
- Other unknown, unanticipated risks given Bitcoin’s limited 13-year history.



4. Money in the real world

It is necessary here to briefly examine what money actually is. In the previous section Bitcoin can be viewed in a couple of different lights. As a self custody digital bearer asset it can be viewed as ‘property’, like gold. Indeed this has long been one of the assertions of the community and it finds favour in law. ‘Money’ though is a far more [slippery concept](#) to grasp. It seems very likely that Bitcoin is evolving as a money, and it’s important to define that, but there are many other kinds of money within the online world which can potentially transfer value within virtual social spaces.

4.1 Defining money

It is hard to find a universally accepted definition of what money is. The best approach is to look at the properties of a thing which is asserted to be a money. In his book ‘A history of money’, Glyn Davies identifies “cognisability, utility, portability, divisibility, indestructibility, stability of value, and homogeneity” [38].

Stroukal examines Bitcoins’ likely value as a money from an Austrian economics perspective and identifies “portability, storability, divisibility, recognizability, homogeneity and scarcity” [39].

A helpfully brief and useful [web page by Desjardins from 2015](#) describes some properties and explains them in layman’s terms below:

- Divisible: Can be divided into smaller units of value.
- Fungible: One unit is viewed as interchangeable with another.
- Portable: Individuals can carry money with them and transfer it to others.
- Durable: An item must be able to withstand being used repeatedly.
- Acceptable: Everyone must be able to use the money for transactions.
- Uniform: All versions of the same denomination must have the same purchasing power.
- Limited in Supply: The supply of money in circulation ensures values remain relatively constant.

4.1.1 Global categories of capital

The legacy moniker “third world” came from a division of the world along economic lines [40]. At the time this was the petrodollar / neo-institutional hegemony [41, 42], vs the economic superpower of the soviet block, and then ‘the rest’; unaligned economic powers.

This old framework has fallen away with the associated terminology, but it’s useful to look at what money ‘is’ from a global viewpoint, because all money is effectively trust in the liability held by some defined counter party.

Right now the dollar system is still predominant, but it seems likely that there are new axes forming, especially around the Chinese Yuan. Use of the dollar system has recently been shown more and more to be contingent on adherence to US defined political principles. This is evidenced most starkly by the seizure of Russian central bank [foreign reserves](#), a new and untried projection of monetary power.

The Chinese Yuan/Renminbi is potentially stepping in where the petrodollar is now waning [43]. The effects of this expansion of economic influence by China, through a potential petro-Yuan, and the belt and road initiative [44], are not yet felt, but the lines are fairly clearly defined. The Euro system is potentially more stable, and less ‘weaponised’ but comes with its own restrictions for use, especially through the IMF. It is notable for instance that the IMF have included a clause in their negotiations with Argentina to ‘discourage’ the use of crypto based money. This is likely a response to the adoption of Bitcoin by El Salvador, something which the IMF is very uncomfortable. The remaining globally interesting monetary system might be Bitcoin (which is best positioned and tested for this role). This is evidenced in the early nation state adoption seen and described to date. It’s too early to tell if this ‘unaligned money’ could constitute a global economic pole, but it’s interesting that some commentators are now even discussing this.

4.2 International money transfer networks

Transferring money from one financial jurisdiction to another is itself a global marketplace which has accreted over the entire course of human history. It’s far less useful here to discuss the mythos of salt and seashells as a mechanisms of international remittance and taxation [45, 46]. Suffice it to say that there are dozens, if not hundreds, of cross border payment companies who make their business from taking a percentage cut of an international money transfer. There are also hundreds not thousands of banks who offer this service as part of their core business portfolio. This section looks at some of the major players, and their mechanism, to contextualise the more recent shifts brought about by technology.

4.2.1 Swift, ISO 20022, and correspondence banking

Society for Worldwide Interbank Financial Communications (SWIFT) was initially formed in 1973 between 239 banks across 15 countries. They needed a way to improve handling of cross border payments. It is now the global [standard](#) for financial message exchange in over 200 countries, and has recently found itself under a fresh spotlight, during the invasion of Ukraine. The system handles around 40 million short, secure, code transmissions a day, which represent crucial data about a transaction and the parties involved. It is used by both banks and major financial institutions to speed up settlement between themselves, on behalf of the clients and customers. It replaced the Telex (wire transfer) system. The new proposed and incoming standard to replace SWIFT is [ISO20022](#) which is a complex

and data rich arrangement. To be clear the SWIFT consortium are promoting this new standard to their 11,000 plus global user base, and there is significant investment and hype from major financial players, but it seems unclear what the actual take-up will or even should be. A group of ‘cryptocurrencies’ are heavily involved in the ISO20022 standard, and there’s been experimentation with private permissioned distributed ledger technologies. It’s actually somewhat unclear what value they bring, and possible that the relationship of these public ledgers to international bank to bank messaging is a marketing distraction. Note that SWIFT, ISO20022, and the associated tokens within crypto are all themselves products which have a business model. They are all intermediaries which will demand a mediating fee somewhere. All of this proposed functionality could be replaced by central bank digital currencies, which will be discussed later in the section.

4.2.2 VISA etc

VISA have announced a “crypto business to business support unit”.

4.2.3 Money transfer operators

[International Money Transfer Operators analysis](#)

western union etc, moneygram, transferwise,

4.2.4 Digital disruptive fintech

It seems that the neobank providers of digital banking apps are likely to converge with native digital asset “wallets”. This is also the thesis advanced by the Ark investments Big Ideas paper.

Strike is a possible the most interesting product in the international fintech arena. It is a ‘global’ money transmitter which uses bank connections in local currencies, but a private version of the Lightning network with settlement on the Bitcoin main chain. In practice users connect the app to their bank and can send money to the bank connected Strike app of another user instantly, and without a fee. This is a far better product than those previously available. In principle it’s open API allows many more application to be integrated into the Strike back end. Twitter already uses this for international tipping (and remittance). It seems that this is a perfect contender for supporting transactions in open metaverse applications, and that may be true, but Strike is currently only available in three countries (USA, El Salvador, Argentina).

Paypal, xoom, Strike, servicing smaller payments, cashapp, venmo, revolut,

4.2.5 Stablecoins

Stablecoins are ‘crypto like’ instruments which are ‘pegged’ at a 1:1 ratio with nationally issued Fiat currencies. In fact they correspond to units of privately issued debt underwritten by a variety of different assets. This is (depending on the issuing company’s model) a far more risky unit of money than the nominal currency that they represent, but they offer significant utility. They allow the user to self custody the cryptographic bearer instrument representing the money themselves, as with blockchain. This may afford the user less friction in that they can transmit the instrument through the newer financial rails which are emerging. The caveat here is that such ‘units’ of money can be frozen by the issuer, and they are subject to the third party risk of the issuer defaulting on the underlying instrument.

It seems like the primary intersection Current state of the art [Libre](#), Mark Carney

synthetic hedgemonic currency, and Keynes baskets

[USDF bank issued private dollar stablecoin](#)

[Paypal](#)

Whatsapp, Novi, USDP etc Crypto dollarisaton (myanmar) [USDC on Bitcoin?](#)

4.3 Central bank digital currencies

If 2021 was the year of the stablecoin then 2022 is likely to be the year of the central bank digital currency (CBDC). CBDCs would likely not exist without the [pressure exerted on central banks](#) by the concept of Bitcoin, and the stablecoins which emerged from the technology.

It now seems plausible that the world is moving toward a plurality of national and private currencies. This text from the [thinktank VoxEU](#) highlights the pressure on central banks not to be ‘left behind’:

“Given the rapid pace of innovations in payments technology and the proliferation of virtual currencies such as bitcoin and ethereum, it might not be prudent for central banks to be passive in their approach to CBDC. If the central bank does not produce any form of digital currency, there is a risk that it loses monetary control, with greater potential for severe economic downturns. With this in mind, central banks are moving expeditiously when they consider the adoption of CBDC.”

CBDCs are wholly digital representations of national currencies, and as such are centralised database entries, endorsed and potentially issued by national governments. It is a rapidly evolving space, and many nations are now scrambling to [catch up](#).

The following text is taken from the March 2021 Biden government “executive order” on digital assets, and defines the current global legislative position well.

“Sec. 4. Policy and Actions Related to United States Central Bank Digital Currencies. (a) The policy of my Administration on a United States CBDC is as follows:

(i) Sovereign money is at the core of a well-functioning financial system, macroeconomic stabilization policies, and economic growth. My Administration places the highest urgency on research and development efforts into the potential design and deployment options of a United States CBDC. These efforts should include assessments of possible benefits and risks for consumers, investors, and businesses; financial stability and systemic risk; payment systems; national security; the ability to exercise human rights; financial inclusion and equity; and the actions required to launch a United States CBDC if doing so is deemed to be in the national interest.

(ii) My Administration sees merit in showcasing United States leadership and participation in international fora related to CBDCs and in multi-country conversations and pilot projects involving CBDCs. Any future dollar payment system should be designed in a way that is consistent with United States priorities (as outlined in section 4(a)(i) of this order) and democratic values, including privacy protections, and that ensures the global financial system has appropriate transparency, connectivity, and platform and architecture interoperability or transferability, as appropriate.

(iii) A United States CBDC may have the potential to support efficient and low-cost transactions, particularly for cross-border funds transfers and payments, and to foster greater access to the financial system, with fewer of the risks posed by private sector-administered digital assets. A United States CBDC that is interoperable with CBDCs issued by other monetary authorities could facilitate faster and lower-cost cross-border payments

and potentially boost economic growth, support the continued centrality of the United States within the international financial system, and help to protect the unique role that the dollar plays in global finance. There are also, however, potential risks and downsides to consider. We should prioritize timely assessments of potential benefits and risks under various designs to ensure that the United States remains a leader in the international financial system.”

In traditional nation state currencies the central banks [control the amount](#) of currency in circulation by issuing debt to private banks, which is then loaned out to individuals. The debt is ‘destroyed’ on the balance sheet to remove currency through the reverse mechanism. They also facilitate government debt [47], and work (theoretically) outside of political control to adjust interest rates, in order to manage growth and flows of money.

Many things which cannot be done with traditional nation state money systems are possible with CDBCs, because they [remove the middleman](#) of private banking between the end user and the policy makers.

- Negative interest rates are possible, such that all of the money can lose purchasing power over time, and at a rate dictated by policy. This “removal of the lower bound” has been discussed by economists over the last couple of decades as interest rate mechanisms have waned in efficacy. It is not possible in the current system, and instead money must be added through [quantitative easing](#), which disproportionately benefits some through Cantillon effects [48, 49].
- Ubiquitous basic income is possible in that money can be issued directly from government to all approved citizens, transferring spending power directly from the government to the people. This also implies efficiency savings for social support mechanisms.
- Asset freezing and confiscation are trivial if CDBCs can replace paper cash money completely, as a bearer asset. Criminals and global ‘bad actors’ could have their assets temporarily or permanently removed, centrally, by suspending the transferability of the digital tokens.
- Targeted bailouts for vital institutions and industries are possible directly from central government policy makers. Currently private banks must be incentivised to make cheap loans available to sectors which require targeted assistance.
- Financial surveillance of every user is possible. In this way a ‘panopticon of money’ can be enacted, and spending rulesets can be applied. For instance, social support money might only be spendable on food, and child support only on goods and services to support childcare. This is a very dystopian set of ideas. Eswar Prasad says “In authoritarian societies, central bank money in digital form could become an additional instrument of government control over citizens rather than just a convenient, safe, and stable medium of exchange.” [50]
- It’s a virtually cost free medium of exchange, since there is no physical instrument which must be shipped, guarded, counted, assayed, and securely destroyed.
- The counterfeiting risk is significantly reduced because of secure cryptographic underpinnings rather than paper or plastic anti counterfeiting technologies.
- Global reach and control is instantly possible for the issuer. This is a big problem especially for a reserve currency such as the dollar. Two thirds of \$100 bills are [thought to](#) reside outside of the USA.
- System level quantitative easing and credit subsidies are made far simpler and less wasteful when centrally dictated.

- Transfer of liability and risk to the holder globally reduces the management costs for global deposits of a currency.
- It may be possible to automate the stability of a currency through continuous adjustment of the ‘peg’ through algorithms or AI.

The UK has signalled that it is not interested in developing a CBDC at this time. It is viewed as a [solution in search of a problem](#), with the Lords economic affairs committee saying:

“The introduction of a UK CBDC would have far-reaching consequences for households, businesses, and the monetary system for decades to come and may pose significant risks depending on how it is designed. These risks include state surveillance of people’s spending choices, financial instability as people convert bank deposits to CBDC during periods of economic stress, an increase in central bank power without sufficient scrutiny, and the creation of a centralised point of failure that would be a target for hostile nation state or criminal actors.”

In the USA this text from Congressman Tom Emmer shows how complex and interesting this debate is becoming.

Today, I introduced a bill prohibiting the Fed from issuing a central bank digital currency directly to individuals. Here’s why it matters:

As other countries, like China, develop CBDCs that fundamentally omit the benefits and protections of cash, it is more important than ever to ensure the United States’ digital currency policy protects financial privacy, maintains the dollar’s dominance, and cultivates innovation.

CBDCs that fail to adhere to these three basic principles could enable an entity like the Federal Reserve to mobilize itself into a retail bank, collect personally identifiable information on users, and track their transactions indefinitely.

Not only does this CBDC model raise “single point of failure” issues, leaving Americans’ financial information vulnerable to attack, but it could be used as a surveillance tool that Americans should never be forced to tolerate from their own government.

Requiring users to open an account at the Fed to access a United States CBDC would put the Fed on an insidious path akin to China’s digital authoritarianism.

Any CBDC implemented by the Fed must be open, permissionless, and private. This means that any digital dollar must be accessible to all, transact on a blockchain that is transparent to all, and maintain the privacy elements of cash.

In order to maintain the dollar’s status as the world’s reserve currency in a digital age, it is important that the United States lead with a posture that prioritizes innovation and does not aim to compete with the private sector.

Simply put, we must prioritize blockchain technology with American characteristics, rather than mimic China’s digital authoritarianism out of fear. Most analysts now seem to think that there is little appetite to replace all of a given currency with a CBDC. It is far more likely that a blend of stablecoins, private bank issued digital currency (with a yield incentive) and some limited CBDC, alongside the new contender Bitcoin, will present a new landscape of user choice. Different models of trust, insurance, yields, acceptability, and potentially privacy, will emerge.

Clearly a global, stable, wholly digital bearer asset in a native currency would be the ideal integration for money in a metaverse application, but it is likely that a transition to such a technology would be complex and painful. It is certainly not ready for consideration now.

[US CDBC whitepaper](#)

4.4 Bitcoin as a money

Since this book seeks to examine transfer of value within a purely digital environment it is necessary to ask the question of whether Bitcoin is money. It is beyond argument that the Bitcoin network is a rugged message passing protocol which achieves a high degree of consensus about the entries on its distributed database. Ascribing monetary value to those database entries is a social consensus problem, and this itself is a contested topic.

Jack Mallers, of Strike [presentation to the IMF](#) identified the following challenges which he claims are solved by the bitcoin monetary network.

- Speed
- Limited transparency and dependability
- High cost
- Lack of interoperability
- Limited Coverage
- Limited accessibility

Mallers further identifies the attributes of the ideal global money.

- Uncensorable
- Unfreezable
- Permissionless
- Borderless
- Liquid
- Digital

The Bitcoin community believes that [Bitcoin is the ultimate money](#), a ‘store of value’, chance to [separate money from state](#), increase [equality of opportunity](#) and [ubiquity of access](#), while others view it as ‘rat poison’, or a [fraudulent Ponzi scheme](#). A notable exclusion from the negative rhetoric is Fidelity, the global investment manager, who have always been positive and have [recently said](#):

“Bitcoin is best understood as a monetary good, and one of the primary investment theses for bitcoin is as the store of value asset in an increasingly digital world.”

The following paraphrases Eric Yakes, author of [‘The 7th Property’](#). Again, this is an Austrian economics perspective, and like much economic theory the underlying premise [is contested](#)[51].

“Paper became money because it was superior to gold in terms of divisibility and portability BUT it lacked scarcity. People reasoned that we could benefit from the greater divisibility/portability of paper money as long as it was redeemable in a form of money that was scarce. This is when money needed to be “backed” by something.

Since we changed money to paper money that wasn’t scarce, it needed to be backed by something that was. Since the repeal of the gold standard, politicians have retarded the meaning of the word because our money is no longer backed by something scarce.

So, what is bitcoin backed by? Nothing.

Sound money, like gold, isn’t “backed”. Only money that lacks inherent monetary properties must be backed by another money that maintains those properties. The idea that our base layer money needs to be backed by something is thinking from the era of paper money. Bitcoin does not require backing, it has inherent monetary properties superior to any other form of money that has ever existed.”



The Price Of One Bitcoin Could Exceed \$1 Million by 2030

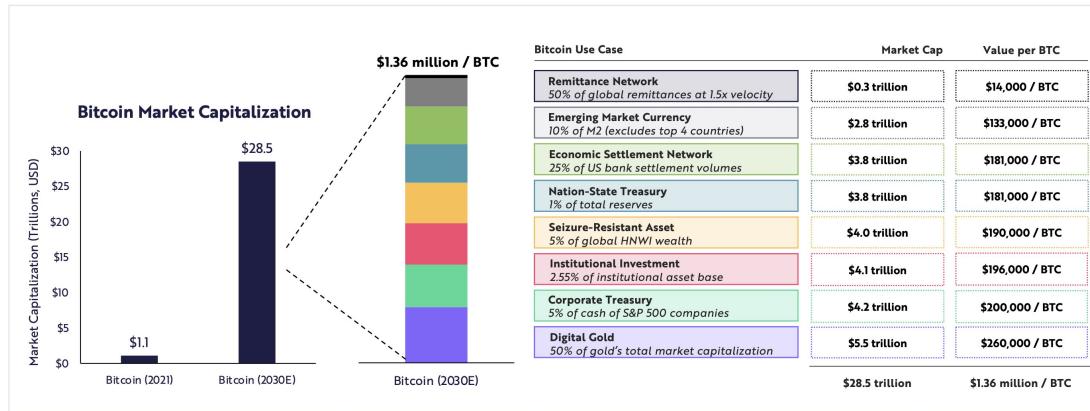


Figure 4.1: Potential market exposure to Bitcoin as a money

Perhaps more than any of these takes, it is worth considering the current public perception of the technology as a money and store of value. This [twitter thread](#) from professional sportsman Saquon Barkley, to his half million followers on the platform, captures the mood. He is one of a handful of athletes now being [paid directly](#) in Bitcoin.

"I want my career earnings to last generations. The average NFL career is 3 years and inflation is real. Saving and preserving money over time is hard, no matter who you are. In today's world: How do we save? This is why I believe in bitcoin. Almost all professional athletes make the majority of their career earnings in their 20s. With a lack of education, inaccessible tools, and inflation, a sad yet common reality is many enter bankruptcy later on. We can do better. We need to improve financial literacy. Bitcoin is a proven, safe, global, and open system that allows anyone to save money. It is the most accessible asset we've ever seen."

Andrew M. Bailey says “*in an ideal world where governments honour the rights of citizens, they don't spy, they don't prohibit transactions, they manage a sound money supply, and they make sound decisions, the value of bitcoin is very low; we're just not in an ideal world*”

The 2022 ARK Big Ideas report again provides some useful market insight. The posit that demand for the money features of Bitcoin could drive the price of the capped supply tokens to around 1M pounds per Bitcoin as in Figure 4.1.

Bitcoin is only a viable money when viewed in the layers described in this book. The base chain layer is the ultimate store of secure value. Whatever layer 2 ultimately emerges is the transactional layer which could replace day to day cash money, while the hypothetical layer 3 might be useful for complex financial mechanisms and contracts operating automatically.

4.5 Risks

It can be seen that following the invasion of Ukraine by Russia, that sanctions of various kinds were applied to the Russian economy. One of these was the previously discussed Swift international settlement network. Another whole category was the removal of support by private businesses domiciled outside of Russia and Ukraine, and pertinent here is that VISA, Mastercard, Paypal, and Western Union all removed support for their product rails. This means that while some cards and services still work, and will likely work again through Chinese proxies in the coming months, considerable disruption will be felt by Russian companies and individuals. This is not to say that this disruption is necessarily wrong, but it is clear now that all of these global financial transfer products and services are contingent on political factors. The same might be true of CBDC products if they gain traction globally. There is certainly no reason why all money within a physically delineated border could not be blocked or cancelled. This is not as true for Bitcoin at this time, though again, with enough political will it is technically plausible to incentivise miners with additional payments to exclude transactions from geolocated wallets. This would be mitigated by Tor, and in a global anonymous network it is very likely that a miner could be found at a higher price for inclusion in the next block.

4.5.1 Hyperbitcoinization

Hyperbitcoinization is a term coined in 2014 by Daniel Krawisz [52]. It is the hypothetical rise of Bitcoin to become the global reserve currency, and the demonetisation of all other store of value assets. This seems unlikely but is backed by a game theoretic analysis of both Bitcoin and current macro economics. Fulcrum Ventures (a venture capital firm) provide a [blog post series](#) about the route this might take. It's beyond the scope of this paper to look at the implications of this possibility, but they are clearly significant if true.

4.6 Does DeFi matter to SMEs

DeFi is decentralised finance, and might only exist because of partial regulatory capture of Bitcoin. If peer to peer Bitcoin secured yield and loans etc were allowed then it seems unlikely that the less secure and more convoluted DeFi products would have found a footing. DeFi has been commonplace over the last few years. It enables trading of value, loans, and interest (yield) without onerous KYC. If Bitcoin's ethos is to develop at a slow and well checked rate, and Ethereum's ethos is to move fast and break things, then DeFi could best be described as throwing mud and hoping some sticks. It is characterised by rapid innovation, huge yields for early adopters, incredibly high risk, and a culture of speculation which leads to products being discarded and/or forked into something else in the pursuit of returns.

Much of the space is now using arcane gamification of traditional financial tools, combined with memes, to promote what are essentially pyramid schemes. Scams are very commonplace. Loss of funds though code errors are perhaps even more prevalent.

The Bank for International Settlements have the stated aim of supporting central banks monetary and financial stability. Their [2021 report on DeFi](#) noted the following key problems.

- ..a “decentralisation illusion” in DeFi due to the inescapable need for centralised governance and the tendency of blockchain consensus mechanisms to concentrate

power. DeFi's inherent governance structures are the natural entry points for public policy.

- DeFi's vulnerabilities are severe because of high leverage, liquidity mismatches, built-in interconnectedness and the lack of shock-absorbing capacity.

These are two excellent and likely true points. In addition access to DeFi is 'usually' through Web2.0 centralised portals (websites) which are just as vulnerable to legal take-down orders and any other centralised technology.

There are more recent DeFi on Bitcoin contenders, but these are vulnerable to the [same attacks](#) and problems in the main.

There is likely no use for this technology for small and medium sized companies on the international stage. It is far more likely that reputation would be damaged. The 'best' of the portfolio of DeFi offerings is perhaps high yield stablecoin accounts, where dollars equivalent tokens are locked up providing very high return rates of up to 20 percent. It's also possible to get loans (by extension business loans) out of such systems at relatively low risks. The best 'distributed' example of this is probably [Lend, at HODLHODL](#), which is a peer to peer loan marketplace. Many more custodial options exist for loans (CASA, BlockFi, Nexo, Ledn, Abra etc). These might not really fit the definition of DeFi at all, but they are potentially useful to companies who have Bitcoin on their balance sheet long term.

4.7 Do DAOs matter to SMEs

A distributed autonomous organisation, or DAO is a governance structure which is built in distributed code on a blockchain smart contract system. Token holders have voting rights proportional to their holding. The first decentralised autonomous organisation was simply called "The DAO" and was launched on the Ethereum network in 2016 after raising around \$100M. [It quickly succumbed to a hack and the money was drained](#). This event was an important moment in the development of Ethereum and resulted in a code fork which preserves two separate versions of the network to this day, though one is falling into obsolescence.

In practice DAOs have very few committed 'stakeholders' and the same names seem to crop up across multiple projects. Some crucial community decisions within large projects only poll a couple of dozen eligible participants. Its might be that the experiment of distributed governance is failing at this stage.

Perhaps more interesting is the use of the DAO concept to crowd fund global projects, currently especially for the acquisition of important art or cultural items. DAOs are also emerging as a way to fund promising technology projects, though this is reminiscent of the 2017 ICO craze which ended badly and is likely to fall foul of regulations.

Within the NFT and digital art space PleaserDAO has quickly established a strong following. "PleasrDAO is a collective of DeFi leaders, early NFT collectors and digital artists who have built a formidable yet benevolent reputation for acquiring culturally significant pieces with a charitable twist.

Opensea wrangle between IPO and governance token.

ConstitutionDAO, Once upon a time in Shaolin etc

4.7.1 Bisq DAO

One of the better designed DAOs is [Bisq DAO](#). It's slightly different design tries to address the issue of overly rigid software intersecting with more intangible and fluid human governance needs. From their website:

"Revenue distribution and decision-making cannot be decentralized with traditional organization structures—they require legal entities, jurisdictions, bank accounts, and more—all of which are central points of failure. The Bisq DAO replaces such legacy infrastructure with cryptographic infrastructure to handle project decision-making and revenue distribution without such central points of failure."

4.7.2 Risks

The most interesting thing about DAOs is that they belong more in this money chapter than they do in blockchain. As we have seen they're finding most success as loosely regulated crowd funding platforms. If a small company did find itself wishing to explore this fringe mechanism for raising capital, then we would certainly recommend keeping a global eye on evolving regulation and the onward legal exposure of the company.



5. Distributed Self Sovereign Identity

This section need to point out that DID/SSI is broken, ION is untested, and everyone's kinda given up. Webs of trust are viable so this means Nostr or Slashtags. Maybe LNURL-Auth can do it.

This is where an external contributor can add information about DID and SSI Distributed self sovereign identity has a great elevator pitch. Individuals should be empowered through technology to manage their own data, without manipulation or exploitation by centralised corporate behemoths. In practice it's a staggeringly complex proposition which increases risk to the individual, decreases convenience, and despite much work, does not even make much sense in it's own terms.

5.1 Applications of DID/SSI

Some of the likely, and discussed applications for DID/SSI are the more inherently private and personally valuable sets of data a individual might generate throughout their life. The theory is that subsets of such data could then be digitally revealed by the individual when required, and that cryptographic verification built into the system would guarantee the veracity of the data to the receiving party. It is also possible to make use of "zero-knowledge proof" such that assertions can be made about the contents of the data without revealing the data itself. A good example of this an age verification challenge, where a threshold age could be asserted without necessarily revealing the date of birth. Other keystone uses of the technology are:

- health documents history
- qualifications and certifications
- financial record and relationships with those of others
- contacts, connections to other people and their appropriate data, including things like shared and personal calendars

It's also possible to extend this key management ethos to all login credentials, and all data currently stored on centralised servers. This is the tension discussed in the chapter about Web3. Proponents think that using something like a DID/SSI stack to manage encryption, decryption and access to data within cloud services gives the user the best of all worlds. They see simply logging in with a cryptographic wallet, and using that same public/private

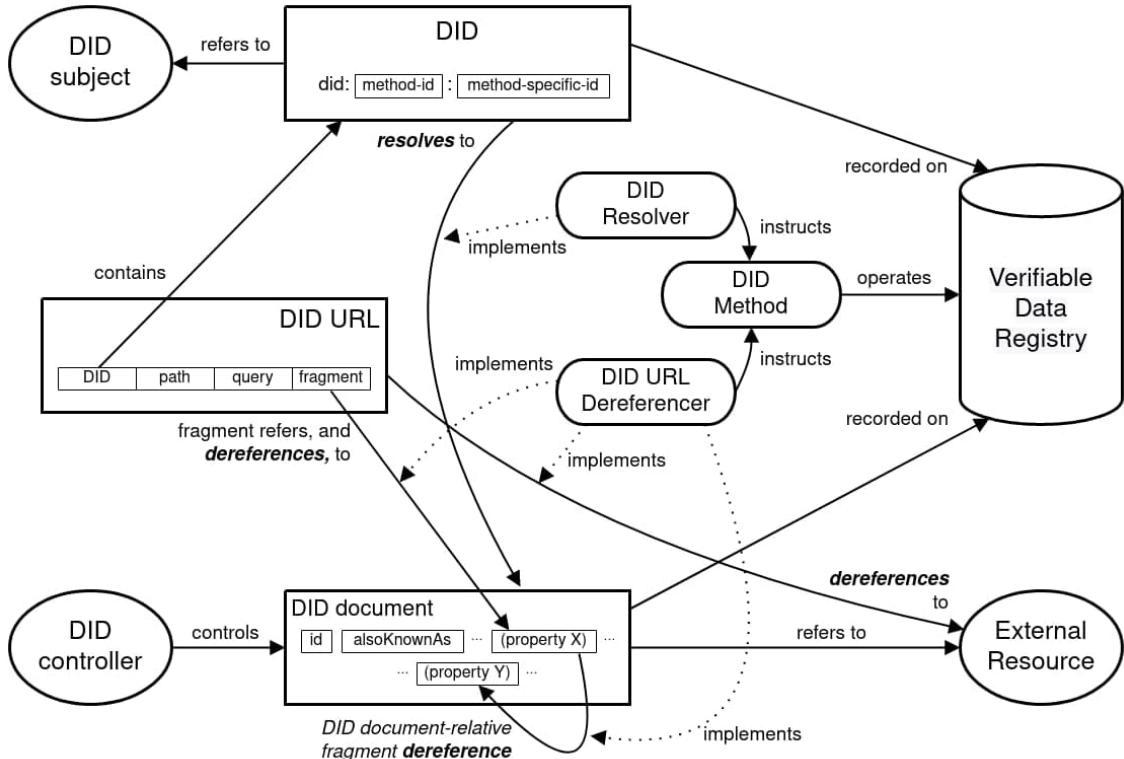


Figure 5.1: Part of the DID SSI specs

key pair to manage the data beyond as some kind of panacea. This is very complex stuff though, and it seems very likely they just haven't thought this through enough.

5.2 Classic DID/SSI

Distributed identity / self sovereign identity has been extensively researched for decades, with hundreds of peer reviewed papers, and extensive support from the [world wide web consortium](#). The academic field now seems quite ossified and has settled on a couple of hundred ‘schema’ which they feel underpin the next layer of development. It is a [complex field](#), and the language and diagrams are arcane and self referential as seen in Figure 5.1.

Moreover the minimal implementation of such proposed systems hints at a [federated model](#) of centralised ‘truth’ to enable persistence of identifiers over time.

The major failing of the DID/SSI work to date is a lack of meaningful use cases with incentives for adoption. This is clearly explained by Lockwood [53] who proposes that the pathway to adoption of ‘classic’ DID/SSI requires an incentive over and above the current identity management on the web. Being distributed is not enough. Especially in the light of questionable assurances of this even being true.

Perhaps most concerning is this [recent exchange](#) on the mailing lists. Here, the two inventors of DID say the following:

“Not a single entity I know that’s doing production deployments has actually vetted did:ion and found it to be production capable. This goes for every DLT-based DID Method out there - even the one we’re working on. I am highly sceptical of anyone that says that any DID Method is ready for production usage at present.”

Agreed — as one of the proponents of DLTs (in particular permissionless public ones) none are mature enough yet for production.”. It seems then that we can rule out use of these technologies?

5.2.0.1 DID principles

The core principles of distributed identity are that there should be persistent identifiers, like real world documents which assert identity, but with extended use cases. These should be permanent, and resolvable everywhere, forever. Underpinning this is cryptographically verifiable and decentralised data, managed by the user, or their trusted proxy. As primitives this makes them lifetime digital assets, that are portable, and unconfiscatable, with no required reliance on a trusted third party. By this stage in the book you should be familiar with these concepts, but application of this fundamental mindset to all personal data and digital interactions is a bigger reach even than money and value.

5.2.0.2 What's in a DID document?

All classic DID is underpinned by a DID document what bootstrap the services it's connected to. It is made up of one or more public keys. The documents can make use of services such as timestamps, cryptographic signatures, proofs, delegations, and authorisations. They should contact the minimum about of information to accomplish the specific task required of them.

5.3 Newer Technologies

5.3.1 Slashtags

Slashtags is a new distributed identity open method being developed by Bitfinex and Tether under the Synonym suite. It uses bitcoin keys for authentication, but communicates a schema through a metadata exchange.

5.3.2 LNURL-Auth

5.3.3 Sovrin

Sovrin foundation are established global players in distributed identity and we wish to apply to join their federation of credential issuers.

5.3.4 Keri

5.3.5 Atala Prism

5.3.6 Microsoft ION

While working at Microsoft on ION Daniel Buchner (now working at Square) or Henry Tsai [said the following](#), which is worth quoting verbatim:

“While ledger-based consensus systems, on the surface, would seem to provide the same general features as one another, there are a few key differences that make some more suitable for critical applications, like the decentralized identifiers of human beings. Some of these considerations and features are:

- The system must be open and permissionless, not a cabal of authorities who can exclude and remove participants.

- The system must be well-tested, and proven secure against attack over a long enough duration to be confident in.
- The system must produce a singular, independently verifiable record that is as immutable as possible, so that reversing the record the system produces is infeasible.
- The system must be widely deployed, with nodes that span the globe, to ensure the record is persisted.
- The system must be self-incentivized, so that nodes continue to operate, process, and secure the record over time. The value from operation must come from the system directly, because outside incentive reliance is itself a vector for attack.
- The cost to attack the system through any game theoretically available means must be high enough that it is infeasible to attempt, and even if an ultra-capitalized attacker did, it would require a weaponized mobilization of force and resources that would be obvious, with options for mitigation.

The outcome:

- Number 1 eliminates private and permissioned ledgers
- Number 2 eliminates just about all other ledgers and blockchains, simply because they are inadequately tested
- For the metrics detailed in 3-6, Bitcoin is so far beyond all other options, it isn't even close - Bitcoin is the most secure option by an absurdly large margin."

On the surface then it might seem that the choice is Bitcoin again, and indeed that the open source Microsoft ION stack is a natural choice.

5.3.7 Atala Prism (ADA Ecosystem)

This [Medium post](#) lists the pre-requisites required for interaction with Eth within Unity. At this stage this hasn't been expanded out. It's based around the .NET 'nethereum' library which can be found on their [github](#).

5.4 Risks & Challenges?

Classic DID/SSI risks fragmentations In all DID scaling to a world where the user is managing potentially thousands of these critical cryptographic data files is daunting. Abstracting the guts of this away to make the user simple, and only caring about right level of information, turns out to be huge problem that nobody has solved It's not clear that users want this In the case of web of trust like Slashtags it's a big piece of work for the users to rate off of their digital interactions with a trust metric.



6. Non Fungible Tokens

Nonfungible tokens allow [digital and new media artists](#) to connect with audiences without gatekeepers. Established mediators and curators of art have been caught totally wrong-footed, and NFTs seem to give a way for them to be cut out completely. There are suggestions of applications beyond this initial digital art scope. This is a compounding, and disrupting paradigm change.

While it is likely that there is currently a speculative bubble, and indeed is waning already, it seems certain that the technology is here to stay.

One of the most interesting companies is Yuga Labs, who launched the incredibly popular Bored Ape Yacht club set of 10,000 algorithmically generated NFTs. These Ethereum based NFTs regularly change hands for hundreds of thousands of pounds.

Yuga recently bought the artistic rights to the commercial reuse of similarly popular ‘Crypto Punks’ set. This is interesting because they have handed the commercial re-use rights to the owners of the individual NFTs.

The community around these collections is incredibly strong, mixing developers, artists, the rich and famous, and the fortunate and early, into a cohesive community who communicate online. The developer ‘good will’ is enormous, and it seems possible that this will lead to faster and broader innovation around the collections, and out into metaverse applications. The brand is strong, and the individual NFT items both benefit from, and reinforce that brand, while adding personal narratives and human interest.

As a gauge of how frothy this market still is it’s interesting to look at the APE token which Yuga just launched. They airdropped 10,000 of the tokens free to each of the 10,000 NFT holders. This instantly created a multi-billion dollar market cap, and a top 50 ‘crypt’ out of thin air. It’s clear that there is both brand, and a market here.

6.1 Digital Art

NFT Key Points

- 1) NFTs are different from ERC-20 tokens, such as DAI or LINK, in that each individual token is completely unique and is not divisible.
- 2) NFTs give the ability to assign or claim ownership of any unique piece of digital data, trackable by using Ethereum’s blockchain as a public ledger. An NFT is minted

- from digital objects as a representation of digital or non-digital assets
- 3) NFTs are minted through smart contracts that assign ownership and manage the transferability of the NFT's. When someone creates or mints an NFT, they execute code stored in smart contracts that conform to different standards, such as ERC-721. This information is added to the blockchain where the NFT is being managed.

NFT Use Cases

Art

Without a doubt the recent surge of interest in NFT's during early 2021, has largely been driven by digital art NFT's.

In 2014, New York artist [Kevin McCoy's *Quantum*](#) is widely recognised as the first piece of art created as an NFT. However it was only during early 2021 that art NFT's started to gain significant attention and by the end of 2021, nearly £31b had been spent on NFT purchases, a considerable and exponential growth given 2020 sales of ~£71m

High profile digital artists such as *Beeple* whose [recent recording break sale](#) of his NFT "The first 5000 days" at Christies (a long established British auction house, specialising in high profile precious work of art) for £52.9m helped bring NFT's into the public spotlight and wider give them global attention.

Art as NFT's offer the following advantages:

- a. **Proof of ownership/creation:** At its most basic level, once a work of art is 'minted' (publishing the art work as a unique token on the blockchain) this function as a proof of ownership and by extension proof of creation.
- b. **Secure Digital Provenance** Tracking ownership and reselling of art; [Provenance](#) (or the chain of custody) now no longer becomes a *detectives game* or open to fraud or misinterpretation, since provenance can increase the value of a piece of art (benefiting both the creator and collector) the use of the blockchain as an open, secure ledger is a far more trustworthy system than traditional methods of artistic provenance that were cobbled together (often consisting of documents spanning private & public sale receipts, art/museum gallery exhibitions and private record keeping). Digital provenance provided when an artist mints a piece of art into an NFT allows artists and collectors to
- c. **Decentralised automated royalty payments:** If a work of art is minted into an NFT and artist's royalty payments can be predetermined and automated in perpetuity via smart contracts. Once the royalty payment rate is set by the artist/creator, future royalties of all sales can be paid directly to the artist/creator account (via a digital wallet) without the need of a third party (traditionally a gallery/agent etc..). The first sale of a piece of art would often (but not always) benefit (financially/critically) the artist, however secondary and subsequent sales would historically only ever financially benefit the collector; the original artist would rarely benefit. However the use of smart contracts NFT smart contracts means that even if piece of art is resold 5, 10 or even a 100,000 times moving through 5, 10 or even a 100,000 different collectors; a pre-determined royalty payment rate set by the creator would still guarantee the artist/creator is paid directly from each and every future sale. If one considers that historically provenance for works of art can spans generations, then NFT smart contracts hold an incredibly potential i.e and artists descendants financially benefit directly from the resale of a piece of work long after the artist/museum's/gallery have turned to dust as long as the original creator's digital wallet is accessible, *the blockchain becomes an everlasting digital patron* .

Computer & Video Games

Computer & Video games are a huge global business. As the industry has matured, secondary markets have emerged, most notably the ‘second hand’ games resale market. The rise of ‘retro’ gaming, has demonstrated .Despite the move to non-physical digital only’ games, the demand remains incredibly high however publisher/developers/retailers aren’t able to directly benefit from the emerging market. The potential of *video games as NFT’s* presents a huge opportunity for publishers, developers and players alike.

- a. Royalty Sales on Pre-owned Games ; A predetermined proportion of any resale of a used game can automated in perpetuity via smart contracts; once these are set by the publisher, future royalties of all sales can be paid directly to the publishers/developers wallets (a digital account) without the need of a third party (traditionally a retail entity). Traditionally only the initial first sale of a game would financially benefit the publisher/developer/retailer, secondary and subsequent sales would only ever financially benefit the purchaser, with many developers/publishers arguing this is hurting the wider industry through the loss of significant income generated by the secondary and subsequent sales, sometimes over the course of decades. However the use of NFT’s smart contracts means that if a game is sold/resold through 10,000 collectors; a pre-determined royalty payment rate set by the publisher would still guarantee the publisher (and or developer/retailer) takes a proportion of any future sales.
- b. **Monetisation of User Generated Content:** Games as a NFT’s offer ability to monetise UGS: User generated content User generated states (i.e save games, user can invest hundreds/thousands of hours on a particular game). Games like Nintendo’s Pokemon, Bungie’s Destiny or Genshin Impact demonstrate that players invest significant amounts of time collecting in game digital assets; often used by players as mark of distinction/status symbol.

However traditionally there has been no real utility or value that went beyond sharing a *digital badge* (i.e triumph/achievement) often on a social media/gamer’s platform profile. NFT’s offer the ideal system for developers/publishers/players to monetise user generated/customised content/data (such as a players unique save game data), simultaneously developing an additional monetised ecosystem to meet player demands (i.e certain players willing to monetise and ‘sell’ their invested time in a particular product/service to other players with little time but willing to pay other players for ‘grinding’ (progressing laborious in game tasks)

The potential to provide publishers/developers with an additional long-term income stream, providing a better ROI on computer & video game development, which in many instances can cost hundreds of millions in development costs spanning 5/10 years, is undeniable.

- c. Play to earn revenue models.
- d. Monetizing In game collectibles: customisable in game assets (vanity items such as cosmetic character skins/clothing or collectible items that offer player advantages(new weapons/vehicles/mods etc,...)

6.2 Other uses

Peter Thiel, the billionaire venture capitalist who founded PayPal has invested in expanded NFT use cases. The first is ‘Royal’ which is experimentally selling [selling limited NFT tokens](#) which contractually entitle the holder to a portion of music artist royalties. The other is a [political funding NFT](#) from Blake Masters to support his senate ambitions.

It is completely reasonable to assert that these use cases could be accomplished without the use of NFT technology, and is part of the hype bubble.

NFT art currently suffers from the same failure of decentralisation already discussed in the Ethererum technology stack, but this is compounded by the normalisation of intermediate art brokers [continuing to custody](#) the NFTs even after sale. They are usually selling a pointer to their own servers. The market is nascent and evolving, but it’s currently not delivering on it’s core promise.

6.3 Is any of this useful?

Maybe, maybe not. This hype cycle isn’t the best place or time to judge this.

6.4 User stories / behaviours

- As a user I want to select a digital asset I find in a metaverse and then be offered an option to purchase the asset so I can look at it in my own spaces.
- As a user I want to click on a digital asset I find in a metaverse and be given the opportunity to buy it as a rare digital representation so that only I and a few others are provably certified to own.
- As a metaverse ‘land’ holder (stakeholder) I would like to reward winners of competitions with real money or digital assets to foster and gamify the system
- As a user I would like to interact with a virtual marketplace where I could swap and trade digital assets with other users so that I continue to feel engaged.
- As a user I would like to create content so I can take it to metaverse locations and monetise it.

6.5 Comparing the technologies

Figure 6.1 on the next page compares the technologies discussed so far in the context of the behaviours in section 7.12.

Looking at these row by row it can be seen that:

- Bitcoin layer 1 is too slow. Additionally the technical challenge of interacting with the network are considerable, and the fees are both high and likely to rise.
- Ethereum ERC721 non fungible tokens might be a possibility for supporting limited edition digital assets. A web interface could be built on cloud servers to handle the shop and back end commitments to the Ethereum network. While fully committed transactions would take a few minutes to process the shop itself could inform the users immediately then email out a certificate of ownership. It would be possible to trade these items on other platforms, and conceivably within an app. Prizes (digital assets) could be issued to users for competition activity. The system would be expensive to use, the fees might vary wildly, and the ecological costs might

be enormous. Batching transactions could overcome this, but that would rely on substantial take-up of the service, and management overhead.

- Building and issuing and Ethereum based token, in an ICO process or similar seems a disproportionate amount of investment and would be expensive to use.
- A custom private blockchain offers no advantages over a database. It is a difficult thing to implement, but would have better performance and use cost to the users.
- BTC pay server is an interesting open source project which allows companies to host a simple lightning / BTC web shop. It's not clear what the use case is here, but it is well engineered and scalable and should be born in mind when thinking about the other problems and systems.
- Lightning network has open source connectivity into Unity, works well with QR codes, is fast, and has low transaction latency. The only issue is that it offers nothing but marketing hype over and above more traditional digital point of sale layers, and is more difficult to connect to money rails in the UK.
- New layer 3 solutions like RGB, stacks, or Synonym might allow cheap and fast in

6.6 User stories / behaviours

- As a user I want to.
- As a user I want to.
- As a user I want to.

6.7 Comparing the technologies

This section is well out of date and needs a tonne of work.

Figure 6.1 on the next page compares the technologies discussed so far in the context of the behaviours in section 7.12.

Looking at these row by row it can be seen that:

- Bitcoin is too slow. Additionally the technical challenge of interacting with the network are considerable, and the fees are both high and rising.
- Ethereum ERC721 non fungible tokens might be a possibility for supporting limited edition digital asset sales. A web interface could be built on cloud servers to handle the shop and back end commitments to the Ethereum network. While fully committed transactions would take a few minutes to process the shop itself could inform the users immediately then email out a certificate of ownership. It would be possible to trade these items on other platforms, and conceivable within the app. Prizes (digital assets) could be issued to users for competition activity. The system would be expensive to use, the fees might vary wildly, and the ecological costs might be enormous. Batching transactions could overcome this, but that would rely on substantial take-up of the service, and management overhead.
- Building and issuing and Ethereum based token, in an ICO process or similar seems a disproportionate amount of investment and would be expensive to use.
- A custom private blockchain offers no advantages over a database. It is a difficult thing to implement, but would have better performance and use cost to the users.
- BTC pay server is an interesting open source project which allows companies to host a simple lightning / BTC web shop. It's not clear what the use case is here, but

	Likely dev time	Cost of Use	Speed of Use	Energy Cost	Open source	Bank connect	Buying things	Prizes	Trading Peer 2 Peer
Bitcoin	N/A	~£10 per use	~H	High	Yes	Maybe (Strike)	Yes	Yes	Maybe
Eth ERC721	N/A	~£10 per use	Mins	Very High	Yes	No	No	Yes	Maybe
Eth token	Weeks Months	~£10 per use	Many Mins	High	Maybe	No	Maybe	Yes	Yes
Custom chain	Months	Very low	Secs	Moderte	Maybe	Maybe	Yes	Yes	Yes
BTC pay server + LND	Days but operational overheads	~£10 per use	Secs	Lower	Yes	Yes	Yes	Yes	No
BTC L2 and L3	Months	Very Low	Secs	Low	Some	Maybe	Yes	Yes	Yes
Colour coins	Months	Very Low	Secs	High	Maybe	Maybe	Yes	Yes	Yes

Table 6.1: This table is basically broken and out of date and will be sorted out soon!

it is well engineered and scalable and should be born in mind when thinking about the other problems and systems.

- Lightning network has open source connectivity into Unity, works well with QR codes, is fast, and has low transaction latency. The only issue is that it offers nothing but marketing hype over and above more traditional digital point of sale layers, and is more difficult to connect to money rails in the UK.
- Coloured coins would allow all of the desirable functionality, but it would be a substantial development challenge, using old open source code, and have high use fees.
- RGB MyCitadel stack seems highly appropriate and is open source but is in development and will not be available until later in the year.

6.8 What are the options

ETH, SOL, ADA, STX, RGB, Liquid?, RSK, so many others, this is gonna be exhausting to write about.

6.9 Risks

6.9.1 Wage slavery / pyramid schemes

Much of the current play to earn models are digital sweat shops. etc

6.9.2 Energy concerns

It is under discussed within the community that ‘minting’ non fungible tokens, predominantly on Ethereum, is energy wasteful, on an already energy inefficient platform. As a

random example this single collector of a few images accounts for nearly 10 kilotonnes of carbon emissions. This is a subtly different problem to the security of the Bitdoooin economic network in that these tokens are clealy non fungible, so cannot be reused again and again in order to reduce the unit environmental impact.

6.10 Why choose bitcoin again

For the purposed of this book, we are going to once again explore Bitcoin as the platform of choice. To be clear it's not the popular, or intuitive choice. It's barely possible to mint NFTs on Bitcoin, and the community is deeply sceptical about it.

because it's the true opensource money option, the cost / barrier to entry is far lower than ethereum, and with new technology such as [NFT integration directly with DIBA on RGB this integrates](#).



7. Metaverses

7.1 History and market need

The word metaverse was coined by the author Neal Stephenson in his 1992 novel Snowcrash. It started popping up soon after in [news articles](#) and research papers [54], but in the last five years it has been finding a new life within a silicon valley narrative.

There were clear precursors to modern social VR, such as [VRML in the 1990's](#) which laid much of the groundwork for 3D content over networked computers.

It might seem that there would be a clear path from there to now, in terms of a metaverse increasingly meaning connected social virtual spaces, but this has not happened. Instead interest in metaverse as a concept waned, MMORG (described later) filled in the utility, and then recently an entirely new definition emerged. The concept of the Metaverse is extremely plastic at this time (Figure 7.1). This book will aim to build toward an understanding of metaverse which really means a useful social mixed reality layer, that allows low friction communication and economic activity, within groups, at a global scale. This focus on value and trust means it's most appropriate to focus on business uses.

This chapter will attempt to frame the context for telepresence (the academic term for communicating through technology), and then explain the increasingly polarised options for metaverse. It's useful to precisely identify the primitives of the product we would like to see here, so this chapter is far more a review of academic literature in the field, culminating in a proposed framework.

7.2 Current telecollaboration

7.2.1 Video conferencing

Video-conferencing has become more popular as technology improves, and with increasing demands for real-time communication across greater distances. The full effects of video-conferencing on human communication are still being explored. Video-conferencing is presumed to be a somewhat richer form of communication than email and telephone, but not quite as informative as face-to-face communication.

In this section we look at the influence of eye contact on communication and how



Figure 7.1: Elon Musk agrees with this on Twitter

video-conferencing mediates both verbal and non-verbal interactions. Facilitation of eye contact is a challenge that must be addressed so that video-conferencing can approach the rich interactions of face-to-face communication. This is an even bigger problem in the emerging metaverse systems, so it's important that we examine the history and trajectory.

There is a tension emerging for companies who do not necessarily need to employ remote meeting technology, but also cannot afford to ignore the competitive advantages that such systems bring. In an experiment performed well before the 2020 global pandemic at CTrip, Bloom et al describe how home working led to a 13% performance increase, of which about 9% was from working more minutes per shift (fewer breaks and sick-days) and 4% from more calls per minute (attributed to a quieter working environment) [55]. Home workers also reported improved work satisfaction and experienced less turnover, but their promotion rate conditional on performance fell. This speaks to a lack of management capability with such systemic change. It's clearly a complex and still barely understood change within business and management.

Due to the success of the experiment, CTrip rolled-out the option to work from home to the whole company, and allowed the experimental employees to re-select between the home or office. Interestingly, over half of them switched, which led to the gains almost doubling to 22%. This highlights the benefits of learning and selection effects when adopting modern management practices like working from home. Increasingly this is becoming a choice issue for prospective employees, and an advantage for hiring managers to be able to offer it.

More recently enterprise collaboration systems provide rich document management, sharing, and collaboration functionality across an organisation. The enterprise ECS system may integrate collaborative video [56]. This is for instance the case with Microsoft Teams / Sharepoint. This integration of ECS should be considered when thinking about social VR systems which wish to support business, value, and trust. It is very much the case that large technology providers are attempting to integrate their 'business back end' systems into their emerging metaverse systems. Open source equivalents are currently lacking.

7.2.1.1 Pandemic drives adoption

The ongoing global COVID-19 pandemic is changing how people work, toward a new global 'normal'. Some ways of working are overdue transformation, and will be naturally

disrupted. In the UK at least it seems that there may be ‘real appetite’ to shift away from old practises. This upheaval will inevitably present both challenges and opportunities.

Highly technical workforces, especially, can [operate from anywhere](#). Little seems to have changed since then, until the recent pandemic forced the hand of global business.

If only a small percentage of companies allow the option of remote working, then they gain a structural advantage, enjoying benefits of reduced travel, lower workplace infection risk across all disease, and global agility for the personnel. Building and estate costs will certainly be reduced. More diversity may be possible. Issues such as sexual harassment and bullying may be reduced. With reduced overheads product quality may increase. If customers are happier with their services, then over time this ‘push’ may mean an enormous shift away from centralised working practises toward distributed working.

Technologies which support this working style were still in their infancy at the beginning of the pandemic. The rush to ‘Zoom’, a previously relatively unknown and insecure [57] web meeting product, shows how naive businesses were in this space.

Connection of multiple users is now far better supported, with Zoom and [Mircosoft Teams](#) alone supporting hundreds of millions of chats a day. This is a 20x increase on market leader Skype’s 2013 figure of [280 million](#) connections per month. Such technologies extend traditional telephony to provide important multi sensory cues. However, these technologies demonstrate shortfalls compared to a live face-to-face meeting, which is generally agreed to be optimal for human-human interaction [58].

While the research community and business are learning how to adapt working practises to web based telepresence [59], there remains little technology support for ad-hoc serendipitous meetings between small groups. It’s possible that Metaverse applications can help to fill this gap, by gamification of social spaces.

7.2.1.2 **Mona Lisa Type Effects**

Almost all traditional group video meeting tools suffer from the so-called Mona Lisa effect describes the phenomenon where the apparent gaze of a portrait or 2 dimensional image always appears to look at the observer regardless of the observer’s position [60, 61, 62]. This situation manifests when the painted or imaged subject is looking into the camera or at the eyes of the painter [63, 64].

Single user-to-user systems based around bidirectional video implicitly align the user’s gaze by constraining the camera to roughly the same location as the display. When viewed away from this ideal axis, it creates the feeling of being looked at regardless of where this observer is [65, 60, 61, 62], or the “collapsed view effect” [66] where perception of gaze transmitted from a 2 dimensional image or video is dependent on the incidence of originating gaze to the transmission medium.

Multiple individuals using one such channel can feel as if they are being looked at simultaneously, leading to a breakdown in the normal non-verbal communication which mediates turn passing [67]. There is research investigating this sensitivity when the gaze is mediated by a technology, finding that “disparity between the optical axis of the camera and the looking direction of a looker should be at most 1.2 degrees in the horizontal direction, and 1.7 degrees in vertical direction to support eye contact” [68, 69]. It seems that humans assume that they are being looked at unless they are sure that they are not [70].

It seems that the effect is truely limited to 2D surfaces. A 3D projection surface (a physical model of a human) designed to address this problem completely removed the Mona Lisa effect [65].

To be clear, there are technological solutions to this problem, but it's useful in the context of discussing metaverse to know that this problem exists. It's known that there are cognitive dissonances around panes of video conference images. Metaverse offers the promise of solving this, making more natural interaction possible, but it's clearly a long way from delivering on those promises.

7.3 What's important for human communication

7.3.1 Vocal

In the following sections we will attempt to rigorously identify just what is important for our proposed application of business centric communication, supportive of trust, and thereby value transfer.

In his book 'Bodily Communication' [71] Michael Argyle divides vocal signals into the following categories:

1. Verbal
2. Non-Verbal Vocalisations
 - a. Linked to Speech
 - i. Prosodic
 - ii. Synchronising
 - iii. Speech Disturbances
 - b. Independent of Speech
 - i. Emotional Noises
 - ii. Paralinguistic (emotion and interpersonal attitudes)
 - iii. Personal voice and quality of accent

Additional to the semantic content of verbal communication there is a rich layer of meaning in pauses, gaps, and overlaps [72] which help to mediate who is speaking and who is listening in multi-party conversation. This mediation of turn passing, to facilitate flow, is by no means a given and is highly dependent on context and other factors [73]. Interruptions are also a major factor in turn passing.

This extra-verbal content [74] extends into physical cues, so-called 'nonverbal' cues, and there are utterances which link the verbal and non-verbal [75]. This will be discussed later, but to an extent, it is impossible to discuss verbal communication without regard to the implicit support which exists around the words themselves.

In the context of all technology-mediated conversation the extra-verbal is easily compromised if technology used to support communication over a distance does not convey the information, or conveys it badly. This can introduce additional complexity [75].

These support structures are pretty much lacking in metaverse XR systems. The goal then here perhaps is to examine the state of the art, and remove as many of the known barriers as possible. Such a process might better support trust, which might better support the kind of economic and activity we seek to engineer.

When examining just verbal / audio communication technology it can be assumed that the physical non-verbal cues are lost, though not necessarily unused. In the absence of non-verbal cues it falls to timely vocal signals to take up the slack when framing and organising the turn passing. For the synchronising of vocal signals between the parties to be effective the systemic delays must remain small. System latency, the inherent delays added by the communication technology, can allow slips or a complete breakdown of 'flow' [76]. This problem can be felt in current social VR platforms, though people don't

necessarily identify the cause of the breakdown correctly. In the main they feel to the users like a bad “audio-only” teleconference.

With that said, the transmission of verbal / audio remains the most critical element for interpersonal communication as the most essential meaning is encoded semantically. There is a debate about ratios of how much information is conveyed through the various human channels [77], but it is reasonable to infer from its ubiquity that support for audio is essential for meaningful communication over a distance. We have seen that it must be timely, to prevent a breakdown of framing, and preferably have sufficient fidelity to convey sub-vocal utterances.

In the context of social immersive VR for business users, this implies that much better microphones are important, and the system should support both angular spatialisation, and respond to distance between interlocutors.

7.3.2 Nonverbal

We have already seen that verbal exchanges take place in a wider context of sub vocal and physical cues. In addition, the spatial relationship between the parties, their focus of attention, their gestures and actions, and the wider context of their environment all play a part in communication [78]. These are identified as follows by Gillies and Slater [79] in their paper on virtual agents.

- Posture and gesture
- Facial expression
- Gaze
- Proxemics
- Head position and orientation
- Interactional synchrony

This is clearly important for our proposed metaverse application. Below we will examine these six areas by looking across the wider available research.

7.3.2.1 Gaze

Of particular importance is judgement of eye gaze which is normally fast, accurate and automatic, operating at multiple levels of cognition through multiple cues [71, 80, 81, 80, 82, 83, 84].

Gaze in particular aids smooth turn passing [85] [86] and lack of support for eye gaze has been found to decrease the efficiency of turn passing by 25% [87].

There are clear patterns to eye gaze in groups, with the person talking, or being talked to, probably also being looked at [88] [89]. To facilitate this groups will tend to position themselves to maximally enable observation of the gaze of the other parties [83]. This intersects with proxemics which will be discussed shortly. In general people look most when they are listening, with short glances of 3-10 seconds [81]. Colburn et al. suggest that gaze direction and the perception of the gaze of others directly impacts social cognition [90] and this has been supported in a follow up study [91].

The importance of gaze is clearly so significant in evolutionary terms that human acuity for eye direction is considered high at 30 sec arc [92] with straight binocular gaze judged more accurately than straight monocular gaze [93], when using stereo vision.

Regarding the judgement of the gaze of others, Symons et al. suggesting that “people are remarkably sensitive to shifts in a person’s eye gaze” in triadic conversation [92]. This perception of the gaze of others operates at a low level and is automatic. Langton et al. cite

research stating that the gaze of others is “able to trigger reflexive shifts of an observer’s visual attention” and further discuss the deep biological underpinnings of gaze processing [89].

When discussing technology-mediated systems, Vertegaal & Ding suggested that understanding the effects of gaze on triadic conversation is “crucial for the design of teleconferencing systems and collaborative virtual environments” [67], and further found correlation between the amount of gaze, and amount of speech. Vertegaal & Slagter suggest that “gaze function(s) as an indicator of conversational attention in multiparty conversations” [88]. It seems like if we are to have useful markets within social immersive environments then support for natural gaze effects should be a priority.

Wilson et al. found that subjects can “discriminate gaze focused on adjacent faces up to [3.5m]” [94]. This perhaps gives us a testable benchmark within a metaverse application which is eye gaze enabled. In this regard Schrammel et al. investigated to what extent embodied agents can elicit the same responses in eye gaze detection [95].

Vertegaal et al. found that task performance was 46% better when gaze was synchronised in their telepresence scenario. As they point out, gaze synchronisation (temporal and spatial) is ‘commendable’ in all such group situations, but the precise utility will depend upon the task [67].

There has been some success in the automatic detection of the focus of attention of participants in multi party meetings [96, 97]. More recently, eye tracking technologies allow the recording and replaying of accurate eye gaze information [98] alongside information about pupil dilation toward determination of honesty and social presence [99]. It seems there are trust and honesty issues conflated with how collaborators in a virtual space are represented.

In summary, gaze awareness does not just mediate verbal communication but rather is a complex channel of communication in its own right. Importantly, gaze has a controlling impact on those who are involved in the communication at any one time, including and excluding even beyond the current participants. Perhaps the systems we propose in this book need to demand eye gaze support, but it is clear that it should be recommended, and that the software selected should support the technology integration in principle.

7.3.2.2 Mutual Gaze

Aygyle and Cook established early work around gaze and mutual gaze, with their seminal book of the same title [80], additionally detailing confounding factors around limitations and inaccuracies in observance of gaze and how this varies with distance [82, 71, 100].

Mutual gaze is considered to be the most sophisticated form of gaze awareness with significant impact on dyadic conversation especially [100, 73, 101]. The effects seem more profound than just helping to mediate flow and attention, with mutual eye gaze aiding in memory recall and the formation of impressions [102].

While reconnection of mutual eye gaze through a technology boundary does not seem completely necessary it is potentially important, with impact on subtle elements of one to one communication, and therefore discrimination of eye gaze direction should be bi-directional if possible, and if possible have sufficient accuracy to judge direct eye contact. In their review Bohannon et al. said that the issue of rejoining eye contact must be addressed in order to fully realise the richness of simulating face to face encounters [102].

Mutual gaze is a challenging affordance as bi-directional connection of gaze is not a trivial problem. It’s perhaps best to view this as at the ‘edge’ of our requirements for a

metaverse.

7.3.2.3 Head Orientation

Orientation of the head (judged by the breaking of bilateral symmetry and alignment of nose) is a key factor when judging attention. Perception of head orientation can be judged to within a couple of degrees [94].

It has been established that head gaze can be detected all the way out to the extremis of peripheral vision, with accurate eye gaze assessment only achievable in central vision [63]. This is less of use for our metaverses at this time, because user field of view is almost always restricted in such systems. More usefully, features of illumination can alter the apparent orientation of the head [103].

Head motion over head orientation is a more nuanced proposition and can be considered a micro gesture [104]. Head tracking systems within head mounted displays can certainly detect these tiny movements, but it's clear that not all of this resolution is passed into shared virtual settings through avatars. It would be beneficial to be able to fine tune this feature within any software selected.

It is possible that 3D displays are better suited to perception of head gaze since it is suggested that they are more suitable for "shape understanding tasks" [105]

Bailenson, Baell, and Blascovich found that giving avatars rendered head movements in a shared virtual environment decreased the amount of talking, possibly as the extra channel of head gaze was opened up. They also reported that subjectively, communication was enhanced [106].

Clearly head orientation is an important indicator of the direction of attention of members of a group and can be discerned even in peripheral vision. This allows the focus of several parties to be followed simultaneously and is an important affordance to replicate on any multi-party communication system.

7.3.2.4 Combined Head and Eye Gaze

Rienks et al. found that head orientation alone does not provide a reliable cue for identification of the speaker in a multiparty setting [107]. Stiefelhagen & Zhu found "that head orientation contributes 68.9% to the overall gaze direction on average" [97], though head and eye gaze seem to be judged interdependently [93]. Langton noted that head and eye gaze are "mutually influential in the analysis of social attention" [89], and it is clear that transmission of 'head gaze' by any mediating system, enhances rather than replaces timely detection of subtle cues. Combined head and eye gaze give the best of both worlds and extend the lateral field of view in which attention can be reliably conveyed to others [63].

7.3.2.5 Other Upper Body: Overview

While it is well evidenced that there are advantages to accurate connection of the gaze between conversational partners [82, 73], there is also a body of evidence that physical communication channels extend beyond the face [73, 108] and include both micro (shrugs, hands and arms), and macro movement of the upper body [109]. Goldin-Meadow suggests that gesturing aids conversational flow by resolving mismatches and aiding cognition [110].

In their technology-mediated experiment which compared face to upper body and face on a flat screen, Nguyen and Canny found that "upper-body framing improves empathy measures and gives results not significantly different from face-to-face under several empathy measures" [108].

The upper body can be broken up as follows:

Facial

Much emotional context can be described by facial expression (display) alone [109, 111], with smooth transition between expressions seemingly important [schiano2004]. This suggests that mediating technologies should support high temporal resolution, or at least that there is a minimum resolution between which transitions between expressions become too 'categorical'. Some aspects of conversational flow appear to be mediated in part by facial expression [ohba1998]. There are gender differences in the perception of facial affect [112].

Gesturing

(such as pointing at objects) paves the way for more complex channels of human communication and is a basic and ubiquitous channel [113]. Conversational hand gestures provide a powerful additional augmentation to verbal content [114].

Posture

Some emotions can be conveyed through upper body configurations alone. Argyle details some of these [71] and makes reference to the posture of the body and the arrangement of the arms (i.e. folded across the chest). These are clearly important cues. Kleinsmith and Bianchi-Berthouze assert that "some affective expressions may be better communicated by the body than the face" [115].

Body Torque

In multi-party conversation, body torque, that is the rotation of the trunk from front facing, can convey aspects of attention and focus [10.2307/40971262].

In summary, visual cues which manifest on the upper body and face can convey meaning, mediate conversation, direct attention, and augment verbal utterances.

7.3.2.6 Effect of Shared Objects on Gaze

Ou et al. detail shared task eye gaze behaviour "in which helpers seek visual evidence for workers' understanding when they lack confidence of that understanding, either from a shared, or common vocabulary" [116].

Murray et al. found that in virtual environments, eye gaze is crucial for discerning what a subject is looking at [117]. This work is shown in Figure 7.2.

It is established that conversation around a shared object or task, especially a complex one, mitigates gaze between parties [80] and this suggests the choice of task is an important factor when examining eye gaze in experimental groups.

7.4 Psychology of Technology-Mediated Interaction

7.4.1 Proxemics

Proxemics is the formal study of the regions of interpersonal space begun in the late 50's by Hall and Sommers and building toward The Hidden Dimension [118], which details bands of space that are implicitly and instinctively created by humans and which have a direct bearing on communication.

Distance between conversational partners, and affiliation, also have a bearing on the level of eye contact [81] with a natural distance equilibrium being established and developed throughout, through both eye contact and a variety of subtle factors. Argyle & Ingham provide levels of expected gaze and mutual gaze against distance [82]. These boundaries are altered by ethnicity [119, 71] and somewhat by gender [120], and age [121],

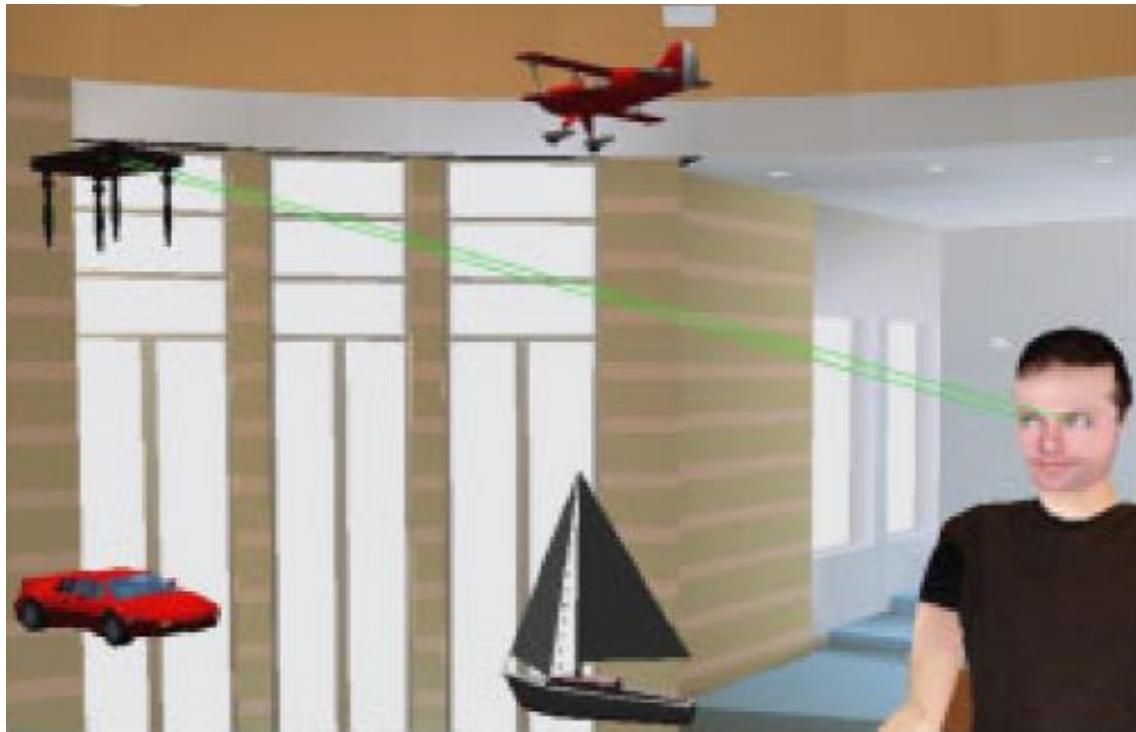


Figure 7.2: Eye tracked eye gaze awareness in VR. Murray et al. used immersive and semi immersive systems alongside eye trackers to examine the ability of two avatars to detect the gaze awareness of a similarly immersed collaborator.

[112].

Even with significant abstraction by communication systems (such as SecondLife) social norms around personal space persist [6790745, bailenson2003interpersonal, 122]. Bailenson & Blascovich found that even in Immersive Collaborative Virtual Environments (ICVE's) "participants respected personal space of the humanoid representation" [6790745] implying that this is a deeply held 'low-level' psychophysical reaction [123].

It is highly applicable to experimental design that the noted 'social space' be considered throughout.

Maeda et al. [124] found that seating position impacts the level of engagement in teleconferencing. Taken together with the potential for reconfiguration within the group as well as perhaps signaling for the attention of participants outside of the confines of the group.

When considering the attention of engaging with people outside the confines of a meeting Hager et al. found that gross expressions can be resolved by humans long distances [Argyle, 125]. It seems that social interaction begins around 7.5m in the so-called 'public space' [118].

7.4.2 Attention

The study of attention is a discrete branch of psychology. It is the study of cognitive selection toward a subjective or objective sub focus, to the relative exclusion of other stimuli. It has been defined as "a range of neural operations that selectively enhance processing of information" [carlston2013]. In the context of interpersonal communication it can be refined to apply to selectively favouring a conversational agent or object or task

above other stimuli in the contextual frame.

Humans can readily determine the focus of attention of others in their space [96] and preservation of the spatial cues which support this are important for technology-mediated communication [126] [97].

Symons et al. suggested that “people are remarkably sensitive to shifts in a person’s eye gaze” in triadic conversation [92]. Wilson et al. found that subjects can “discriminate gaze focused on adjacent faces up to [3.5m]” [94]

The interplay between conversational partners, especially the reciprocal perception of attention, is dubbed the perceptual crossing [127, 128].

This is a complex field of study with gender, age, and ethnicity all impacting the behaviour of interpersonal attention [Hofmann20061683, Pan2008a, 129, 121, 71]. Verte-gaal has done a great deal of work on awareness and attention in technology-mediated situations and the work of his group is cited throughout the thesis [130]. As an example it is still such a challenge to “get” attention through mediated channels of communication, that some research [131, 126] and many commercial systems such as ‘blackboard collaborate’, use tell tale signals (such as a microphone icon) to indicate when a participant is actively contributing. Some are automatic, but many are still manual, requiring that a user effectively hold up a virtual hand to signal their wish to communicate.

Langton et al. cite research stating that the gaze of others is “able to trigger reflexive shifts of an observer’s visual attention”. Novick et al. performed analysis on task hand-off gaze patterns which is useful for extension into shared task experiment design [86].

Attention and the sensing of attention rely on the reconnection of non-verbal cues across the whole of the space in which communication occurs.

Regarding the attention of others, Fagal et al demonstrated that eye visibility impacts collaborative task performance when considering a shared task [101].

7.4.3 Behaviour

Hedge et al. suggested that gaze interactions between strangers and friends may be different which could have an impact on research recruitment [85]. Voida et al. elaborate that prior relationships can cause “internal fault lines” in group working [132]. When new relationships are formed the “primary concern is one of uncertainty reduction or increasing predictability about the behavior of both themselves and others in the interaction [berger1975]. This concept of smoothness in the conversation is a recurring theme, with better engineered systems introducing less extraneous artefacts into the communication, and so disturbing the flow less. In a similar vein the actor-observer effect describes the mismatch between expectations which can creep into conversation. Conversations mediated by technology can be especially prone to diverging perceptions of the causes of behaviour [133]. Interacting subjects progress conversation through so-called ‘perception-action’ loops which are open to predictive modelling through discrete hidden Markov models [134].

It may be that the perception-behaviour link where unconscious mirroring of posture bolsters empathy between conversational partners, especially when working collaboratively [135], and the extent to which posture is represented through a communication medium may be important.

Landsberger posited the Hawthorne effect [Parsons922]. Put simply this is a short term increase in productivity that may occur as a result of being watched or appreciated. This has broader implications for experimentation where tasks are observed or are perceived to

be observed.

The impression of being watched changes gaze patterns during experimentation, with even implied observation through an eye tracker modifying behaviour [[risko2011eyes](#)].

There is much historic work describing “the anatomy of cooperation” [[136](#)] and this may be useful in analysing experimental data around common collaborative tasks.

Cuddihy and Walters defined an early model for assessing desktop interaction mechanisms for social virtual environments [[137](#)].

7.4.3.1 Perception Of Honesty

Hancock et al. state that we are most likely to lie, and to be lied to, on the telephone [[138](#)]. Technology used for communication impacts interpersonal honesty. It seems that at some level humans know this; lack of eye contact leads to feelings of deception, impacting trust [[139](#)]. This has a major impact on video conferencing which often does not support mutual gaze. Trust is crucial for business interactions.

Further there are universal expressions, micro-expressions, and blink rate which can betray hidden emotions [[porter2008reading](#)], though the effects are subtle and there is a general lack of awareness by humans of their abilities in this regard [[139](#)]. Absence of support for such instinctive cues inhibits trust [[140](#)]. Support for these rapid and transient facial features demands high resolution reproduction in both resolution and time domains. There is detectable difference in a participant’s ability to detect deception when between video conference mediated communication and that mediated by avatars [[Steptoe2010b](#)]. Systems should aim for maximally faithful reproduction.

7.4.4 Dyadic

7.4.4.1 Point to Point Telephony (Audio only)

The ubiquitous technology to mediate conversation is, of course, the telephone. The 2015 Ericsson mobility report [[2015](#)] states that there are a staggering 7.4 billion mobile subscriptions. More people have access to mobile phones than to working toilets. Jouppi and Pan designed a system which focused attention on spatially correct high definition audio. They found “significant improvement over traditional audio conferencing technology, primarily due to the increased dynamic range and directionality. [[jouppi2002mutually](#)]. Aoki et al. also describe an audio only system with support for spatial cues [[141](#)].

7.4.4.2 Point to Point Video Conferencing

O’Malley et al. showed that face-to-face and video mediated employed visual cues for mutual understanding, and that addition of video to the audio channel aided confidence and mutual understanding. However, video mediated did not provide the clear cues of being co-located [[O’Malley1996](#)].

Dourish et al. make a case for not using face to face as a baseline for comparison, but rather that analysis of the efficacy of remote tele-collaboration tools should be made in a wider context of connected multimedia tools and ‘emergent communicative practises’ [[Dourish1993](#)]. While this is an interesting viewpoint it does not necessarily map well to a recreation of the ad-hoc meeting.

There is established literature on human sensitivity to eye contact in both 2D and 3D VC [[Eijk2010](#), [142](#)], with an accepted minimum of 5-10 degrees before observers can reliably sense they are not being looked at [[70](#)]. Roberts et al. suggested that at the limit of social gaze distance (4m) the maximum angular separation between people standing

shoulder to shoulder in the real world would be around 4 degrees[143].

Sellen found limited impact on turn passing when adding a visual channel to audio between two people when using Hydra, an early system which provided multiple video conference displays in an intuitive spatial distribution[Sellen1992a]. She did however, find that the design of the video system affected the ability to hold multi-party conversations [Sellen1995a].

Monk and Gale describe in detail experiments which they used for examining gaze awareness in communication which is mediated and unmediated by technology. They found that gaze awareness increased message understanding [84].

Both Kuster et al. and Gemmel et al. have successfully demonstrated software systems which can adjust eye gaze to correct for off axis capture in real time video systems[Kuster2012b, 144].

Shahid et al. conducted a study on pairs of children playing games with and without video mediation and concluded that the availability of mutual gaze affordance enriched social presence and fun, while its absence dramatically affects the quality of the interaction. They used the ‘Networked Minds’, a social presence questionnaire.

7.4.5 Triadic and Small Group

Early enthusiasm in the 1970’s for video conferencing, as a medium for small group interaction quickly turned to disillusionment. It was agreed after a flurry of initial research that the systems at the time offered no particular advantage over audio only communication, and at considerable cost [145].

Something in the breakdown of normal visual cues seems to impact the ability of the technology to support flowing group interaction. Nonetheless, some non-verbal communication is supported in VC with limited success.

Additional screens and cameras can partially overcome the limitation of no multi-party support (that of addressing a room full of people on a single screen) by making available more bidirectional channels. For instance, every remote user can be a head on a screen with a corresponding camera. The positioning of the screens must then necessarily match the physical organization of the remote room.

Egido provides an early review of the failure of VC for group activity, with the “misrepresentation of the technology as a substitute for face to face” still being valid today [146].

Commercial systems such as Cisco Telepresence Rooms cluster their cameras above the centre screen of three for meetings using their telecollaboration product, while admitting that this only works well for the central seat of the three screens. They also group multiple people on a single screen in what Workhoven et al. dub a “non-isotropic” configuration [147]. They maintain that this is a suitable trade off as the focus of the meeting is more generally toward the important contributor in the central seat [CiscoSystems]. This does not necessarily follow for less formal meeting paradigms.

In small groups, it is more difficult to align non-verbal cues between all parties, and at the same time, it is more important because the hand-offs between parties are more numerous and important in groups. A breakdown in conversational flow in such circumstances is harder to solve. A perception of the next person to talk must be resolved for all parties and agreed upon to some extent.

However, most of the conventional single camera, and expensive multi camera VC systems, suffer a fundamental limitation in that the offset between the camera sight lines

and the lines of actual sight introduce incongruities that the brain must compensate for [58].

7.4.6 Presence, Co-presence, and Social Presence

Presence is a heavily cited historic indicator of engagement in virtual reality, though the precise meaning has been interpreted differently by different specialisms [148, 149]. It is generally agreed to be the 'sense of being' in a virtual environment [150]. Slater extends this to include the "extent to which the VE becomes dominant".

Beck et al. reviewed 108 articles and synthesised an ontology of presence [148] which at its simplest is as follows:

1. Sentient presence
 - a. Physical interaction
 - b. Mental interaction
2. Non-sentient
 - a. Physical immersion
 - b. Mental immersion = psychological state

When presence is applied to interaction it may be split into Telepresence, and Co/Social presence [heeter1992being, 151]. Co-presence and/or social presence is the sense of "being there with another", and describes the automatic responses to complex social cues [doi:10.1080/01449299508914633, fulk1987social, haythornthwaite1995work]. Social presence (and co-presence) refers in this research context to social presence which is mediated by technology (even extending to text based chat [152]), and has its foundations in psychological mechanisms which engender mutualism in the 'real'. This is analysed in depth by Nowak [153]. An examination of telepresence, co-presence and social presence necessarily revisits some of the knowledge already elaborated.

The boundaries between the three are blurred in research with conflicting results presented [154]. Biocca et al. attempted to enumerate the different levels and interpretations surrounding these vague words [155], and to distill them into a more robust theory which better lends itself to measurement. They suggest a solid understanding of the surrounding psychological requirements which need support in a mediated setting, and then a scope that is detailed and limited to the mediated situation.

Since 'social presence' has been subject to varied definitions [155] it is useful here to consider a single definition from the literature which defines it as "the ability of participants in the community of inquiry to project their personal characteristics into the community, thereby presenting themselves to the other participants as real people." [156, 148]. Similarly to specifically define co-presence for this research it is taken to be the degree to which participants in a virtual environment are "accessible, available, and subject to one another" [155].

Social presence has received much attention and there are established questionnaires used in the field for measurement of the levels of perceived social presence yet the definitions here also remain broad, with some confusion about what is being measured [Biocca2003a].

Telepresence meanwhile is interaction with a different (usually remote) environment which may or may not be virtual, and may or may not contain a separate social/co-presence component.

Even in simple videoconferencing Bondareva and Bouwhuis stated (as part of an experimental design) that the following determinants are important to create social presence

[[jouppi2002mutually](#), [157](#)].

1. Direct eye contact is preserved
2. Wide visual field
3. Both remote participants appear life size
4. Possibility for participants to see the upper body of the interlocutor
5. High quality image and correct colour reproduction
6. Audio with high S/N ratio
7. Directional sound field
8. Minimization of the video and audio signal asynchrony
9. Availability of a shared working space.

Bondareva et al. went on to describe a person-to-person telepresence system with a semi-silvered mirror to reconnect eye gaze, which they claimed increased social presence indicators. Interestingly they chose a checklist of interpersonal interactions which they used against recordings of conversations through the system [[157](#)].

The idea of social presence as an indicator of the efficacy of the system, suggests the use of social presence questionnaires in the evaluation of the system [[155](#)]. Subjective questionnaires are however troublesome in measuring effectiveness of virtual agents and embodiments, with even nonsensical questions producing seemingly valid results [[158](#)]. Usoh et al. found that 'the real' produced only marginally higher presence results than the virtual [[Usoh2000a](#)].

Nowak states that "A satisfactory level of co-presence with another mind can be achieved with conscious awareness that the interaction is mediated" and asserts that while the mediation may influence the degree of co-presence it is not a prohibiting factor [[153](#)].

Baren and IJsselsteijn [[Baren, 159](#)] list 20 useful presence questionnaires in 2004 of which "Networked Minds" seemed most appropriate for the research. Hauber et al. employed the "Networked Minds" Social Presence questionnaire experimentally and found that while the measure could successfully discriminate between triadic conversation that is mediated or unmediated by technology, it could not find a difference between 2D and 3D mediated interfaces [[160, 152](#)].

In summary, social presence and co-presence are important historic measures of the efficacy of a communication system. Use of the term in literature peaked between 1999 and 2006 according to Google's ngram viewer and has been slowly falling off since. The questionnaire methodology has been challenged in recent research and while more objective measurement may be appropriate, the networked minds questions seem to be able to differentiate real from virtual interactions [[159](#)].

7.4.7 Other Systems to Support Business

There have been many attempts to support group working and rich data sharing between dispersed groups in a business setting. So called 'smart spaces' allow interaction with different displays for different activities and add in some ability to communicate with remote or even mobile collaborators on shared documents [[161](#)], with additional challenges for multi-disciplinary groups who are perhaps less familiar with one or more of the technology barriers involved [[162](#)].

Early systems like clearboard [[163](#)] demonstrated the potential for smart whiteboards with a webcam component for peer to peer collaborative working. Indeed it is possible to support this modality with Skype and a smartboard system (and up to deployments such as Accessgrid). They remain relatively unpopular however.

7.4.8 Mutual Gaze in Telepresence

We have seen that transmission of attention can broadly impact communication in subtle ways, impacting empathy, trust, cognition, and co-working patterns. Mutual gaze (looking into one another's eyes), is currently the high water mark for technology-mediated conversation.

Many attempts have been made to re-unite mutual eye gaze when using tele-conferencing systems. In their 2015 review of approaches Regenbrecht and Langlotz found that none of the methods they examined were completely ideal [164]. They found most promise in 2D and 3D interpolation techniques, which will be discussed in detail later, but they opined that such systems were very much ongoing research and lacked sufficient optimisation.

A popular approach uses the so called 'Peppers Ghost' phenomenon [[steinmeyer2013science](#)], where a semi silvered mirror presents an image to the eye of the observer, but allows a camera to view through from behind the angled mirror surface. The earliest example of this is Rosenthal's two way television system in 1947 [[rosenthal1947two](#)], though Buxton et al. 'Reciprocal Video Tunnel' from 1992 is more often cited [[buxton1992telepresence](#)]. This optical characteristic isn't supported by retroreflective projection technology, and besides requires careful control of light levels either side of the semi-silvered surface.

The early GAZE-2 system (which makes use of Pepper's ghost) is novel in that it uses an eye tracker to select the correct camera from several trained on the remote user. This ensures that the correct returned gaze (within the ability of the system) is returned to the correct user on the other end of the network [[Vertegaal2003](#)]. Mutual gaze capability is later highlighted as an affordance supported or unsupported by key research and commercial systems.

7.4.9 Other Systems to Support Business

There have been many attempts to support group working and rich data sharing between dispersed groups in a business setting. So called 'smart spaces' allow interaction with different displays for different activities and add in some ability to communicate with remote or even mobile collaborators on shared documents [161], with additional challenges for multi-disciplinary groups who are perhaps less familiar with one or more of the technology barriers involved [162].

Early systems like clearboard [163] demonstrated the potential for smart whiteboards with a webcam component for peer-to-peer collaborative working. Indeed it is possible to support this modality with Skype and a smartboard system (and up to deployments such as Accessgrid). They remain relatively unpopular however.

7.4.9.1 Ad-hoc in SecondLife, immersive VR

something about the bitcoin VR meetups and other ad-hoc groupings that are emerging.

7.4.9.2 Tabletop and Shared Task

In early telepresence research Buxton and William argued through examples that "effective telepresence depends on quality sharing of both person and task space [[Buxton1992a](#)].

In their triadic shared virtual workspace Tang et al. found difficulty in reading shared text using a 'round the table' configuration, a marked preference for working collaboratively on the same side of the table. They also found additional confusion as to the identity of remote participants [165]. Tse et al. found that pairs can work well over a shared digital tabletop, successfully overcoming a single user interface to interleave tasks [166].

Tang et al. demonstrate that collaborators engage and disengage around a group activity through several distinct, recognizable mechanisms with unique characteristics [167]. They state that tabletop interfaces should offer a variety of tools to facilitate this fluidity.

Camblend is a shared workspace with panoramic high resolution video. It maintains some spatial cues between locations by keeping a shared object in the video feeds [Norris2013a, 168]. Participants successfully resolved co-orientation within the system.

The t-room system implemented by Luff et al. surrounds co-located participants standing at a shared digital table with life sized body and head video representations of remote collaborators [169] but found that there were incongruities in the spatial and temporal matching between the collaborators which broke the flow of conversation. Tuddenham et al. found that co-located collaborators naturally devolved 'territory' of working when sharing a task space, and that this did not happen the same way with a tele-present collaborator [tuddenham2009territorial]. Instead remote collaboration adapted to use a patchwork of ownership of a shared task. It seems obvious to say that task ownership is a function of working space, but it is interesting that the research found no measurable difference in performance when the patchwork coping strategy was employed.

The nature of a shared collaborative task and/or interface directly impacts the style of interaction between collaborators. This will have a bearing on the choice of task for experimentation [Jetter2011a, 170].

7.4.10 Beyond 2D Screens

Displays need not be limited to 2 dimensional screens and can be enhanced in various ways.

Stereoscopy allows an illusion of depth to be added to a 2D image by exploiting the stereo depth processing characteristics of the human vision system. This technical approach is not perfect as it does not fully recreate the convergence and focus expected by the eyes and brain.

There are multiple approaches to separating the left and right eye images, these primarily being active (where a signal selectively blanks the input to left then right eyes in synchronicity with the display), passive, where either selective spectrum or selective polarisation of light allow different portions of a display access to different eyes, or physical arrangements which present different displays (or slices of light as in lenticular systems) to different eyes.

These barrier stereoscopy / lenticular displays use vertical light barriers built into the display to create multiple discrete channels of display which are accessed by moving horizontally with respect to the display. In this way it is possible to generate either a left/right eye image pair for 'autostereoscopic' viewing, or with the addition of head tracking and small motors. With these techniques multiple viewpoint or an adaptive realtime viewpoint update can be presented without the glasses required for active or passive stereoscopic systems.

7.4.10.1 Spatially Faithful Group

Hauber et al. combined videoconferencing, tabletop, and social presence analysis and tested the addition of 3D. They found a nuanced response when comparing 2D and 3D approaches to spatiality: 3D showed improved presence over 2D (chiefly through gaze support), while 2D demonstrated improved task performance because of task focus [171].

I3DVC reconstructs participants from multiple cameras and places them isotropically

(spatially faithful) [Kauff2002a, 172]. The system uses a large projection screen, a custom table, and carefully defined seating positions. They discussed an “extended perception space” which used identical equipment in the remote spaces in a tightly coupled collaborative ‘booth’. It employed head tracking and multi camera reconstruction alongside large screens built into the booth. This system exemplified the physical restrictions which are required to limit the problems of looking into another space through the screen. Fuchs et al. demonstrated a similar system over a wide area network but achieved only limited resolution and frame rate with the technology of the day [Fuchs:2002ww].

University of Southern California used a technically demanding real-time set-up with 3D face scanning and an autostereoscopic 3D display to generate multiple ‘face tracked’ viewpoints [jones2009headspin]. This had the disadvantage of displaying a disembodied head.

MAJIC is an early comparable system to support small groups with life size spatially correct video, but without multiple viewpoints onto the remote collaborators it was a one to ‘some’ system rather than ‘some’ to one. Additionally users were rooted to defined locations [Okada:1994et, 173].

Multiview In order to reconnect directional cues of all kinds it is necessary for each party in the group to have a spatially correct view of the remote user which is particular for them. This requires a multi-view display, which has applications beyond telepresence but are used extensively in research which attempts to address these issues.

Nguyen and Canny demonstrated the ‘Multiview’ system [66]. Multiview is a spatially segmented system, that is, it presents different views to people standing in different locations simultaneously. They found similar task performance in trust tasks to face to face meetings, while a similar approach without spatial segmentation was seen to negatively impact performance.

In addition to spatial segmentation of viewpoints [gotsch2018telehuman2] it is possible to isolate viewpoints in the time domain. Different tracked users can be presented with their individual view of a virtual scene for a few milliseconds per eye, before another viewpoint is shown to another user. Up to six such viewpoints are supported in the c1x6 system [174] Similarly MM+Space offered 4 Degree-Of-Freedom Kinetic Display to recreate Multiparty Conversation Spaces [175]

7.4.10.2 Robots, Shader Lamp, and Hybrid

Virtuality human representation extends beyond simple displays into robotic embodiments (which need not be humanoid [176]), shape mapped projection dubbed “shader lamps”, and hybridisations of the two.

Uncanniness

When employing simulation representations of humans it may be the case that there is an element of weirdness to some of these systems, especially those that currently represent a head without a body. Mori has demonstrated The Uncanny Valley [177] effect in which imperfect representations of humans elicit revulsion in certain observers. This provides a toolkit for inspecting potentially ‘weird’ representations, especially if they are ‘eerie’ and is testable through Mori’s GODSPEED questionnaire.

With an improved analysis of the shape of the likeability curve estimated later showing a more nuanced response from respondents where anthropomorphism of characters demonstrated increased likeability even against a human baseline [Bartneck2009a, 178].

A mismatch in the human realism of face and voice also produces an Uncanny Valley

response [179].

However, there is a possibility that Mori's hypothesis may be too simplistic for practical everyday use in CG and robotics research since anthropomorphism can be ascribed to many and interdependent features such as movement and content of interaction [180].

Bartneck et al. also performed tests which suggest that the original Uncanny Valley assertions may be incorrect, and that it may be inappropriate to map human responses to human simulacrum to such a simplistic scale. They suggest that the measure has been a convenient 'escape route' for researchers [180]. Their suggestion that the measure should not hold back the development of more realistic robots holds less bearing for the main thrust of this telepresence research which seeks to capture issues with imperfect video representation rather than test the validity of an approximation.

Interestingly Ho et al. performed tests on a variety of facial representations using images. They found that facial performance is a 'double edged sword' with realism being important to robotic representations, but there also being a significant Uncanny Valley effect around 'eerie, creepy, and strange' which can be avoided by good design [181].

More humanlike representations exhibiting higher realism produce more positive social interactions when subjective measures are used [Yee2007a] but not when objective measures are used. This suggests that questionnaires may be more important when assessing potential uncanniness.

Ho et al. also identified problems with the original GODSPEED indices used to measure Uncanny Valley effects. They proposed a new set of measures which they found to be generally valid [Ho:2010:RUV:1853385.1853509], though they admit they were only tested with a single set of stimuli.

A far more objective method would be to measure user responses to humans, robots, and representations with functional near-infrared spectroscopy and while this has been attempted it is early exploratory research [182], an emotional response to 'eerie' was discovered.

Embodiment through robots

Robots which carry a videoconference style screen showing a head can add mobility and this extends the available cues [Lee2011b, 183, 184, 185, 186]. Interestingly Desai and Uhlik maintain that the overriding modality should be high quality audio [187].

Tsui et al. asked 96 participants to rate how personal and interactive they found interfaces to be. Interestingly they rated videoconferencing as both more personal and more interactive than telepresence robots, suggesting that there is a problem with the overall representation or embodiment [188].

Kristoffersson et al. applied the Networked Minds questionnaire to judge presence of a telepresence robot for participants with little or no experience of videoconferencing. Their results were encouraging, though they identified that the acuity of the audio channel needing improvement [189].

There are a very few lifelike robots which can be used for telepresence, and even these are judged to be uncanny [190]. This is only an issue for a human likeness since anthropomorphic proxies such as robots and toys perform well [177].

Physical & Hybrid embodiment

Embodiment through hybridisation of real-time video and physical animatronic mannequins has been investigated as a way to bring the remote person into the space in a more convincing way [Lincoln:2010it, Raskar2001a, 191]. These include telepresence robots [Lee2011b, 190, 184], head in a jar implementations such as SphereAvatar

[pan2014comparing, Pan2012a, 192] and BiReality [193], UCL's Gaze Preserving Situated Multi-View Telepresence System [Pan2014a], or screen on a stick style representations [186].

Nagendran et al. present a 3D continuum of these systems into which they suggest all such systems can be rated from artificial to real on the three axes, shape, intelligence, and appearance [Nagendran].

Itoh et al. describe a 'face robot' to convey captured human emotion over a distance. It uses an 'average face' and actuators to manipulate feature points [194]. It seems that this is an outlier method for communication of facial affect but demonstrates that there are many development paths to a more tangible human display.

Shader lamps

Projection mapping is a computational augmented projection technique where consideration of the relative positions and angles of complex surfaces allows the projection from single or multiple sources to augment the physical shapes onto which they appear. It was first considered by the Disney corporation in 1969 [projectionmappingweb2013] and was given prominence by Raskar and Fuchs with "office of the future" [195] and later by Raskar and other researchers [Low2001, Raskar2001a]. It has since gained considerable commercial popularity in live entertainment [googleStatsProjectionMapping].

Shader lamps [Raskar2001a] is the more formal academic designation for projection mapping. It is possible to use the technique alongside reconstruction to project onto a white facial mannequin. Researchers have attempted to use the technology for remote patient diagnostic, projecting onto styrofoam heads [196].

Bandyopadhyay et al. demonstrated [Bandyopadhyay] that it is possible to track objects and projection map [197] onto them in real time. This is beyond the scope of the proposed projection onto furniture since we wish to keep the system as simple as possible, but could be useful for shared tasks in the future work.

Lincoln et al. employed animatronic avatars which they projected with shader lamps. This combination recreated facial expression and head movement though they were limited in speed and range of control of the remote head [Lincoln:2010it].

While shader lamps are an important and useful technology, there are limitations imposed by its use. In particular if a realtime video feed or reconstruction of a subject is used then that scanned subject must either remain still enough to be correctly mapped onto geometry on the remote side (useful for some virtual patients for instance [198], or else there must be a computational adjustment made for their changing position to make them appear static, or the projection surface must move to match their movement as in Lincoln et al. .

7.4.10.3 Holography and Volumetric

Blanche et al. have done a great deal of research into holographic and volumetric displays using lasers, rotating surfaces, and light field technology [tay2008updatable, 199]. They are actively seeking to use their technologies for telepresence and their work is very interesting.

Similarly Jones et al. "HeadSPIN" is a one-to-many 3D video teleconferencing system [jones2009headspin] which uses a rotating display to render the holographic head of a remote party. They achieve transmissible and usable framerate using structured light scanning of a remote collaborator as they view a 2D screen which they say shows a spatially correct view of the onlooking parties.

Eldes et al. used a rotating display to present multi-view autostereoscopic projected images to users [eldes2013multi].

Seelinder is an interesting system which uses parallax barriers to render a head which an onlooking viewer can walk around. The system uses 360 high resolution still images which means a new spatially segmented view of the head every 1 degrees of arc. They claim the system is capable of playback of video and this head in a jar multi-view system clearly has merit but is comparatively small, and as yet untested for telepresence [200].

These systems do not satisfy the requirement to render upper body for the viewers and are not situated (as described soon).

7.4.10.4 Reconstructed Viewpoint

Although it is possible to manipulate a 2D view of a remote participant to bring the camera, eye, and screen into correct alignment better results can be obtained by capturing a 3D representation of the remote user, then modifying that. Several groups have investigated stereo reconstruction of face and upper body, so-called immersive videoconferencing [201].

Multiple viewpoints from multiple cameras allow either playback from different cameras on a continuous basis, so-called free viewpoint TV [202]. This multiple camera approach can be extended with the use of light field to interpolate between real or virtual cameras [203, 204], or else algorithmic generation and display of a polygonal mesh through photogrammetry can be undertaken. This can now be performed fast enough to compute a continuous surface from a video stream with little latency [Criminisi:2003ji]. Engineers from Google presented an efficient method for network transmission of light field scenes at Siggraph [broxton2020immersive]. These multi camera systems aim to create geometry for a single virtual viewpoint (such as a face in video conferencing), which is correct to align eye gaze. This is sometimes called “3D video”, which is different to stereoscopic video used in cinema and commercial TV.

A depth map, point cloud, then polygon mesh can be calculated for one side of a person from two adjacent cameras which have slightly different views of a subject. This so-called ‘narrow baseline’ technique examines sub-pixel disparities between the images [205].

Multiple ‘narrow baseline’ pairs can create stereo geometry reliably from a tightly coupled pair of cameras which are integrated with geometry from another pair with an independent viewpoint [Okutomi:1993ez, Cooke2002a, 206].

Alternatively there is the so-called “wide baseline” approach in which a pair of cameras with a significant distance between them capture both sides of the face or body. Sadagic et al. demonstrated a system which brought 2 camera reconstructed [207] people into a desk based collaborative space. The user addressing these two tele-present collaborators could move around in their chair while being presented with both passive stereo depth cues, and parallax cues appropriate to the desk environment. The system’s main limitation was the necessity for the user who addressed the two desk screens to wear both stereo glasses and a camera on their head. This made their system unbalanced. There are also problems with geometry and texture management using a wide baseline approach since the cameras are resolving for different lighting and views of the scene.

Pan, Steptoe and Steed found similar results with their spherical display with a decrease in trust toward avatar mediated conversation when viewing 2D displays at oblique angles [pan2014comparing].

Microsoft Kinect sensors demonstrate effectiveness in real-time scanning with between 1 and 5 deployed in research systems [208]. Multiple Kinect v1 sensors interfere with

one another, and though this problem is partially addressed with v2 hardware there are constraints with temperature drift causing frequency shifts which allow interference to creep in. There are workarounds for this issue but no completely reliable implementation.

For some time it has been possible to create geometric models of the human form using a technique called “shape from silhouette” [Petit:2008tva, Laurentini1994a, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219]

Similar effect can also be accomplished for the human shape using depth cameras such as the Microsoft Kinect v2 [220], and indeed this seems now to be the prevailing method. Sensors in this vein continue to improve.

Structured light systems provide still another technique, back as early as the 1980’s [Hu:1986ww], and as recently as the Microsoft Kinect v1 sensor. Deformations in projected visible or infrared patterns can be compared to the known baseline and shape reverse engineered from the differences. 3D-Board is an interesting example since two standing collaborators can interact through touch with a digital whiteboard which also forms the perceived barrier between them [221].

Rendering and telepresence

Cooke et al. investigate the difference between so-called narrow and wide baseline camera capture for dealing with the more complicated elements of reconstructed geometry such as hands which gesture. They maintain that multiple stereo camera pairs (each narrow baseline) arranged as a wide baseline system, with post processing to remove artifacts is of most use [Cooke_image-basedrendering].

Petit et al. adopt a wide baseline system with conventional green screen background segmentation and get good 3D video results. Because they are using a chroma based technique for their silhouette system they cannot exclude objects in the scene by ‘training’ the system [Petit2010].

Kuster et al. use a depth based approach and have a compelling demonstrator which uses what they call an “anisotropic transparent background back projection foil”. They rear project a Kinect reconstructed head and upper body in stereoscopic 3D onto a transparent film suspended so as to appear to bring the remote person to the edge of a table. This brings the 3D avatar into a space at 15 frames per second across significant network distances but is again peer to peer[Kuster2012a].

Similarly Maimone and Fuchs use large tiled displays and Kinect sensors but attempt to ‘merge’ two spaces by scanning both the person and the remote space to create perspective correct viewing through a virtual window into the other space. Without multi-view this system is one to one, although they demonstrate multiple people in the background of the space [Maimone2011a].

The Fusion4D system successfully captured complex and challenging scenes such as dancers with flowing clothing for playback from arbitrary angles [222].

withyou is an experimental capture and playback system which uses the Octave multimodal suite. This system uses wide baseline single cameras and CPU based shape from silhouette reconstruction alongside a novel network transport system to send a full 3D video polygonal hull to another rendering location [140]. Previous tests on the system suggested that the capture and playback made it possible to judge the eye gaze of the reconstructed subjects to within normal social/biological limits [143].

A particularly well developed example is the blue-c system which enables telepresence through multi-camera 3D video. The users at multiple locations are scanned and transmitted as a point cloud captured during projector blanking frames [223]. The system is flexible

enough to use either CAVE style surround or a projected screen style arrangement.

The ViPiD system interestingly uses Chromatte in its intended role to improve segmentation for their capture system [224].

Schreer et al demonstrate a capable multi camera capture system suitable for small group telepresence, though its technical overhead is high [Schreer:2008ty]. They discuss the need for multi-view capable screens only insomuch as they agree one must be developed for the system.

Maimone et al demonstrate a system uses a HMD system to project a reconstructed person into the correct seat at a table in augmented reality [225].

A depth map, point cloud, then polygon mesh can be calculated for one side of a person from two adjacent cameras which have slightly different views of a subject. This so-called ‘narrow baseline’ technique examines sub-pixel disparities between the images [205].

Multiple ‘narrow baseline’ pairs can create stereo geometry reliably from a tightly coupled pair of cameras which are integrated with geometry from another pair with an independent viewpoint [Okutomi:1993ez, Cooke2002a, 206].

Alternatively there is the so-called “wide baseline” approach in which a pair of cameras with a significant distance between them capture both sides of the face or body. Sadagic et al. demonstrated a system which brought 2 camera reconstructed [207] people into a desk based collaborative space. The user addressing these two tele-present collaborators could move around in their chair while being presented with both passive stereo depth cues, and parallax cues appropriate to the desk environment. The system’s main limitation was the necessity for the user who addressed the two desk screens to wear both stereo glasses and a camera on their head. This made their system unbalanced. There are also problems with geometry and texture management using a wide baseline approach since the cameras are resolving for different lighting and views of the scene.

Pan, Steptoe and Steed found similar results with their spherical display with a decrease in trust toward avatar mediated conversation when viewing 2D displays at oblique angles [pan2014comparing].

Microsoft Kinect sensors demonstrate effectiveness in real-time scanning with between 1 and 5 deployed in research systems [208]. Multiple Kinect v1 sensors interfere with one another, and though this problem is partially addressed with v2 hardware there are constraints with temperature drift causing frequency shifts which allow interference to creep in. There are workarounds for this issue but no completely reliable implementation.

For some time it has been possible to create geometric models of the human form using a technique called “shape from silhouette” [Petit:2008tva, Laurentini1994a, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219]

Similar effect can also be accomplished for the human shape using depth cameras such as the Microsoft Kinect v2 [220], and indeed this seems now to be the prevailing method. Sensors in this vein continue to improve.

Structured light systems provide still another technique, back as early as the 1980’s [Hu:1986ww], and as recently as the Microsoft Kinect v1 sensor. Deformations in projected visible or infrared patterns can be compared to the known baseline and shape reverse engineered from the differences. 3D-Board is an interesting example since two standing collaborators can interact through touch with a digital whiteboard which also forms the perceived barrier between them [221].

Rendering and telepresence

Cooke et al. investigate the difference between so-called narrow and wide baseline

camera capture for dealing with the more complicated elements of reconstructed geometry such as hands which gesture. They maintain that multiple stereo camera pairs (each narrow baseline) arranged as a wide baseline system, with post processing to remove artifacts is of most use [**Cooke_image-basedrendering**].

Petit et al. adopt a wide baseline system with conventional green screen background segmentation and get good 3D video results. Because they are using a chroma based technique for their silhouette system they cannot exclude objects in the scene by ‘training’ the system [**Petit2010**].

Kuster et al. use a depth based approach and have a compelling demonstrator which uses what they call an “anisotropic transparent background back projection foil”. They rear project a Kinect reconstructed head and upper body in stereoscopic 3D onto a transparent film suspended so as to appear to bring the remote person to the edge of a table. This brings the 3D avatar into a space at 15 frames per second across significant network distances but is again peer-to-peer [**Kuster2012a**].

Similarly Maimone and Fuchs use large tiled displays and Kinect sensors but attempt to ‘merge’ two spaces by scanning both the person and the remote space to create perspective correct viewing through a virtual window into the other space. Without multi-view this system is one to one, although they demonstrate multiple people in the background of the space [**Maimone2011a**].

The Fusion4D system successfully captured complex and challenging scenes such as dancers with flowing clothing for playback from arbitrary angles [222].

withyou is an experimental capture and playback system which uses the Octave multimodal suite. This system uses wide baseline single cameras and CPU based shape from silhouette reconstruction alongside a novel network transport system to send a full 3D video polygonal hull to another rendering location [140]. Previous tests on the system suggested that the capture and playback made it possible to judge the eye gaze of the reconstructed subjects to within normal social/biological limits [143].

A particularly well developed example is the blue-c system which enables telepresence through multi-camera 3D video. The users at multiple locations are scanned and transmitted as a point cloud captured during projector blanking frames [223]. The system is flexible enough to use either CAVE style surround or a projected screen style arrangement.

The ViPiD system interestingly uses Chromatte in its intended role to improve segmentation for their capture system [224].

Schreer et al demonstrate a capable multi camera capture system suitable for small group telepresence, though its technical overhead is high [**Schreer:2008ty**]. They discuss the need for multi-view capable screens only insomuch as they agree one must be developed for the system.

Maimone et al demonstrate a system uses a HMD system to project a reconstructed person into the correct seat at a table in augmented reality [225].

7.4.10.5 Immersive Collaborative Virtual Environments

Simple online virtual environments with an external viewpoint and many avatars interacting online generated much hype in the early days as a potential means to support group meetings, even back as far as the VRML standard [226]. SecondLife in particular generated significant interest. Erickson et al. went so far as to say that a virtual conference hosted in SecondLife was ‘fairly successful’ [227]. For whatever reason however these systems have fallen out of favour with the public and research communities. It might be that the external

viewpoint perspective, or the clunky internal viewpoint are a barrier to communication.

In contrast 'immersive' collaborative virtual environments place the user physically into the virtual scene. These systems are less scalable than online virtual communities which can take better advantage of distributed resources [**Grimstead2005a**, **Benford1998**]. Goebbel et al. designed a taxonomy for what they termed simply CVE's. In their design they provide an excellent high level description of the ICVE as spaces which provide "distributed collaborative teams with a virtual space where they could meet as if face-to-face, co-exist and collaborate while sharing and manipulating virtual data", crucially designed in a way to 'disburden' the users' senses [228] and reduce fragmentation of the shared perceptual environment [229]. The ability to collaborate in such systems extends even to closely coupled physical tasks [230].

There are various technology systems which demonstrate heightened immersion and presence in a virtual environment while allowing interaction between parties who may or may not be physically co-located [**Murray2009a**]. Attempts at bringing people together in VR extend back to the early days of the technology and systems such as DIVE [**Benford:1995hh**].

Bailenson et al. noted that while increased realism of such avatars increases co-presence it decreases self disclosure [231]. Such systems seem to compromise social interaction even as the realism increases. Avatar representations increased in fidelity and eventually it became possible to share avatar representations of participants wearing body tracking [232] and eye tracking equipment [233, 234]. This enabled tracked, viewpoint independent interaction with reconnection of eye gaze cues within VR in a spatially faithful way [**Steptoe2010b**, 235, 236, 237, 98].

It is also possible (at least in research implementations) to fully reconstruct people as 3D models, and connect more than two users together utilising life-size multi-view video [140, 214, 238, 223, 239]. These 3D video representations of users can share a virtual space in which spatial and directional information are maintained in a natural way [140, 58].

Fairchild et al. have developed a system for ICVE and desktop based on earlier work by Duckworth et al. [**Duckworth:2012tx**], adapted for use in the Telethrone system [**Fairchild2016**]. **Eye tracking in ICVE** More recently, advanced eye tracking technologies allow the recording and replaying of accurate eye gaze information [**Ciger2004a**, 98, 237] alongside information about pupil dilation toward determination of honesty and social presence [**Steptoe2010a**]. Heldal found that collaborative tasks manifested fewer disturbances due to "misunderstandings of reference or action" when using more immersive systems [240].

7.4.10.6 Situated Displays

Between the complexity of ICVEs and the more ubiquitous screen based VC technology there now exist so-called situated displays.

Conversation does not exist in isolation, but rather in the context of semiotic resources. Participants make constant reference to the surrounding assets through mutual orientation, gesture, and diverse cues [**Goodwin2000a**]. So called situated displays seek to embed the represented participant within the spatial and contextual framework of the conversation such that the referential cues are better supported. This has many implications, but chief amongst these is support for a spatially faithful conversational environment supportive of gaze [**Pan2014a**, **Pan2012a**, **Zhang2013a**, 192].

Such displays place a representation of the remote user into a space, theoretically allowing all participants to physically interact with the ‘contextual configurations’ around them [Goodwin2000a]. This is a relatively new field of research. These could include the aforementioned telepresence robots [Lee2011b, 190, 184], head in a jar implementations such as SphereAvatar [pan2014comparing, Pan2012a, 192], and Gaze Preserving Situated Multi-View Telepresence System [Pan2014a]. Sphereavatar [192] demonstrates that there are problems with accurate mapping, distortion, and projection, and movement of the captured actors outside of very tight bounds.

Telehuman brings the whole body of a standing remote user into a space via a cylindrical display with a single tracked observer viewpoint [241].

7.4.10.7 Retro-reflective Projective Technology (RPT)

Retro-reflective materials such as Chromatte(tm) cloth reflect light back along the angle of incidence. An everyday application of such material is high visibility jackets.

Inami et al. described the first use of RPT in 2000 with their visuo-haptic display [inami2000visuo]. This system used head mounted projectors to augment RPT shapes in the real world. Krum et al. describe the REFLCT system [242] which uses large retro-reflective surfaces to “provide[s] users with a personal, perspective correct view of virtual elements that can be used to present social interactions with virtual humans”. They use helmet mounted projectors and describe a military training application in a large volume which allows faithful transmission of attention and gestures from the virtual to the real. They also briefly describe augmenting a facial mannequin by projecting onto RPT adhered to the surface. They point out that the optical characteristics of the material maintain polarisation and so could support passive stereoscopy.

Tachi describes an augmented reality system where a helmet mounted projector places an image of a remote human onto a robot. This is a system they describe as Teleexistence RPT [243, 244]. This system demonstrates a single user, wearing a head mounted projector, viewing only a head which is captured from a single viewpoint. It is however enabled for motion by means of robotics under the Chromatte cloth like the hybrid robotic systems described earlier.

Hua et al. demonstrated elements of ‘SCAPE’, a tracked head mounted projection system surrounded by RPT cloth which could surround multiple physical users in a shared immersive experience [245]. Hypothetically this system could be extended to include what they term ‘passive remote users’ projected into the views of the co-located and immersed users. The set-up however is quite complex, and involves wearing headgear, so is again less suitable for informal meetings.

Surman et al. discuss a potential development of their retroreflection based auto-stereoscopic display (which head tracks a single user to present a depth enabled stereo pair). They suggest that multiple users could address a large screen, but think that their stereopsis would break down and the system would be extremely challenging [JSID:JSID295].

Although less pertinent to this research there are other interesting deployments using RPT. Darken et al. and later Hahn in his PhD thesis in the same group describe a novel system which mounts a chromakey LED light ring and camera on a HMD [246]. The camera takes in a mocked up physical cockpit and windows which are made of Chromatte cloth. This video image is very easy to segment to replace the green-screen windows with a simulated external view. The pilot trainee sees their own hands and the cockpit instrument panels (albeit in monoscopic video), alongside the generated external view.

This augmented reality is much cheaper to deploy than a full simulator set-up.

7.4.10.8 Furniture as a Mediating Display

Paul Sermon experimented with projection onto beds in his telematic dreaming work [247].

Room2room from Microsoft Labs [147] demonstrates the utility of projecting onto furniture by building upon their previous automatic projection mapping research [248] using their Kinect system and projectors. This system is seen in Figure ???. Their system provides spatially correct viewing through reconstruction but only on a point to point basis with a single user at each end.

7.4.10.9 Augmented Reality Collaboration

Lehment et al described what they term a “consensus reality” for head mounted AR in which they compute the correct locations for remote participants as 3D video representations overlaid on the real [249].

7.4.11 Network Issues

Bradner & Mark established that latency matters in remote telecollaboration [Bradner:2002:WDM:587078.587110] while an extensive body of work notes that system latency impacts key factors such as interruptability [250]. Gergle et al. found that remote collaborating pairs of people were tolerant of delays in visual feedback up to 939ms, after which shared task performance suffered considerably [251]. Overall it is clear that for communication systems, especially those with potential for shared or collaborative task, those aspects of the system latency which can be controlled and/or minimised should be.

Early research around COVEN and DIVE described the benefits of multicast for multiparty collaborative technology systems. While multicast still has advantages in latency critical set-ups the bandwidth/throughput has to a degree ‘caught up’, while the technical demands of a wide area multicast network remain the same. The DIVE system has a standalone WAN application layer network transport called DIVEBONE which attempted to address this issue [Greenhalgh2001a]. Steed et al. describe a hybrid multi/uni cast system which allows local multicast with unicast on the WAN [252].

Bradner used social impact theory to suggest it is important to either say the distally located person is geographically remote to emulate real systems, or else explain they are in the next room to exclude this factor [Bradner:2002:WDM:587078.587110], while an extensive body of work notes that system latency impacts key factors such as interruptibility [250].

Gasparello et al. [Gasparello:2011in] then later Moore et al. [Moore:2010jt] worked within the wider research group on novel network transport systems for synchronised 2D and 3D video information to support telepresence reconstruction in the later withyou system [140].

When Lamboray extended their blue-c 3D video system they considered that latency of up to 200ms was acceptable so long as jitter (changes in latency) remained low [Park1999, lamboray2005data].

7.5 Support for Less Formal?

Outside of the sphere of simple webcam systems such as Zoom and their ilk there has been little attention invested in advanced telecommunication tools for informal settings. This is

exactly the kind of thing necessary to enable home working in the future.

Slovak et al. describe a group-to-group videoconferencing system called GColl in which they highlight mutual gaze support and modest technical requirements [253]. GColl is a monitor based system which seeks to align mutual gaze by rendering to a small window within 5 degrees of the mounted camera. It seems to address the right problems but is a questionable development over and above a careful videoconferencing set-up. De Greef and Ijsselsteijn describe a system which provides collaborative working tools alongside video and audio conferencing for the home [DeGreef2001]. Judge and Neustaedter examined how 21 families used videoconferencing in a home setting and found that requirement for planning and availability for using the tools inhibited their use. They suggest an 'always on' system might better mediate availability [254].

Interestingly Lottridge et al. identified that it is the empty reflective moments during mundane activity which might benefit most from the ability to strike up intimate conversations between separated couples [255]. Whether and how much this supposition is analogous to the so-called 'water cooler meeting' moments in business remains an open question but is perhaps worth bearing in mind.

There is, of course, Room2Room which provides a far better informal home experience, but is inherently one to one [147]. The same is true of 'Holoportation', a reconstructed and tracked 3D video based system designed by Microsoft for head mounted display and seemingly aimed at the home market [orts2016holoportation]. Similarly utilising Microsoft hardware, but on a tight budget, HomeProxy uses a simple combination of consumer devices to render a reconstructed remote partner to a cabinet style display on a desk [256]. Users reported that they appreciated the look and feel of the HomeProxy system which is designed to fit in sympathetically to home furnishings.

Gaver explains that "the "space" created by audio-video technologies is significantly different from spaces as found in hallways, offices or meeting rooms", conveying a limited subset of visual and auditory information, preventing movement and exploration, and is often arbitrary and discontinuous. He says that "these properties shape the possibilities for collaboration." [257].

In a work setting Fish et al. said (in 1993) that "informal communication cuts across organisational boundaries and often happens spontaneously." They reference research which suggests that informal communication is more frequent, expressive, and interactive [258]. Fish went on to look at early telepresence systems which might support informal, but by any standard these systems are too clunky to be deemed capable of such. Although this has been recognised as an issue for decades it could nonetheless seem that less formal business meetings are of insufficient utility to attract the technology and research to better support them. Computer Supported Collaborative Working (CSCW) is a well known acronym but there seems to be no single point of entry to research into less formal systems. This could be a function of the technology being better suited to deployment in a controlled formal setting.

The paper "Ad hoc versus established groups in an electronic meeting system environment" turns out to be using the other definition of ad-hoc meetings which is meetings convened from random members with no past and no future together. The study is also dated being from 1990, but still provides insight. In the paper Dennis et al. reviewed what little research existed at the time noted that meeting outcomes varied between studies as to whether ad-hoc or more established and formal meetings were more productive [259].

Fayard and Weeks examined the affordances of informal interaction (indeed exactly

the ‘water cooler meetings’ that the BBC were talking about) [260]. They identified: proximity, privacy, and legitimacy as key affordances which should be supported, and highlighted: functional centrality, semi-enclosure, reciprocal visibility, easy access and egress, multiple shared resources as important further affordances of the environment. The affordance of legitimacy is interesting. This means that the location of the informal meeting should support the feeling that it is legitimate to be there and strike up a conversation. A potential model example is explored in Horgan’s book ‘Excellence by design: Transforming workplace and work practice’. ‘The LX common’ was a semi communal space for informal meetings. It was semi-enclosed to give some privacy, but located on a busy thoroughfare and housed the photocopiers, printers, shared reference material, kitchen etc. People who passed felt free to listen and occasionally join in. Three rules evolved from the research: Traffic through the common was acceptable at any time, anyone was free to join any meeting in the common, anyone was free to leave any meeting in the common at any time [horgen1999excellence].

Supporting spatial aspects such as mutual gaze or ‘full gaze awareness’ [84] in this way normally demands large purpose built installations which poorly support casual or ad hoc meeting paradigms [261, 58]. It seems clear that there is a justification for development of simple affordable technologies which better mediates communication over distance for less formal situations.

7.5.1 Support for Dynamic Meetings

Voida et al. discuss natural sub grouping in partially distributed teams [132]. These ‘fault lines’ within meetings are a natural component of more informal meeting structure with groups potentially forming and dispersing naturally over time.

Greenburg and Jerald found in group interaction with trainers that members often changed their seating position to facilitate discussion goals [greenberg1976role]. This may not be true of all group meetings but it is under supported in telepresence systems.

This dynamic reorganisation of meetings is potentially an affordance for ad-hoc and casual meeting styles. It demands the ability to reorganise chairs such that they still have faithful views onto remote participants. There is no support for this found in literature.

7.6 Summary

Attention is mediated in a large part by non-verbal cues. A large body of research across psychology tells us that Extra Verbal Content is important [Goodwin2000a] for mediating attention [Argyle, Hofmann20061683, Pan2008a, 96, 126, 97, 127, 128, 129, 121, 130], conversational flow, and transmission of mood and manner.

Support for these cues is limited in conventional commercial systems which convey the image of a person over a distance through technology.

Eye gaze, mutual eye gaze, and facial expression [Argyle, ohba1998facial, 125] are particularly rich information channels, as are body torque, gesture, and head gaze. Proxemics shows us that interpersonal distance is an important feature of communication and can evolve as situations change [118].

It is especially difficult to support mutual eye gaze well because of the Mona Lisa effect where a viewer of a screen showing a camera view of a remote participant will only perceive that they are being looked at when the remote user is looking directly into the camera lens which captures them. Fixing this disconnect of mutual gaze in telepresence

systems is an active area in both research and commercial system.

Telepresence solutions aimed at collaborative working have traditionally attempted to join remote spaces so that people in each can look into the other, seeing each other, at best, as if through a window. Immersive Collaborative Virtual Environments (ICVE) have been used to join remote spaces so that they coincide. Free-viewpoint video has been combined with immersive projection technology to allow people to move around spaces, seemingly together [140]. For these spaces to be aligned so that they coincide, removing the impression of being separated by glass, users must wear stereo glasses.

More recently, situated displays have attempted to give the impression of placing people within each other's room, but either lack the spatial faithfulness available within ICVE, or else do not render the whole of the remote collaborator.

Situated displays can place a representation of the remote user within a space. Most informative to the research at the time of the literature review was the work of Pan et al. [pan2014comparing, Pan2012a, Pan2014a, 192]. Paul Sermon explored situating live telepresence video onto furniture [247].

The bulk of commodity systems for telecommunication remain dyadic, that is, person to person, though this is changing fast in the face of the 2020 global pandemic. More interesting and difficult is small group interaction, so-called 'teleconferencing'. This is a difficult field, with known issues in all of the market leading systems. Mona Lisa effects, latency, jitter, and the requirement of a lot of expensive equipment to be maintained alongside a booking system for its use amongst the most problematic.

Inami pioneered the use of RPT [inami2000visuo] and this was built on by others, most notably Tachi [243, 244]. Tachi's projection of a face onto a robotically manipulated surface has some similarity to Krum et al's REFLCT system [242] in that observers move around to gain viewpoint independent and spatially segmented views on a retro-reflective surface, but the application is very different. Retro reflection (light field) is used in the research leading Telehuman 2 [gotsch2018telehuman2], and the small group Telethrone system [o2018telethrone].

The Holovizio real-time 3D light field transmission system[3dlightfield] is seen in Table ??.

Maimone et al. presented their "Encumbrance free telepresence" [6162881], which was not captured in the initial review but is now added to Table ???. It can be seen that this system provides many of the desired affordances but falls short on being able to support small groups. During the peer-review process the "SCAPE" system [1255811] was highlighted. This comparatively complex system combines immersive projection from head mounted projectors (onto surrounding walls and objects) with tracking and virtual avatars to create multiple viewpoints into a shared virtuality.

7.6.1 Comparing Key Systems and their Affordances

Affordances are properties of the environment that offer actions to appropriate organisms [gibson2014ecological]. In this research the Gibson definition of 'an action possibility' is adopted rather than the more restrictive Norman definition of 'a perceived suggestion' [norman2002psychopathology]. McGrenere and Ho sum up the difference in definitions of affordances by considering a hidden door in a paneled room [262]. In Gibson's definition the existence of the door (or parallax cues, or stereoscopy, or eye gaze for instance in this research) does not need to be explicitly known to the user, just its existence is enough. In Norman's view the door would need to have a handle to suggest the action to the user. This

is a less relevant HCI definition for this research.

7.7 Theoretical Framework

7.7.1 Problem Statement

It is clear that there are multiple factors which contribute to successful human-human communication. These factors remain important in telecommunication supported by technology, and are variously supported, unsupported, or modified by particular technologies.

Of particular importance is interpersonal gaze [Kleinke1986a, 100, 101], and gaze is an excellent dependant variable for experimentation. Non-verbal cues are also important across multiple modalities of sight, sound [75], and position of interlocutors [83], extending to the whole body [Kleinke1986a, 108].

The Mona Lisa effect demands that some novel system which allows spatial signals to be maintained must be employed [65].

While formal meeting paradigms are supported to an extent by commercially deployed systems this does not suit all meeting styles. Such systems are generally booked well in advance and so meetings tend toward a formal structure. These meetings seem to demand many smaller supporting meetings between parties or groups of parties.

The ‘problem’ is a supporting technology for small less formal groups. One which connects home and work spaces without bringing in those backgrounds, creating a level playing field. A fully pervasive system could also allow dynamism and movement, connection of natural non vocal cues, without too much encumbering technology overhead.

7.7.2 Core Assumptions

Figure 7.3 shows the interlocking relationships between baseline communication where the participants are present, and technology which attempts to support across distance.

Of most interest to this research is the centre of the Venn where meeting styles which are less formal, and perhaps dynamic, may occur. Looking at these items one by one gives us our core assumptions.

1. Gaze

Gaze is broadly agreed to be highly important for mediating flow. Mutual gaze is a rich emotional channel. The research must consider gaze. All of the researchers listed around the Venn have at some point engaged with this topic.

2. Attention

The non-verbal communication channel employed in ‘attention’ is assumed based upon the literature to be critical to smoothly leaving and entering a fast flowing conversation where concentration around a defined problem may be high (gesturing to a chair for instance). Again, all of the listed researchers have made reference to attention in their work.

3. Spatial (immersive)

Support for spatiality is important in a group setting so that directional non-verbal cues can find their target. The topic of spatial relationships between interlocutors cuts across all of the researchers, but this is not true of immersion. Immersion in a shared virtuality can certainly support the underlying requirements spatial, but the technical infrastructure required is out of scope (so this is struck through on the diagram). Roberts and Steed are the main expertise referenced even though this

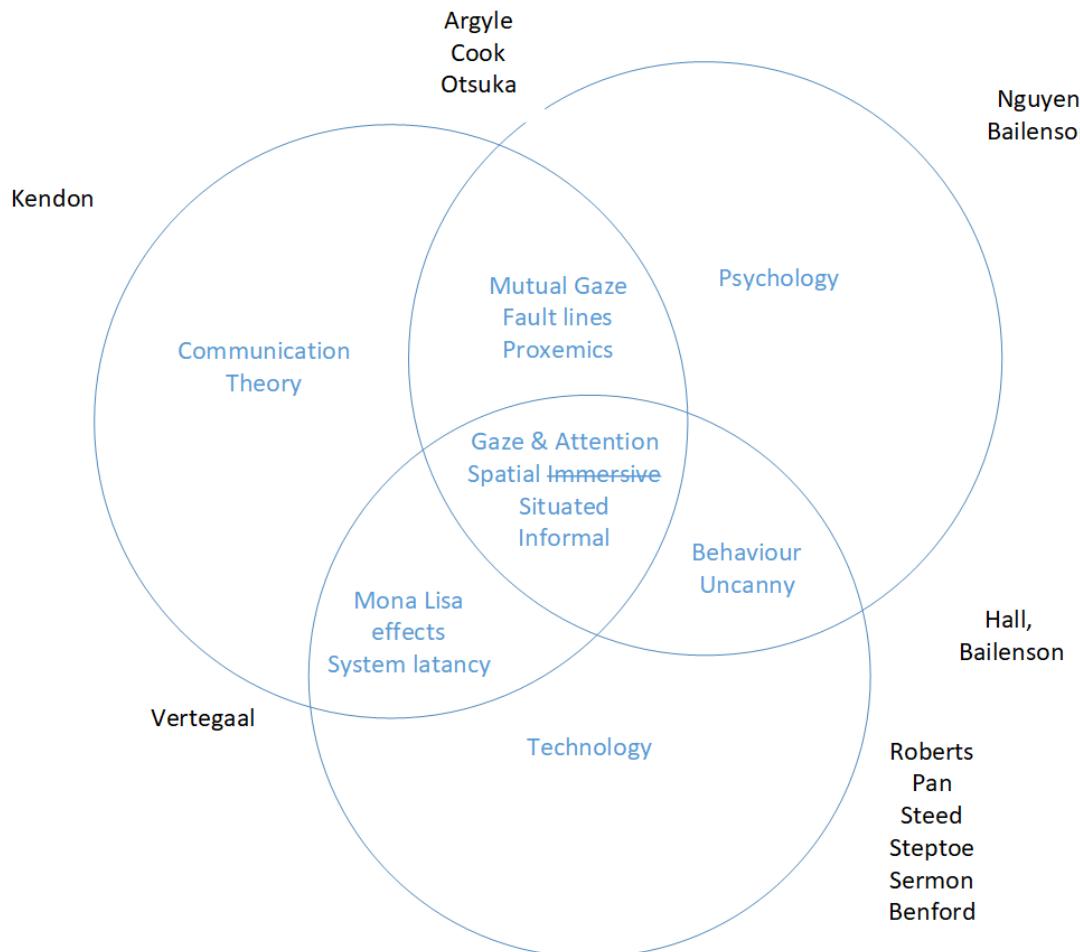


Figure 7.3: The Venn diagram shows areas of research which have been identified in blue. These interlock and overlap as shown. The most relevant identified researchers from the literature are shown in black close to the fields of study which they represent. This diagram is a view of the core assumptions for the research, with the most important fields at the centre.

element is not expanded in the research.

4. Situated

Situated displays are those which are appropriate for their surrounding context, in this case the informal meeting. Roberts, Pan, Steed and Steptoe seem the most relevant researchers in these technology spaces.

5. Informal

Based on the literature proxemics is believed to be relevant in a meeting where subgroups can be instantiated and destroyed as the meeting evolves, and those where people can be invited in from outside the physical bounds of the meeting (informal spaces). Hall is the best source for this work. If it is assumed that people may come and go, and subgroups may be convened then Sermon and Benford are the best references through their work blending real and virtual spaces. This may be more consistent with less organised meetings such as those convened on demand (ad-hoc).

7.7.3 Peripheral Assumptions

Surrounding the centre of the Venn are additional relevant topics from social science branches of theory

From verbal communication

It is assumed that the directionality of sound is important [141], and this will be engineered into the experimental design. It is assumed that movement of the lips is an indicator and this is tied to latency and frame rate in the vision system.

From non-verbal communication

It is assumed that eye gaze is of high importance, and that this information channel is supported by head gaze and body torque to a high degree. It is further assumed that mutual eye gaze is of less relevance in a multi party meeting where there is a common focus for attention but can be significant for turn passing. It is assumed that upper body framing and support for transmission of micro and macro gesturing is important for signaling attention in the broader group, and for message passing in subgroups.

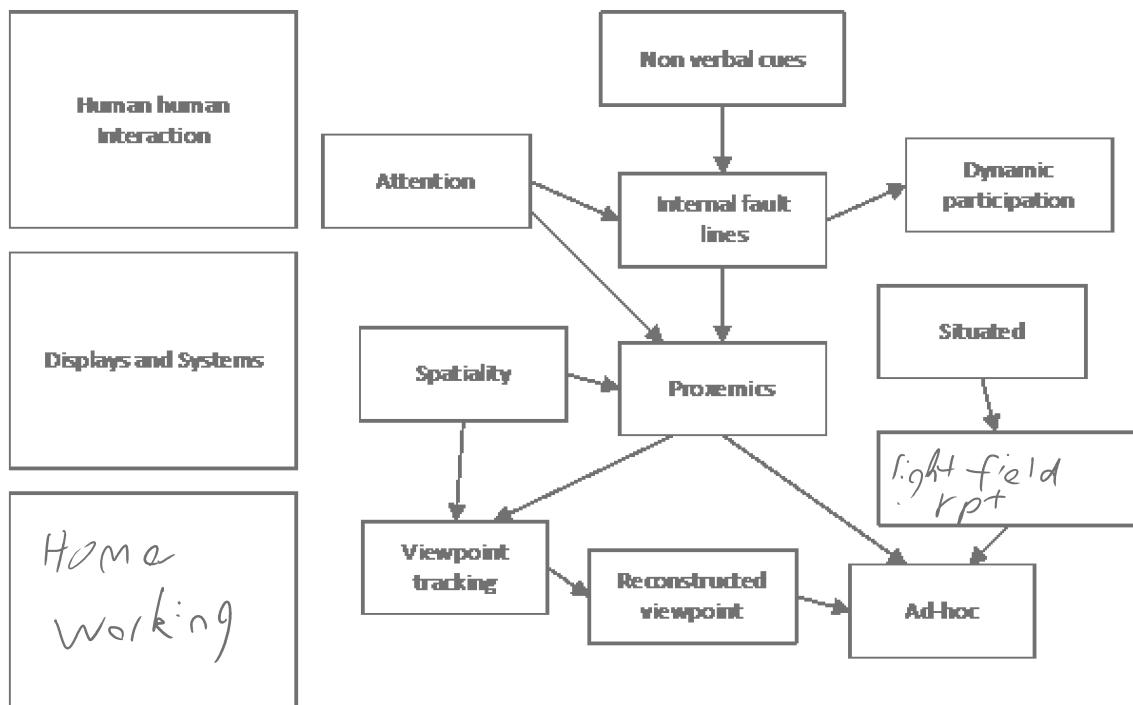


Figure 7.4: Connected concepts revealed during the review, and roughly how they relate to one another. The web moves top left to bottom right as the attention focuses. The left hand column shows the broad flow from interpersonal communication in the real (top) to home working (bottom).

7.8 Post 'Meta' metaverse

The current media around “metaverse” has been seeded by Mark Zuckerberg’s rebranding of his Facebook company to ‘Meta’, and his planned investment in the technology. The second order hype is likely a speculative play by major companies on the future of the internet. There has been a reactive pushback against this by the wider tech community who are concerns about monetisation of biometrics. [Observers do not trust](#) these ‘Web2’

players with such a potentially powerful social medium. It is very plausible that this is all just a marketing play that goes nowhere and fizzles out. It is by no means clear that people want to spend time socialising globally in virtual and mixed reality. These major companies are making an asymmetric bet that if there is a move into virtual worlds, then they need to be stakeholders in the gatekeeping capabilities of those worlds.

Meta, Disney plus, Sportswear manufacturers

[Can enough be done to prevent abuse?](#)

It seems like there are four major interpretations of the word.

Facebook have recently rebranding their parent company as ‘Meta’ and they are aggressively promoting “The Metaverse” as a shared social VR space, chiefly of their design. In Stephenson’s ‘Snow Crash’ the Hero Protagonist (drolly called Hiro Protagonist) spends much of the novel in a dystopian virtual environment called the metaverse. It is unclear if Facebook is deliberately embracing the irony of aping such a dystopian image, but certainly their known predisposition for corporate surveillance, alongside their attempt at a global digital money is ringing alarm bells.

The Grayscale investment trust [published a report](#) which views Metaverse as a potential trillion dollar global industry. Such industry reports are given to hyperbole, but it seems the technology is becoming the focus of technology investment narratives.

7.8.1 Mixed reality as a metaverse

[Spatial anchors](#) allow digital objects to be overlaid persistently in the real world. With a global ‘shared truth’ of such objects a different kind of metaverse can arise. One such example is the forthcoming [AVVYLAND](#).

Peloton as a metaverse?

7.9 Digital Land Metaverses

One of the most intuitive ways to view a metaverse is as a virtual landscape. This is how metaverse was portrayed in the original Neal Stephenson use of the word.

7.9.1 Legacy Web2

7.9.1.1 [Secondlife](#)

7.9.1.2 [Roblox](#)

7.9.1.3 [Minecraft](#)

7.9.1.4 [Fortnite](#)

7.9.2 The new stuff

7.9.2.1 [Decentraland](#)

JPMorgan Chase, the investment bank, has just opened a ‘lounge’ in the Decentraland metaverse. This coincided with a [major report](#) on the potential opportunities.

7.9.2.2 Sonmumspace**7.9.2.3 The Sandbox**

7.10 Global enterprise perspective

Meta(Facebook)

Nvidia Omniverse is [free for creators](#), Unity etc

Microsoft have just bought Activision / Blizzard for around seventy billion dollars. This has been communicated by Microsoft executives as a "Metaverse play", leveraging their internal game item markets, and their massive multiplayer game worlds to build toward a closed metaverse experience like the one Meta is planning. This builds on the success of early experiments like the Fortnite based music concerts, which attracted millions of concurrent users to live events.

7.11 NFT as metaverse narrative

Within the NFT community it is normalised to refer to ownership of digital tokens as participation in a metaverse. This CNBC article highlights the confusion, as this major news outlet refers to [Walmart prepares to offer NFTs](#) as an entry "into the metaverse".

7.12 MMORG games and NFTs

Traditional gamers have pushed back on the seemingly useful idea of integrating NFTs with traditional games. This may be in part because Ethereum mining has kept graphics card prices high for a decade.

HBAR partnerships The following text is from Justin Kan, co-founder of twitch:

NFTs are a better business model for games. Many gamers seem to be raging hard against game studios selling NFTs. But NFTs are also better for players. Here's why I think blockchain games will be the predominant business model in gaming in ten years. NFTs are a better business model for funding games . Example: recently I invested in a new web3 game SynCityHQ. They are building a mafia metaverse and raised \$3M in their initial NFT drop.

NFTs give studios access to a new capital market for raising capital from the crowd.NFTs can be a better ongoing model for games. Web3 games will open economies, and by building the games on open and programmable assets (tokens + NFTs) they will create far more economic value than they could from any one game. Imagine Fortnite, but other developers can build experiences on top of the V-Bucks and skins. Epic would get a royalty every time any transaction happens. As big as Fortnite is today, Open Fortnite could be much bigger, because it will be a true platform. NFTs are better for gamers. Allowing gamers to have ownership of the assets they buy and earn in game allows them to participate in the potential growth of a game. It lets gamers preserve some economic value when they switch to playing something new. But what about the criticisms of NFTs? Here are my thoughts on the common FUDs: "It's just a money grab on the part of the studios!" Game studios already switched over to the model of selling in-game items, cosmetics, etc to players long ago. But currently the digital stuff players are buying isn't re-sellable. NFT ownership is strictly better for players. "The games aren't real games." This reminds me of the criticism of free-to-play in 2008, when the games were Mafia Wars / FarmVille. We haven't had time for great developers to create incredible experiences yet. Everyone investing in games knows there are great teams building. "Game NFTs aren't really decentralized because they rely on models / assets inside centralized game clients." Crypto is as much a movement as it is a technology. Putting items on a blockchain is what gives people trust that they have participatory ownership...which make people willing to buy in to the game. These assets are "backed" by blockchain. The fact that these item collections are NFTs will make other people willing to build on top of them. "NFTs are bad for the environment." Solana and L2s solve this. NFT games are better for players and for game developers. Like the free-to-play revolution changed gaming, so will blockchain. The games of the future will be fully robust, with open and programmable economies.

- As a user I want to select a digital asset I find in the AR/VR world and then be offered an option to purchase the asset so I can look at it in my own spaces.
- As a user I want to click on a digital asset I find in the AR app and be given the opportunity to buy it as a rare digital representation so that only I and a few others are provably certified to own.
- As a user I would like to transfer economic value to people and entities I meet in the metaverse such that it is agreed by all parties quickly that value has been provably transferred.
- As a user I would like to access an online marketplace in the metaverse where I swap and trade digital assets with other users so that I continue to feel engaged.
- As a user I would like to create content (inside or outside of the metaverse) so I can take it to metaverse and monetise it.
- As a content creator or influencer I would like to engage with live audiences within the metaverse, and monetise my opinions and knowledge in real-time. I would like

to have a way to split this money with co-collaborators in real time.

7.13 Crypto metaverses

[Report on Axie Infinity](#)

[Pavia Metaverse](#) Probably the best example in the market at this time with connecting users with one another through blockchain is [Satoshi's Games 'Litenite'](#). Litenite is a 'battle royale' game which allows users to earn Satoshi's through the Lightning network.

Similarly, and potentially more significantly, Zebedee have brought Lightning based micropayments to [Counter Strike](#), which adds a financial layer directly to eSport, itself a multi billion pound global industry.

The Zebedee model is interesting in that they provide simple onboarding, and management solutions, for gaming and metaverse application developers. There is doubtless an opportunity to utilise their business model in the proposed stacks in the paper, but it would be at odds with the free and open source product methodology in this paper. Their CTO said the following in a [recent podcast](#):

"We all had very similar ideas around being able to put Lightning network capabilities into games. You're essentially putting real value inside of the game. So whether it's a point inside of the game or as a real end game economy, as a game developer, you don't have to worry about the mechanics around it."

You can use a real life currency: the money that exists in the world and that value carries regardless of which game you're in, regardless if you're in the real world or inside of a virtual world, like a game. We just think that vision is just sort of opening. Now there's so much more, if you extrapolate it into the Meatverse we would love for Zebedee to be a big platform provider and enabler for a lot of these."

[Spells of Genesis](#) is a long running card RPG trading game on mobiles which allows "ownership" of items and cards through non-fungible token ecosystems.

There are also hundreds of casinos which operate within and even on blockchain networks. These feel out of scope as they are a different and somewhat regulated offering.

7.13.1 Immersive and third person XR

In considering the needs of business to business and business to client social VR is it useful to compare software platforms:

7.13.1.1 Second Life

Notable because it's the original and has a decently mature marketplace.

7.13.1.2 Spatial

- Very compelling. Wins at wow.
- Great avatars, user generated
- AR first design
- Limited scenes
- Smaller groups (12?)
- Limited headset support
- Intuitive meeting support tools
- No back end integration

7.13.1.3 MeetinVR

- Good enough graphics, pretty mature system
- OK indicative avatars, user selected
- VR first design
- Limited scenes
- Smaller groups (12?)
- Quest and PC
- Writing and gestures supported
- Some basic enterprise tools integration
- Bring in 3D objects
- Need to apply for a license?

7.13.1.4 Glue

- Better enterprise security integration
- Larger environments, potential for breakouts in the same space. Workshop capable
- 3D object support, screen sharing, some collaborative tools
- Apply for a license
- Fairly basic graphics
- Basic avatars
- Quest and PC
- Writing and gestures supported
- Mac support

7.13.1.5 Mozilla Hubs

- Open source, bigger scale, more complex
- Choose avatars, or import your own
- Environments are provided, or can be designed
- Useful for larger conferences with hundreds or thousands of members but is commensurately more complex
- Quest and PC
- Larger scenes within scenes

7.13.1.6 FramesVR

- Really simple to join
- Basic avatars
- Bit buggy
- 3D object support, screen sharing, some collaborative tools
- Quest and PC
- Larger scenes within scenes
- Runs in the browser

7.13.1.7 AltSpace

- Microsoft social meeting platform
- Very good custom avatar design
- Great world building editor in the engine
- Doesn't really support business integration so it's a bit out of scope
- Huge numbers (many thousands) possible so it's great for global events
- Mac support

7.13.1.8 Engage

- Great polished graphics
- Fully customisable avatars
- Limited scenes
- Presentation to groups for education and learning
- PC first, quest is side loadable but that's a technical issue
- BigScreen VR
- Seated in observation points in a defined shared theatre
- Screen sharing virtual communal screen watching, aimed at gamers, film watching
- up to 12 user

7.13.1.9 VRChat

7.13.1.10 NEOSVR

Notable because it's trying to integrate crypto marketplaces

7.13.1.11 Meta Horizon Worlds & Workrooms

Horizon Worlds is the Meta (Facebook) metaverse, and Workrooms its business offering and a subset of the “Worlds” global system. It is currently a walled garden without connection to the outside digital world, and arguably not therefore a metaverse.

The Financial Times took a look at their patent applications and noted that the travel is toward increased user behaviour tracking, and targeted advertising.

Facebook actually have a poor history on innovation and diversification of their business model. This model has previously been tracking users to target ads on their platform, while increasing and maintaining attention using machine learning algorithms.

It makes complete sense then to analyse the move by Meta into 3D social spaces as an attempt to front run the technology using their huge investment capacity. Facebook have recently taken a huge hit to their share price. Nothing seems to change in the underlying business except Zuckerbergs well publicised shift to supporting a money losing gamble on the Metaverse. It is by no means clear that users want this, that Meta will be able to better target ads on this new platform, or that the markets are willing to trust Zuckerberg on this proactive move.

With all this said the investment and management capacity and capability at Meta cannot be dismissed. It is very likely that Meta will be able to rapidly deploy a 3D social space, and that its development will continue to be strong for years. The main interface for Horizon Worlds is through the Meta owned and developer Oculus headset, which is excellent and reasonably affordable. It has been quite poorly received [by reviewers](#) but will likely improve, especially if users are encouraged to innovate.

7.13.1.12 Vircadia

We have chosen Vircadia as our development platform for this investigation as it's a community supported free and open source project with some support for economic interaction.

7.13.2 WebXR/WebGL/WebGPU

7.13.2.1 Playcanvas

Monetisation of WebXR Playcanvas - Telefonica

7.13.3 Integration with web and game engines

(other integrations?)

7.14 User stories

7.15 Recommendations for value transfer

There are many claims about what blockchain technologies can enable. In the tens of thousands of attempts at utility over the last decade there are surprisingly few chains able to claim any value at all, and those that can often have significant problems in other areas. It is not the intention of this analysis to poke at problems. The primary use case within the context of a shared social space (metaverse) is low friction transmission of value. Remember that interlocutors and entities might have globally different physical locations outside of the real world. What is needed is instant “settlement” of digital bearer instruments within the context of the digital environment. This is fortunately what Bitcoin was designed to do.

7.15.1 Streaming micropayments

Lightning enables automatic streaming of micropayments while avatars are in certain areas or engaged in chargeable activities.

7.16 Bitcoin market gap

Nobody is currently deploying the discussed technologies purely on Bitcoin because it's slower evolution is still catching up. This is a market gap and the following section shows how this might be done.

7.16.1 Money

7.16.2 Identity proofs

Microsoft ION? NOSTR?

7.16.3 Digital object tracking

Connect

7.16.4 Object transfer and trading

Additionally it is highly possible that exchange of digital goods (art and objects and provable services) will be required. This is also in scope within this document.

7.17 Risks

[Online regulations still apply](#)



8. Hardware and software choices to support this

We propose that the software be deployable into a server based virtual network running on a single system. Our hardware recommendation to run a balanced selection of elements is a quad core 16GB system with two 1TB SSD drives. The system diagram can be seen in Figure 8.1.

8.1 Networking layer

8.1.1 Proxmox

Proxmox open source server virtualisation allows the deployment of several virtual machines within a hardware system. Each of these VMs can have a different balance of performance, ease of use, and security. The VMs are given only the access they need to perform the task which they are specialised for. This improves overall security. Resilience is improved since elements of the cluster can be shutdown, reinstanced, and upgraded, without impacting the whole. Snapshots and backups are simplified.

8.1.2 VyOS firewall

VyOS firewall is a Proxmox aware threat management gateway and firewall which allows more nuanced interfacing with corporate systems, while maximising the security of the VM network which sits behind it in the virtual cluster.

8.1.3 Privacy aspects and using TOR etc

Currently some compromises with privacy may be necessary. Power users vs standard users. KYC/AML. [tor nodes](#)

8.2 Bitcoin value management stack

8.2.1 Bitcoin Core on NixOS

8.2.2 Electrum server

Options on the Blockstream website

8.2.3 C-Lightning and CLBoss

Both Donner Labs and Zebedee have code packages which allow interaction with the lightning network within Unity. In all likelihood users would have to run a lightning / Bitcoin node and have their users interact with it. This would allow instantaneous transactions of Satoshis (the Bitcoin unit of account) between users.

8.2.4 LDK based Sensei

[Sensei](#) is a new node implementation which uses a hub and spoke model to allow light nodes to operate lightning channels without running the whole Bitcoin stack themselves. This is potentially useful for the client side of the deployment.

8.2.5 LNBits with RGB and metaverse plugin

8.2.6 Backup & Watchtower

8.2.7 Object and media tracking: RGB

The world database in the shared rooms in the metaverse is the global object master, with RGB client side validation inheriting and validating against these objects.

8.3 Identity

8.3.1 nostr

User authentication and sideband communication will be through nostr. This allows completely private end to end encrypted chat sessions between participants.

8.4 Metaverse

8.4.1 Vircadia

8.4.2 Wolvic

“The goal of the Wolvic project is to create a full-featured browser exclusively for standalone AR and VR headsets.”

Wolvic is a continuation of the now defunct Firefox Reality Browser project. It is open source.

8.4.3 WebM

[Free open source web video player](#)

8.5 Messengers and boards

8.5.1 Nostr

8.5.2 Cyperpost

[Cyperpost](#) is a social platform that uses ecdsa keys for e2ee. It’s private first so no plain text on servers so no scope for censorship so it’s okay to be centralized. All accounts are made with a single ID pubkey derived from a social Mnemonic, so you maintain uniform

ID on all platforms.

8.5.3 Matrix

Matrix is the main self hosted solution for chat and also uses ecdsa keys.

8.6 Addressing identified risks

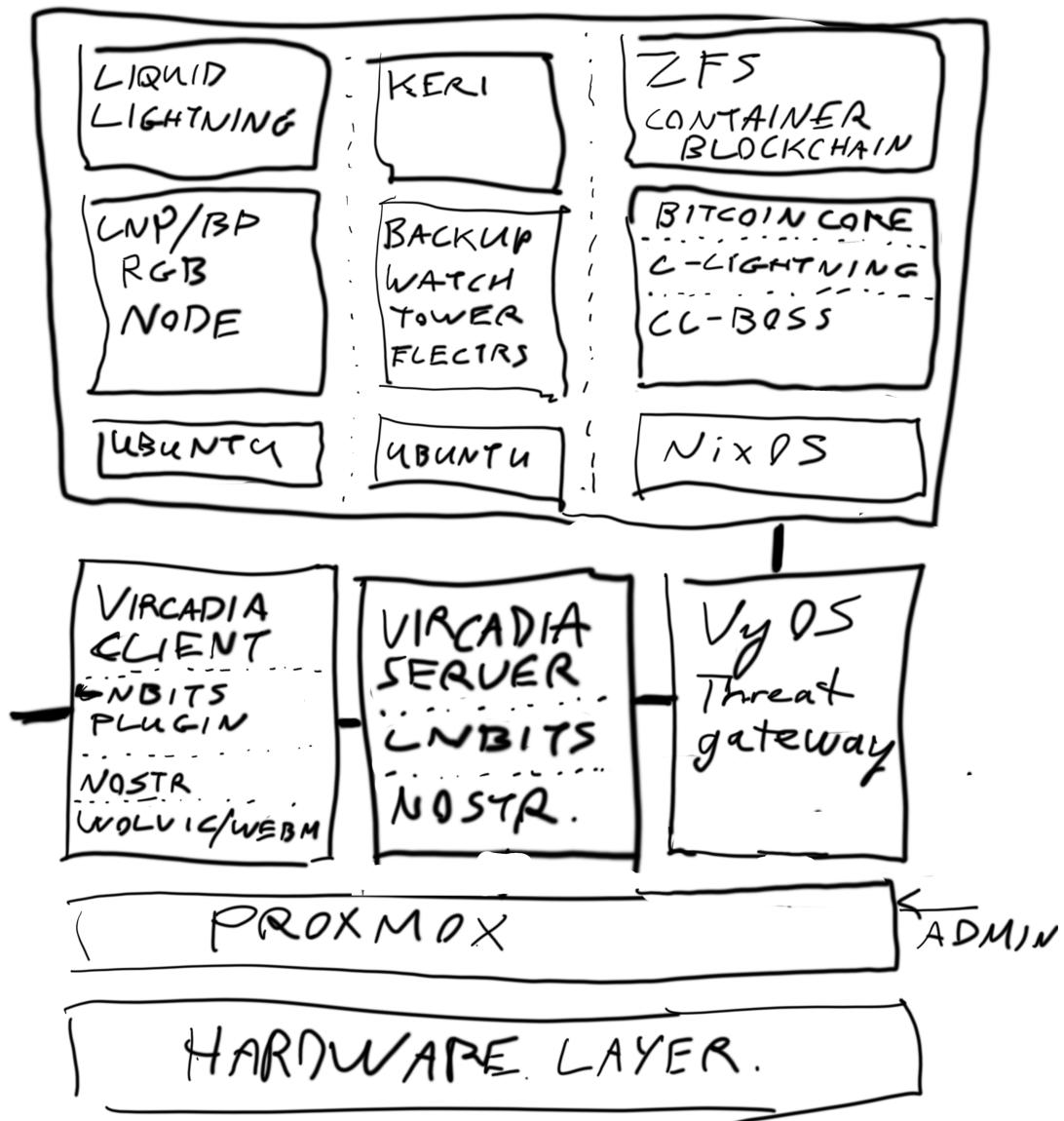
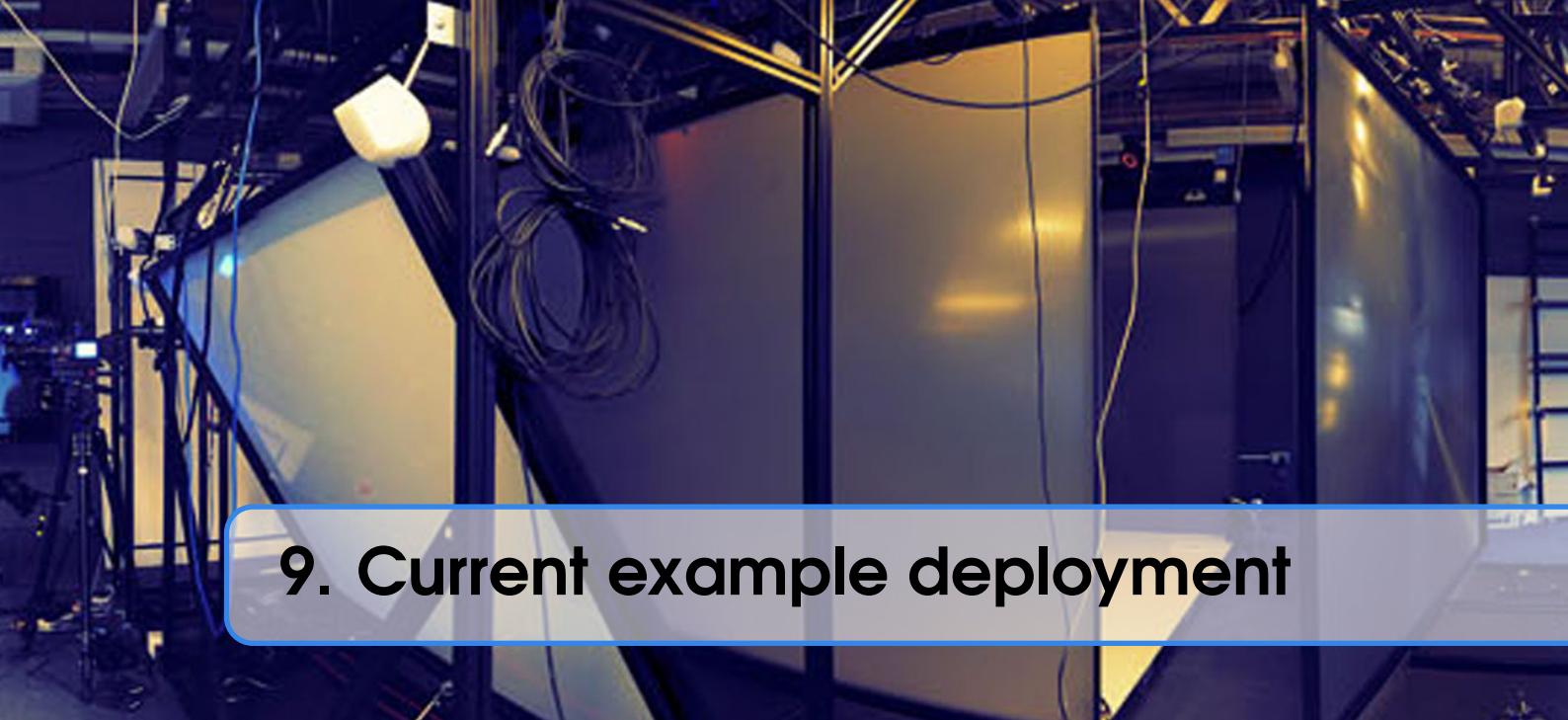


Figure 8.1: Proposed deployment of the software within the VMs on a single hardware system.



9. Current example deployment

9.1 GitHub

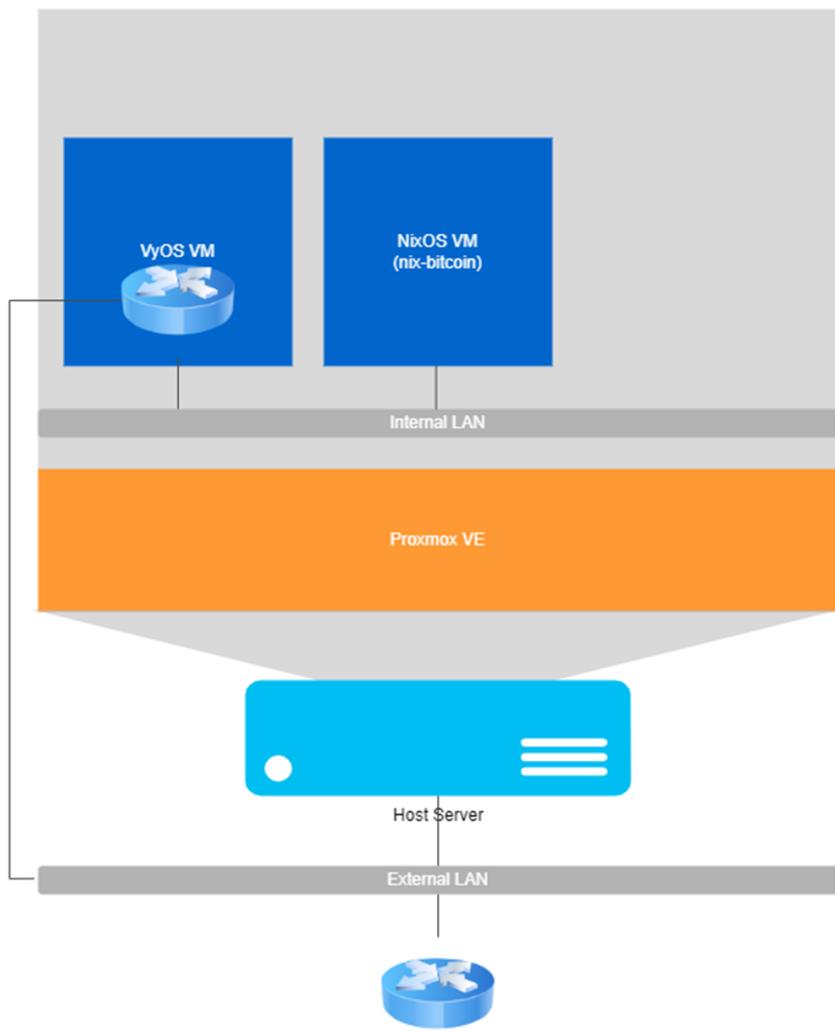


Figure 9.1: Current diagram of the proxmox as seen on the github.



10. Our proposition

This chapter identifies an intersectional space across the described technologies, and proposes a valuable and novel software stack, which can enable exploration and product development. It is useful to briefly look at some of the potential applications which might benefit from value and trust exchange within an global shared social space.

10.1 Potential applications

- Art / NFT galleries with instant sales

This application allows artists and content creator communities to display and sell NFT and fungible art to global consumer audiences, instantly.

- Large scale conference center
 - Academic conferences
 - Political conference
 - Commercial expo

In a hypothetical virtual conference centre a true marketplace of ideas could be enacted, with participants being paid directly by their audience based on the proximity to the presentation.

- Group entertainment
 - Global social puzzle gaming with prizes
 - Music festivals and gigs - Pay live artists and DJs in real time depending on location within the extended landscape of the venue. Split to music producer a portion of the value
 - Mixed reality theatre
 - murder mystery
 - Mixed reality live immersive MMORG games
 - Bingo and mass participation gameshows
 - Immersive brand storytelling metaverses
 - Escape rooms
- Debating townhall meetings (with voting etc)
- Mixed reality information metaverse
 - AR based city tours with collectibles

- AR based collectibles for trails and heritage (museums, libraries) with location specific donations.
- Retail applications
 - Proxy for physical market
 - AR home delivery market interface within physical marketplaces
- Global course / Education provision
 - Explore the universe as a group of spaceship or planet characters
 - Explore biology and physics at a microscopic and nanoscopic level
- Micro tasking marketplace
- Code bounty marketplace
- Micro remittance role sharing (business PA / reception etc)
- Careers fair with credential passing
- Auctions in mixed reality
- eSports and live sports
- Gambling, betting markets, and financial leverage markets

10.1.1 Global cybersec course delivery

Isolating and building out one example here:

- Elements for the infrastructure: Economic layer, asset layer, content interface, user management, data storage, microsites loaded in Wolvin and webm, accessibility schema, network security, backups, secure messaging. Deployable framework with high modularity. Some more ossified elements for surity, some less so for malleability and open opportunity. Figure 10.2.
- Course delivery in XR, how to we develop a platform, marketplace, framework for open contribution.
- WebXR, Vircadia, any snap in metaverse middleware that is free and open source (action to compare the two).
- Define an interface schema for bolting in any commercial or FOSS metaverse engine.
- VR marketplace (outside the scope of the VR engine) without a trusted third party.
- Cryptographically managed learning deliverables (coursework as NFT).
- Secure messaging and group messaging using cryptographic keys. Check this stuff with the distributed computing science people in the group (action on John)
- work toward an exemplar MVP which is then "in the wild"
- Platform for educators
- Define scheme, documentation, best practice, interfaces, functional objects, pedagogy, accessibility, multi-language.
- Define user management system for educators and client learners.
- Identify the pain points which current FOSS elements which need development time/money
- separate the UI/engine from the graphical assets, and the educational / pedagogical components, accessibility, and the value and asset transfer layers.
- Desktop systems are the primary target (low end system)
- define schema for accessibility. Colour, subtitles, immersion concerns which can be applied to metaverse rooms through API?
- Start to define the hybrid presentation model we favour. Avatars? Micro sites? A combination of the two? Balance of guided vs unguided experience. Do we need to test the correct way to do delivery? Is there prior art we can draw on? I feel I should

know. Is this part of the research that's being done here?

- Big work package on schema vs key and user management to enforce rules in spaces. Only participants who have provably paid should have access to learning material, the ability to input into the assessment system, and the tokenised learning outcome 'NFT' or proof.
- Proof that XR system improve learning outcomes. Also that the proposed systems for micro-transactions and user and schema management give additional headroom for teaching.

Notes on build-out The world database in the shared rooms in the metaverse is the global object master, educational materials, videos, audio content and branded objects are fungible tokens authentically proved by rgb client side validation between parties, only validated ones will be persisted in shared rooms like conferences and classes according to the room schema. That allows educators to monetise their content. That can work on lightning. NFT objects between parties like content crafted by participants (coursework, homework) are not on lightning and will attract main chain fees but are rarer. User authentication and communication will be through nostr.

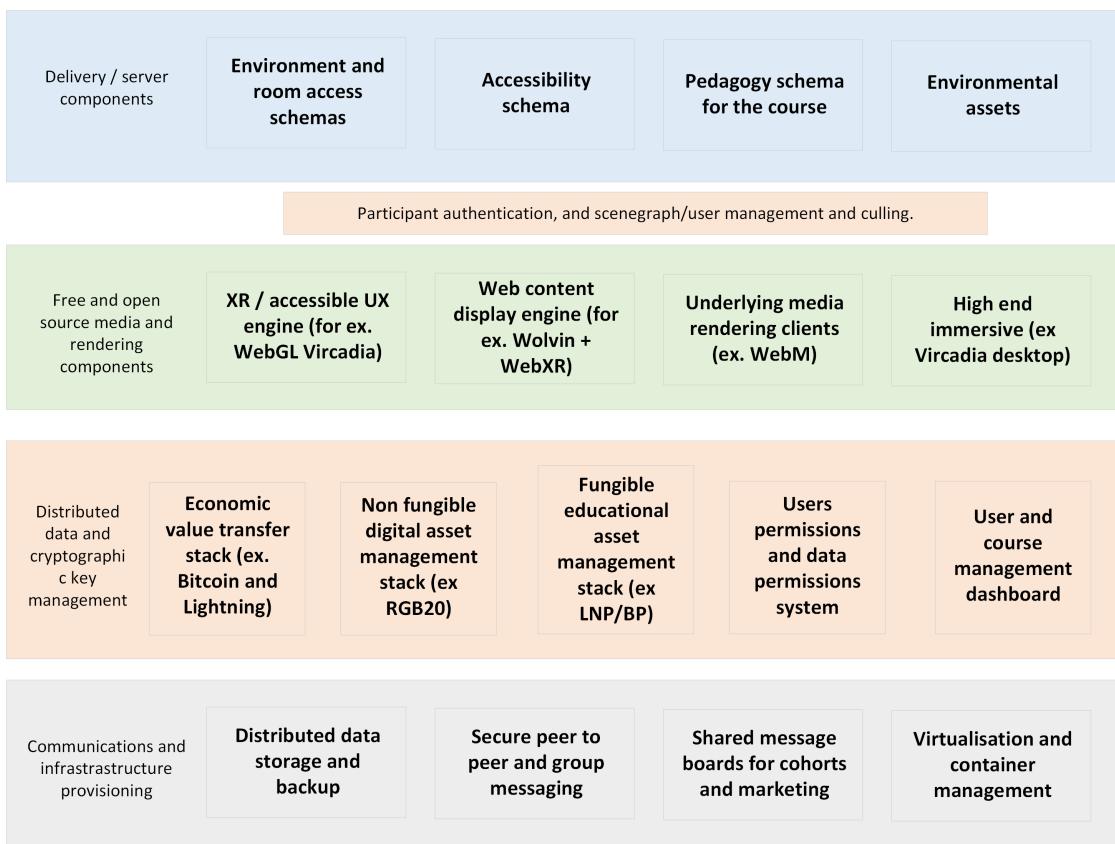


Figure 10.1: Functional elements for infrastructure.

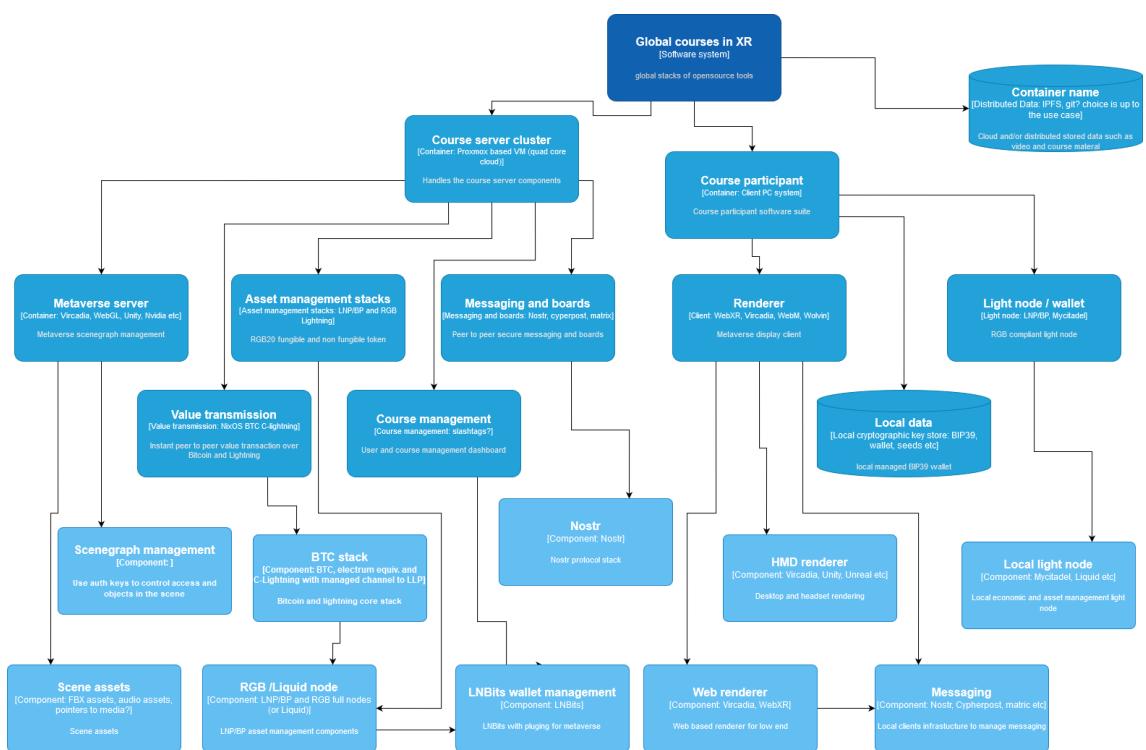


Figure 10.2: Client server C4 diagrams.

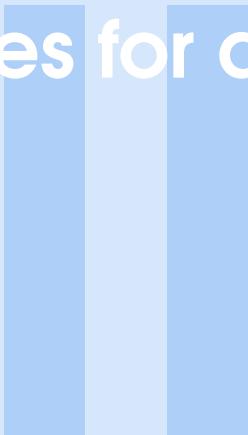


11. Conclusions



12. Future work

Guides for deploying the software



12.1 Lab

12.1.1 Overview

This document details the process of creating the system detailed in the accompanying paper. It is intended to be complete. It is a how to guide.

12.1.1.1 Summary of software

Summarise the software and functionality

12.1.2 Prerequisites

Ensure that the BIOS / firmware / etc of the hardware you intend to use is up to date.

12.1.3 Network details

In the example setup provided here there are currently two networks:

1. The virtual server resides in a LAN with the following details:

192.168.x.0/24

Replace x with an integer between 0 and 254

This LAN has a gateway to the Internet and DNS server configured. Of course, it could be replaced with a direct connection to the Internet, though for research and development purposes it is often better to work within a clean LAN and manage access to the Internet as required.

2. There is a virtual network configured on the virtual machine host upon which virtual machines can reside:

This virtual network is not configured to bridge with the physical network adapter rather a virtual machine is configured as a gateway to route IP traffic through. This provides a level of isolation. More on this later (@todo).

12.1.4 Server configuration

12.1.4.1 Server hardware details

@todo

12.1.4.2 Disk configuration details

@todo

12.1.5 Proxmox VE

12.1.5.1 Installation and configuration

Version used: 7.1

Keep in mind that this setup uses the Proxmox VE installer (<https://www.proxmox.com/en/proxmox-ve/get-started>) which, as noted on the site, is a bare-metal installer and will erase all data on at least one disk. There are alternative methods to install Proxmox VE but these are not covered here.

A brief summary of the steps taken using Proxmox VE 7.1:

Dialogue 1 Choose the target harddisk (/dev/sda in this case).

Dialogue 2 Select country, time zone, keyboard layout.

Dialogue 3 Set a password (this is the root password, see proxmox hardening section), and email address.

Dialogue 4 Select a Network Interface Card (NIC) on which the management interface will be available and provide a hostname, IP address, gateway and DNS server.

In this example the following settings were used:

Hostname: proxmoximus.local IP address: 192.168.x.220 / 24 Gateway: 192.168.x.254
DNS server: 192.168.x.254

Either replace x with the integer used earlier and update the last octet of the gateway and server with that that corresponds to your setup (assuming the setup is local and has a local dns server or forwarder) or configure the values according to your intended setup.

Once the install has completed and the system has rebooted it is time to begin configuring. This is done (almost entirely) via the web interface, in this case, available at <https://proxmoximus.local:8006>

It is also possible to login to a shell via the local terminal and SSH (which is enabled by default @todo: in hardening, add keys and remove ability to login with password).

12.1.5.2 Software updates

If you are in a testing and non-production environment then it is possible access updates without a subscription as detailed here: https://pve.proxmox.com/wiki/Package_Repositories. Update `/etc/apt/sources.list` as detailed under the Proxmox VE No-Subscription Repository. This can be achieved via the local terminal, SSH or web interface (via Shell option).

For example, edit the file:

```
nano /etc/apt/sources.list
```

Add the following:

```
# PVE pve-no-subscription repository provided by proxmox.com,
# NOT recommended for production use
deb http://download.proxmox.com/debian/pve bullseye pve-no-subscription
```

To the existing:

```
deb http://ftp.uk.debian.org/debian bullseye main contrib
```

```
deb http://ftp.uk.debian.org/debian bullseye-updates main contrib
```

```
# security updates
```

```
deb http://security.debian.org bullseye-security main contrib
```

Resulting in:

```
deb http://ftp.debian.org/debian bullseye main contrib
deb http://ftp.debian.org/debian bullseye-updates main contrib
```

```
# PVE pve-no-subscription repository provided by proxmox.com,
# NOT recommended for production use
deb http://download.proxmox.com/debian/pve bullseye pve-no-subscription
```

```
# security updates
```

```
deb http://security.debian.org/debian-security bullseye-security main contrib
```

The Proxmox VE system will now retrieve updates for both itself and the base Debian system.

Then from a shell run:

```
$ apt update  
$ apt upgrade
```

@todo: determine if system needs a reboot

12.1.5.3 Proxmox VE hardening

Links

@todo

Adding users

Add SSH keys and remove ability to login with password

12.1.6 Setup an internal only network in Proxmox VE

From the Web GUI navigate to Datacenter -> your server -> Network

From the menu select Create then Linux Bridge

Input the desired IPv4/CIDR in this case 192.168.y.0/24 and add a comment if desired (“Internal network” was used here). Note that y must not be the same as x previously used.

Name was left as vmbr1

Credit: <https://dannyda.com/2020/06/01/how-to-create-an-internal-only-isolated-network-for-guest-os-virtual-machines-vm-on-proxmox-ve-pve-like-in-vmware-workstation-host-only-network-but-different/>

12.1.7 Install and configure Internet gateway server virtual machine

VyOS was selected (<https://vyos.io/>)

12.1.7.1 Create an ISO of the stable version (as of writing 1.3.0)

@todo: the built version seemed to be a nightly release, is it possible to add a tag to get a stable build?

Follow the build instructions:

<https://docs.vyos.io/en/latest/contributing/build-vyos.html>

This document does not list this version (goes up to 10 “buster”) but Debian 11 “bullseye” was successfully used in this setup.

Run the following commands:

```
$ apt install git  
$ apt install build-essential
```

Follow the instructions here <https://docs.docker.com/engine/install/debian/> to install Docker

Run the following commands:

```
$ git clone -b equuleus --single-branch https://github.com/vyos/vyos-build  
$ docker run --rm -it --privileged -v $(pwd):/vyos -w /vyos vyos/vyos-build:equuleus
```

Then in the Docker terminal run the following commands:

```
./configure --architecture amd64  
sudo make iso
```

12.1.7.2 Upload the ISO image to the Proxmox VE server

1. Via the web GUI navigate to Datacenter -> your server -> local.
2. In the right hand pane select ISO Images and then upload.
3. Upload the ISO image

Tip: you can also pass the checksum to the Proxmox VE upload tool

12.1.7.3 Create VyOS virtual machine

1. From the top right of the web GUI select Create VM
2. In the appearing dialogue type a Name “VyOS” and optionally select advanced and Start at boot
3. On the next tab select the target ISO image
4. On the System tab leave everything as default
5. In the Disk tab leave the defaults (this exceeds requirements <https://docs.vyos.io/en/latest/installation/installing-vyos.html>)
6. On the CPU tab:
 Sockets: 1, Cores: 2
7. On the Memory tab
 Memory: 4096MiB
8. On the Network tab
 Choose the bridge with the internet vmbr0 (it is possible to add the second later) and leave the defaults including firewall
 Confirm all the settings on the next tab but **do not** select start after created
 Navigate to the newly created VM on the left-hand pane then selected Hardware from the menu that is presented on the right. Choose Add and then Network Device. In the dialogue that appears select the Internal network bridge (vmbr1 in this case) that was created earlier and leave all other options as is.
 So, the VM will have the following Network Devices:
 net0: Internet
 net1: Internal only
9. Start the VM and connect the console (top right)
10. Login with vyos and vyos
 Run the command:

```
$ install image
```

11. Follow the instructions
12. Set the CD/DVD to none in Web GUI
13. Reboot

12.1.7.4 Configure VyOS

Open a noVNC window to the host

Login with vyos and vyos

Switch to configure mode:

```
vyos@vyos$ configure  
vyos@vyos#
```

Then configure as desired. Below is configuration used in the setup here (if you use for inspiration do take care to replace the x and y octet values correctly with previously chosen values):

```
set interfaces ethernet eth0 address '192.168.x.221/24'  
set interfaces ethernet eth0 description 'OUTSIDE'  
set protocols static route 0.0.0.0/0 next-hop 192.168.x.254 distance 1  
set service dns forwarding system  
set service dns forwarding name-server 192.168.x.254  
set service dns forwarding listen-address 192.168.y.1  
set service dns forwarding allow-from 192.168.y.0/24  
set system name -server 192.168.x.254  
  
set interfaces ethernet eth1 address '192.168.y.1/24'  
set interfaces ethernet eth1 description 'INSIDE'  
  
set nat source rule 100 outbound-interface eth0  
set nat source rule 100 source address 192.168.y.0/24  
set nat source rule 100 translation address masquerade  
  
set service ssh listen-address 0.0.0.0
```

Once done remember to commit the config (correcting any misconfiguration) and save.

```
commit  
save
```

Inspiration for the above was taken from: [@todo: hardening, IDS, IPS](https://bertvv.github.io/cheat-sheets/VyOS.html)

12.1.8 Install and configure a Debian virtual machine

This VM can be used for various tasks such as software compilation and testing of the networks. In this setup the Debian VM was used to test connectivity to the VyOS gateway and the Internet. It is also used in the subsequent stages to deploy a nix-bitcoin node.

In Proxmox VE create a new virtual machine and configure the network device to use the bridge ‘vmbr1’.

Then install Debian and configure the network adapter within the VM with the following settings:

IP address: 192.168.y.2 Gateway: 192.168.y.1 DNS: 192.168.y.1

Test that the VM has Internet connectivity.

12.1.9 Deploying the nix-bitcoin node

This deployment follows the documentation:

<https://github.com/fort-nix/nix-bitcoin/#get-started>

Take note of the hardware requirements:

<https://github.com/fort-nix/nix-bitcoin/blob/master/docs/hardware.md>

In the main, the install guide (<https://github.com/fort-nix/nix-bitcoin/blob/master/docs/install.md>) is followed verbatim and notes with a reference to particular sections are added where appropriate.

A small exception in regards to this setup is that a separate virtual disk (located on a different physical drive mirror (RAID 1)) was used to store the bitcoin database - this is optional and details are provided on how to achieve it. Also detailed is how to configure the network when using the minimal image.

12.1.9.1 Acquiring NixOS

Following [section 1.1](#) make sure the latest NixOS is obtained i.e. do not just copy the whole wget command outright and make sure to verify the hash against trusted sources before using the image.

Download the minimal ISO image (<https://nixos.org/download.html>)

Verify the hash

Upload the ISO to Proxmox VE server

12.1.9.2 Create a new VM

Name: NixOS

Follow the setup and leave everything as default until the CPU page. The following configuration was used, which should exceed the minimum requirements:

Cores: 4

Memory: 4096MiB = 4.2GB

Network: vmbr1 (Internal Network)

Do NOT check the select the start the VM checkbox

Next, an additional drive will be configured in Proxmox VE. This will then be used to store the bitcoin database within the NixOS VM.

Select Datacenter -> server name and then from the right pane Disks -> LVM-Thin. Then select Create: Thinspace

From the dialogue select the disk and type a name “data” was used in this setup. This provisions a vg with the name *data* and a name *data* @todo: review

Navigate back to the VM created and choose Hardware and then Add -> Hard Disk

Choose “data” from Storage and then set the size to 560 GiB which equates to about 600GB

Now, continue from section 1.3 in the install instructions

Start the VM and connect a console

`sudo -i`

With the SeaBios that was used in this setup the file does not exist and Legacy Boot (MBR) should be followed (option 2)

Note: no consideration is currently given for encrypted partitions within the Proxmox VE setup

Enable the OpenSSH daemon

```
services.openssh.permitRootLogin = "yes";
```

Configure the network config in configuration.nix (remember to replace y with the chosen value)

```
networking.useDHCP = false;
networking.interfaces.ens18.useDHCP = false;

networking.interfaces.ens18.ipv4.addresses = [ {
    address= "192.168.y.3";
    prefixLength = 24;
} ];
networking.defaultGateway = "192.168.y.1";
networking.nameservers = ["192.168.y.1"];
networking.hostName = "nixicon";
```

Although the IP above will be assigned once the nix-bitcoin is deployed the installation cannot continue without a connection to the Internet so that needs to be configured:

```
$ ifconfig ens18 192.168.y.3
$ ifconfig ens18 255.255.255.0
$ ip route add 192.168.y.0/24 dev ens18 scope link src 192.168.y.3
```

Then add the nameserver:

```
nano /etc/resolv.conf
```

Add:

```
nameserver 192.168.y.1
```

Once the above is complete and successful networking is verified
Run the following command:

```
nixos-install
```

and then reboot.

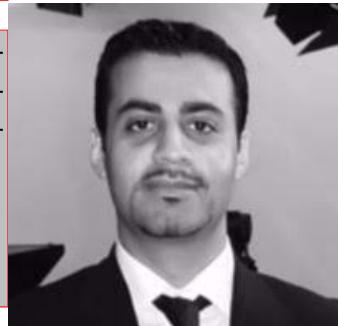
12.2 Acknowledgements and thanks

12.3 Author Biographies

Dr John O'Hare is a results driven, certified Prince2 Agile Practitioner. Leveraging proven analytical ability, and drawing on 23 years of experience at the University of Salford. Successful as a leader and an influential team member in both project and customer-facing roles. As a product manager he specialises in systems design, procurement, tendering and bid writing for research funding, running complex heterogeneous research systems, research and development, and supporting academic staff / research students to undertake theirs. Completed a PhD in “[Attention in Telepresence](#)”, uniting the gaze of remote collaborators, through furniture. Recently pursuing research opportunities in value transfer mechanisms for ‘Metaverses’.



Dr Umran Ali currently works as a senior lecturer in creative media, and continues to explore virtual natural environment design through teaching and research, maintaining a deep interest in the meaning, impact, and design of natural spaces, in particular around video games. A keen video game collector and player, and a landscape photographer.



Bibliography

Bibliography

- [1] Kiku Jones and Lori NK Leonard. "Trust in consumer-to-consumer electronic commerce". In: *Information & management* 45.2 (2008), pages 88–95 (cited on page 15).
- [2] Mark Berners-Lee Tim; Fischetti. *Weaving the web*. https://archive.org/details/isbn_9780062515872/mode/2up. Accessed: 2021-02-10. 1999 (cited on page 17).
- [3] Sean S Costigan. *World Without Mind: The Existential Threat of Big Tech (Franklin Foer)*. 2018 (cited on page 17).
- [4] Bryan Ford, Pyda Srisuresh, and Dan Kegel. "Peer-to-Peer Communication Across Network Address Translators." In: *USENIX Annual Technical Conference, General Track*. 2005, pages 179–192 (cited on page 23).
- [5] Enis Karaarslan and Eylul Adiguzel. "Blockchain based DNS and PKI solutions". In: *IEEE Communications Standards Magazine* 2.3 (2018), pages 52–57 (cited on page 24).
- [6] Michel Rauchs et al. "Distributed ledger technology systems: A conceptual framework". In: *Available at SSRN* 3230013 (2018) (cited on page 26).
- [7] Whitfield Diffie and Martin Hellman. "New directions in cryptography". In: *IEEE transactions on Information Theory* 22.6 (1976), pages 644–654 (cited on pages 26, 34).
- [8] Ralph C Merkle. "Secure communications over insecure channels". In: *Communications of the ACM* 21.4 (1978), pages 294–299 (cited on page 26).
- [9] David Burnham. *The rise of the computer state*. Random House Inc., 1983 (cited on page 26).
- [10] David Chaum. "Security without identification: Transaction systems to make big brother obsolete". In: *Communications of the ACM* 28.10 (1985), pages 1030–1044 (cited on page 26).
- [11] Don Lavoie. "Prefatory Note: The Origins of" The Agorics Project". In: *Market Process*, v8, Spring (1990), pages 116–119 (cited on page 26).
- [12] Phil Salin. *Costs and Computers*. 1991. URL: <http://cdn.oreillystatic.com/radar/r1/11-91.pdf> (cited on page 26).
- [13] Various. *Cypher Punks Mailing List Archives*. 1990. URL: <https://mailing-list-archive.cryptoanarchy.wiki> (cited on page 26).
- [14] Adam Back et al. "Hashcash-a denial of service counter-measure". In: (2002) (cited on page 26).
- [15] Wei Dai. "b-money, 1998". In: URL <http://www.weidai.com/bmoney.txt>. (Last access: 08.04. 2019) (1998) (cited on page 26).

- [16] Nick Szabo. “Formalizing and securing relationships on public networks”. In: *First monday* (1997) (cited on page 26).
- [17] Jon Callas et al. *OpenPGP message format*. Technical report. RFC 2440, November, 1998 (cited on page 26).
- [18] Satoshi Nakamoto. “Re: Bitcoin P2P e-cash paper”. In: *Email posted to listserv 9* (2008), page 04 (cited on pages 26, 32).
- [19] Fran Casino, Thomas K Dasaklis, and Constantinos Patsakis. “A systematic literature review of blockchain-based applications: Current status, classification and open issues”. In: *Telematics and informatics* 36 (2019), pages 55–81 (cited on page 26).
- [20] David Columbia. “Cryptocurrency Is Garbage. So Is Blockchain.” In: *So Is Blockchain.(June 16, 2020)* (2020) (cited on page 28).
- [21] Vitalik Buterin et al. “Ethereum white paper”. In: *Github repository 1* (2013), pages 22–23 (cited on page 28).
- [22] Yulin Liu et al. “Empirical Analysis of EIP-1559: Transaction Fees, Waiting Time, and Consensus Security”. In: *arXiv preprint arXiv:2201.05574* (2022) (cited on page 31).
- [23] Stuart Haber and W Scott Stornetta. “How to time-stamp a digital document”. In: *Conference on the Theory and Application of Cryptography*. Springer. 1990, pages 437–455 (cited on page 32).
- [24] Yuji Ijiri. “A framework for triple-entry bookkeeping”. In: *Accounting Review* (1986), pages 745–759 (cited on page 32).
- [25] Alessio Faccia and Narcisa Roxana Mosteanu. “Accounting and blockchain technology: from double-entry to triple-entry”. In: *The Business & Management Review* 10.2 (2019), pages 108–116 (cited on page 32).
- [26] Alan Sangster. “The earliest known treatise on double entry bookkeeping by Marino de Raphaeli”. In: *Accounting Historians Journal* 42.2 (2015), pages 1–33 (cited on page 32).
- [27] Igor Makarov and Antoinette Schoar. *Blockchain Analysis of the Bitcoin Market*. Technical report. National Bureau of Economic Research, 2021 (cited on page 32).
- [28] Kyle Croman et al. “On scaling decentralized blockchains”. In: *International conference on financial cryptography and data security*. Springer. 2016, pages 106–125 (cited on page 33).
- [29] Sergi Delgado-Segura et al. “Analysis of the bitcoin utxo set”. In: *International Conference on Financial Cryptography and Data Security*. Springer. 2018, pages 78–91 (cited on page 33).
- [30] Oxford Analytica. “El Salvador bitcoin experiment comes with risks”. In: *Emerald Expert Briefings oxan-db* (2021) (cited on page 33).
- [31] Saul Griffith. *Electrify: An Optimist’s Playbook for Our Clean Energy Future*. MIT Press, 2021 (cited on page 36).
- [32] Dirk G Baur and Josua Oll. “Bitcoin investments and climate change: a financial and carbon intensity perspective”. In: *Finance Research Letters* (2021), page 102575 (cited on page 36).

- [33] Apolline Blandin et al. “3rd global cryptoasset benchmarking study”. In: *Available at SSRN 3700822* (2020) (cited on page 36).
- [34] Carlos L Bastian-Pinto et al. “Hedging renewable energy investments with Bitcoin mining”. In: *Renewable and Sustainable Energy Reviews* 138 (2021), page 110520 (cited on page 38).
- [35] Matteo Benetton, Giovanni Compiani, and Adair Morse. “When Cryptomining Comes to Town: High Electricity-Use Spillovers to the Local Economy”. In: *Available at SSRN 3779720* (2021) (cited on page 38).
- [36] Joseph Poon and Thaddeus Dryja. *The bitcoin lightning network: Scalable off-chain instant payments*. 2016 (cited on page 41).
- [37] Philipp Zabka et al. “Short Paper: A Centrality Analysis of the Lightning Network”. In: (2022) (cited on page 41).
- [38] Glyn Davies. *History of money*. University of Wales Press, 2010 (cited on page 54).
- [39] Dominik Stroukal et al. “Can Bitcoin become money? Its money functions and the regression theorem”. In: *International Journal of Business and Management* 6.1 (2018), pages 36–53 (cited on page 54).
- [40] Brian Roger Tomlinson. “What Was the Third World?” In: *Journal of Contemporary History* 38.2 (2003), pages 307–321 (cited on page 55).
- [41] Ricardo J Caballero, Emmanuel Farhi, and Pierre-Olivier Gourinchas. *Financial crash, commodity prices and global imbalances*. Technical report. National Bureau of Economic Research, 2008 (cited on page 55).
- [42] David E Spiro. *The hidden hand of American hegemony*. Cornell University Press, 2019 (cited on page 55).
- [43] John A Mathews and Mark Selden. “China: The emergence of the Petroyan and the challenge to US dollar hegemony”. In: *The Asia-Pacific Journal* 16.22/3 (2018), pages 1–12 (cited on page 55).
- [44] Yiping Huang. “Understanding China’s Belt & Road initiative: motivation, framework and assessment”. In: *China Economic Review* 40 (2016), pages 314–321 (cited on page 55).
- [45] Peter Gainsford. “Salt and salary: were Roman soldiers paid in salt?” In: *Kiwi Hellenist: Modern Myths about the Ancient World. Retrieved 11* (2017) (cited on page 55).
- [46] Dror Goldberg. “Famous myths of” fiat money””. In: *Journal of Money, Credit and Banking* (2005), pages 957–967 (cited on page 55).
- [47] Andrew J Filardo, Madhusudan S Mohanty, and Ramon Moreno. “Central bank and government debt management: issues for monetary policy”. In: *BIS Paper* 67d (2012) (cited on page 58).
- [48] Richard Cantillon. *Essai sur la nature du commerce en général*. éditeur non identifié, 1756 (cited on page 58).
- [49] Michael David Bordo. “Some aspects of the monetary economics of Richard Cantillon”. In: *Journal of Monetary Economics* 12.2 (1983), pages 235–258 (cited on page 58).

- [50] Eswar S Prasad. *The Future of Money: How the Digital Revolution is Transforming Currencies and Finance*. Harvard University Press, 2021 (cited on page 58).
- [51] Mathilde Maurel and Gunther Schnabl. “Keynesian and Austrian perspectives on crisis, shock adjustment, exchange rate regime and (long-term) growth”. In: *Open Economies Review* 23.5 (2012), pages 847–868 (cited on page 60).
- [52] Daniel Krawisz. “Hyperbitcoinization”. In: *Online verfügbar unter: https://nakamotoin* (2014) (cited on page 62).
- [53] Mick Lockwood. “Exploring value propositions to drive Self-Sovereign Identity adoption”. In: *Frontiers in Blockchain* 4 (2021), page 4 (cited on page 66).
- [54] Hilary McLellan. “Avatars, Affordances, and Interfaces: Virtual Reality Tools for Learning.” In: (1993) (cited on page 76).
- [55] Nicholas Bloom et al. “Does working from home work? Evidence from a Chinese experiment”. In: *Q. J. Econ.* 130.1 (2015), pages 165–218 (cited on page 77).
- [56] Shiv Prakash et al. “Characteristic of enterprise collaboration system and its implementation issues in business management”. In: *International Journal of Business Intelligence and Data Mining* 16.1 (2020), pages 49–65 (cited on page 77).
- [57] Adam Aiken. “Zooming in on privacy concerns: Video app Zoom is surging in popularity. In our rush to stay connected, we need to make security checks and not reveal more than we think”. In: *Index on Censorship* 49.2 (2020), pages 24–27 (cited on page 78).
- [58] R Wolff et al. “Communicating Eye Gaze across a Distance without Rooting Participants to the Spot”. In: *2008 12th IEEE/ACM International Symposium on Distributed Simulation and Real-Time Applications*. Ieee, #oct# 2008, pages 111–118 (cited on pages 78, 88, 99, 103).
- [59] RS Oeppen, G Shaw, and PA Brennan. “Human factors recognition at virtual meetings and video conferencing: how to get the best performance from yourself and others”. In: *British Journal of Oral and Maxillofacial Surgery* (2020) (cited on page 78).
- [60] Dhanraj Vishwanath, Ahna R Girshick, and Martin S Banks. “Why pictures look right when viewed from the wrong place”. In: volume 8. Nature Publishing Group, 2005, pages 1401–1410 (cited on page 78).
- [61] Stuart M Anstis, John W Mayhew, and Tania Morley. “The Perception of Where a Face or Television Portrait is Looking”. In: volume 82. JSTOR, 1969, pages 474–489 (cited on page 78).
- [62] William Hyde Wollaston. “On the apparent direction of eyes in a portrait”. In: JSTOR, 1824, pages 247–256 (cited on page 78).
- [63] Jack M Loomis et al. “Psychophysics of perceiving eye-gaze and head direction with peripheral vision: implications for the dynamics of eye-gaze behavior”. en. In: *Perception* 37.9 (2008), pages 1443–1457 (cited on pages 78, 82).
- [64] Chris Fullwood and Gwyneth Doherty-Sneddon. “Effect of gazing at the camera during a video link on recall”. en. In: *Appl. Ergon.* 37.2 (#mar# 2006), pages 167–175 (cited on page 78).

- [65] Samer Al Moubayed, Jens Edlund, and Jonas Beskow. “Taming Mona Lisa”. In: volume 1. 2012, pages 1–25 (cited on pages 78, 105).
- [66] David Nguyen and John Canny. “MultiView: spatially faithful group video conferencing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’05. Portland, Oregon, USA: Association for Computing Machinery, #apr# 2005, pages 799–808 (cited on pages 78, 92).
- [67] Roel Vertegaal and Yaping Ding. “Explaining effects of eye gaze on mediated group conversations: amount or synchronization?” In: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. CSCW ’02. New Orleans, Louisiana, USA: Association for Computing Machinery, #nov# 2002, pages 41–48 (cited on pages 78, 81).
- [68] R van Eijk et al. “Human sensitivity to eye contact in 2D and 3D videoconferencing”. In: *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*. #jun# 2010, pages 76–81 (cited on page 78).
- [69] Simon W Bock, Peter Dicke, and Peter Thier. “How precise is gaze following in humans?” en. In: *Vision Res.* 48.7 (#mar# 2008), pages 946–957 (cited on page 78).
- [70] Milton Chen. “Leveraging the asymmetric sensitivity of eye contact for video-conference”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’02. Minneapolis, Minnesota, USA: Association for Computing Machinery, #apr# 2002, pages 49–56 (cited on pages 78, 86).
- [71] Michael Argyle. *Bodily communication*. Methuen, 1988 (cited on pages 79–81, 83, 85).
- [72] Mattias Heldner and Jens Edlund. “Pauses, gaps and overlaps in conversations”. In: *J. Phon.* 38.4 (#oct# 2010), pages 555–568 (cited on page 79).
- [73] C L Kleinke. “Gaze and eye contact: a research review”. en. In: *Psychol. Bull.* 100.1 (#jul# 1986), pages 78–100 (cited on pages 79, 81, 82).
- [74] Stella Ting-Toomey and Leeva C Chung. *Understanding intercultural communication*. Oxford University Press New York, NY, 2012 (cited on page 79).
- [75] Kazuhiko Otsuka, Yoshinao Takemae, and Junji Yamato. “A probabilistic inference of multiparty-conversation structure based on Markov-switching models of gaze patterns, head directions, and utterances”. In: *Proceedings of the 7th international conference on Multimodal interfaces*. ICMI ’05. Toronto, Italy: Association for Computing Machinery, #oct# 2005, pages 191–198 (cited on pages 79, 105).
- [76] Yasuhiro Katagiri. “Aiduti in Japanese multi-party design conversations”. In: *Proceedings of the Workshop on Embodied Language Processing*. 2007, pages 9–16 (cited on page 79).
- [77] Jack M Loomis, Roberta L Klatzky, and Nicholas A Giudice. “-Sensory Substitution of Vision: Importance of Perceptual and Cognitive Processing”. In: *Assistive technology for blindness and low vision*. CRC Press, 2012, pages 179–210 (cited on page 80).
- [78] Charles Goodwin. “Action and embodiment within situated human interaction”. In: *J. Pragmat.* 32.10 (#sep# 2000), pages 1489–1522 (cited on page 80).

- [79] Marco Gillies, Mel Slater, et al. “Non-verbal communication for correlational characters”. In: (2005) (cited on page 80).
- [80] Michael Argyle and Mark Cook. *Gaze and Mutual Gaze*. en. Cambridge University Press, #jan# 1976 (cited on pages 80, 81, 83).
- [81] Michael Argyle and Janet Dean. “Eye-contact, Distance And Affiliation”. In: JSTOR, 1965, pages 289–304 (cited on pages 80, 83).
- [82] Michael Argyle and Roger Ingham. *Gaze, Mutual Gaze, and Proximity*. 1969 (cited on pages 80–83).
- [83] A Kendon. “Some functions of gaze-direction in social interaction.” In: *Acta Psychol.* 26.1 (1967), pages 22–63 (cited on pages 80, 105).
- [84] Andrew F Monk and Caroline Gale. “A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation”. In: *Discourse Process*. 33.3 (#may# 2002), pages 257–278 (cited on pages 80, 87, 103).
- [85] B J Hedge, B S Everitt, and Christopher D Frith. “The Role Of Gaze In Dialogue”. In: volume 42. Elsevier, 1978, pages 453–475 (cited on pages 80, 85).
- [86] D G Novick, B Hansen, and K Ward. “Coordinating turn-taking with gaze”. In: *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*. Volume 3. #oct# 1996, 1888–1891 vol.3 (cited on pages 80, 85).
- [87] Roel Vertegaal, Gerrit Van der Veer, and Harro Vons. “Effects of gaze on multi-party mediated communication”. In: *Graphics interface*. Morgan Kaufmann, 2000, pages 95–102 (cited on page 80).
- [88] Roel Vertegaal et al. “Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: Association for Computing Machinery, #mar# 2001, pages 301–308 (cited on pages 80, 81).
- [89] S R Langton. “The mutual influence of gaze and head orientation in the analysis of social attention direction”. en. In: *Q. J. Exp. Psychol. A* 53.3 (#aug# 2000), pages 825–845 (cited on pages 80–82).
- [90] Alex Colburn, Michael F Cohen, and Steven Drucker. “The Role Of Eye Gaze In Avatar Mediated Conversational Interfaces”. In: 2000 (cited on page 80).
- [91] C Neil Macrae et al. “Are you looking at me? Eye gaze and person perception”. en. In: *Psychol. Sci.* 13.5 (#sep# 2002), pages 460–464 (cited on page 80).
- [92] Lawrence A Symons et al. “What are you looking at? Acuity for triadic eye gaze”. In: volume 131. 2004, pages 451–469 (cited on pages 80, 85).
- [93] Nathan L Kluttz et al. “The effect of head turn on the perception of gaze”. en. In: *Vision Res.* 49.15 (#jul# 2009), pages 1979–1993 (cited on pages 80, 82).
- [94] Hugh R Wilson et al. “Perception of head orientation”. In: volume 40. Elsevier, 2000, pages 459–472 (cited on pages 81, 82, 85).
- [95] Johann Schrammel et al. ““Look!”: using the gaze direction of embodied agents”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: Association for Computing Machinery, #apr# 2007, pages 1187–1190 (cited on page 81).

- [96] Rainer Stiefelhagen, Jie Yang, and Alex Waibel. “Estimating focus of attention based on gaze and sound”. In: *Proceedings of the 2001 workshop on Perceptive user interfaces*. PUI ’01. Orlando, Florida, USA: Association for Computing Machinery, #nov# 2001, pages 1–9 (cited on pages 81, 85, 103).
- [97] R Stiefelhagen. “Tracking focus of attention in meetings”. In: *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*. #oct# 2002, pages 273–280 (cited on pages 81, 82, 85, 103).
- [98] W Steptoe et al. “Eye Tracking for Avatar Eye Gaze Control During Object-Focused Multiparty Interaction in Immersive Collaborative Virtual Environments”. In: *2009 IEEE Virtual Reality Conference*. #mar# 2009, pages 83–90 (cited on pages 81, 99).
- [99] William Arthur Hugh Steptoe. “Eye tracking and avatar-mediated communication in immersive collaborative virtual environments”. PhD thesis. University College London (University of London), 2010 (cited on page 81).
- [100] Mark Cook. “Gaze and mutual gaze in social encounters”. In: *Am. Sci.* 65.3 (1977), pages 328–333 (cited on pages 81, 105).
- [101] Sascha Fagel et al. “On the importance of eye gaze in a face-to-face collaborative task”. In: *Proceedings of the 3rd international workshop on Affective interaction in natural environments*. AFFINE ’10. Firenze, Italy: Association for Computing Machinery, #oct# 2010, pages 81–86 (cited on pages 81, 85, 105).
- [102] Leanne S Bohannon et al. “Eye contact and video-mediated communication: A review”. In: *Displays* 34.2 (#apr# 2013), pages 177–185 (cited on page 81).
- [103] N F Troje and U Siebeck. “Illumination-induced apparent shift in orientation of human heads”. en. In: *Perception* 27.6 (1998), pages 671–680 (cited on page 82).
- [104] Steven M Boker et al. “Something in the way we move: Motion dynamics, not perceived sex, influence head movements in conversation”. en. In: *J. Exp. Psychol. Hum. Percept. Perform.* 37.3 (#jun# 2011), pages 874–891 (cited on page 82).
- [105] M St John et al. “The use of 2D and 3D displays for shape-understanding versus relative-position tasks”. en. In: *Hum. Factors* 43.1 (2001), pages 79–98 (cited on page 82).
- [106] Jeremy N Bailenson, Andrew C Beall, and Jim Blascovich. “Gaze And Task Performance In Shared Virtual Environments”. In: volume 13. 2002, pages 313–320 (cited on page 82).
- [107] Rutger Rienks, Ronald Poppe, and Dirk Heylen. “Differences in head orientation behavior for speakers and listeners: An experiment in a virtual environment”. #jan# 2010 (cited on page 82).
- [108] David T Nguyen and John Canny. “More than face-to-face: empathy effects of video framing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’09. Boston, MA, USA: Association for Computing Machinery, #apr# 2009, pages 423–432 (cited on pages 82, 105).
- [109] P Ekman. “Facial expression and emotion”. en. In: *Am. Psychol.* 48.4 (#apr# 1993), pages 384–392 (cited on pages 82, 83).

- [110] S Goldin-Meadow. “The role of gesture in communication and thinking”. In: volume 3. 1999, pages 419–429 (cited on page 82).
- [111] Nicole Chovil. “Discourse ?oriented facial displays in conversation”. In: *Research on Language and Social Interaction* 25.1-4 (#jan# 1991), pages 163–194 (cited on page 83).
- [112] Stefan G Hofmann, Michael Suvak, and Brett T Litz. “Sex differences in face recognition and influence of facial affect”. In: *Pers. Individ. Dif.* 40.8 (#jun# 2006), pages 1683–1690 (cited on pages 83, 84).
- [113] Jana M Iverson and Susan Goldin-Meadow. “Gesture paves the way for language development”. en. In: *Psychol. Sci.* 16.5 (#may# 2005), pages 367–371 (cited on page 83).
- [114] Robert M Krauss, Yihsiu Chen, and Purnima Chawla. “Nonverbal Behavior and Nonverbal Communication: What do Conversational Hand Gestures Tell Us?” In: *Advances in Experimental Social Psychology*. Edited by Mark P Zanna. Volume 28. Academic Press, #jan# 1996, pages 389–450 (cited on page 83).
- [115] A Kleinsmith and N Bianchi-Berthouze. “Affective Body Expression Perception and Recognition: A Survey”. In: *IEEE Transactions on Affective Computing* 4.1 (#jan# 2013), pages 15–33 (cited on page 83).
- [116] Jiazhi Ou et al. “Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’05. Portland, Oregon, USA: Association for Computing Machinery, #apr# 2005, pages 231–240 (cited on page 83).
- [117] Norman Murray et al. “Eye gaze in virtual environments: evaluating the need and initial work on implementation”. In: *Concurr. Comput.* 21.11 (#aug# 2009), pages 1437–1449 (cited on page 83).
- [118] Edward Twitchell Hall and Edward T Hall. *The Hidden Dimension*. Volume 1990. Anchor Books New York, 1969 (cited on pages 83, 84, 103).
- [119] O Michael Watson and Theodore D Graves. “Quantitative Research in Proxemic Behavior”. In: *Am. Anthropol.* 68.4 (#aug# 1966), pages 971–985 (cited on page 83).
- [120] Nicola Bruno and Michela Muzzolini. “Proxemics Revisited: Similar Effects of Arms Length on Men’s and Women’s Personal Distances”. In: *Universal Journal of Psychology* 1.2 (2013), pages 46–52 (cited on page 83).
- [121] Gillian Slessor, Louise H Phillips, and Rebecca Bull. “Age-related declines in basic social perception: evidence from tasks assessing eye-gaze processing”. en. In: *Psychol. Aging* 23.4 (#dec# 2008), pages 812–822 (cited on pages 83, 85, 103).
- [122] Nick Yee, Jeremy N Bailenson, and Kathryn Rickertsen. “A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’07. San Jose, California, USA: Association for Computing Machinery, #apr# 2007, pages 1–10 (cited on page 84).

- [123] Jim Blascovich. “A theoretical model of social influence for increasing the utility of collaborative virtual environments”. In: *Proceedings of the 4th international conference on Collaborative virtual environments*. CVE ’02. Bonn, Germany: Association for Computing Machinery, #sep# 2002, pages 25–30 (cited on page 84).
- [124] Hiroyuki Maeda et al. “Real World Video Avatar: Transmission And Presentation Of Human Figure”. In: *Virtual Reality, 2004. Proceedings. IEEE*. 2004, pages 237–238 (cited on page 84).
- [125] Joseph C Hager and Paul Ekman. “Long-distance of transmission of facial affect signals”. In: *Ethol. Sociobiol.* 1.1 (#oct# 1979), pages 77–82 (cited on pages 84, 103).
- [126] Abigail Sellen, Bill Buxton, and John Arnott. “Using spatial cues to improve videoconferencing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’92. Monterey, California, USA: Association for Computing Machinery, #jun# 1992, pages 651–652 (cited on pages 85, 103).
- [127] Eva Deckers et al. “Designing for perceptual crossing: designing and comparing three behaviors”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’13. Paris, France: Association for Computing Machinery, #apr# 2013, pages 1901–1910 (cited on pages 85, 103).
- [128] J J Gibson and A D Pick. “Perception of another person’s looking behavior”. en. In: *Am. J. Psychol.* 76.3 (#sep# 1963), pages 386–394 (cited on pages 85, 103).
- [129] Gary Bente, W C Donaghy, and Dorit Suwelack. “Sex Differences In Body Movement And Visual Attention: An Integrated Analysis Of Movement And Gaze In Mixed-sex Dyads”. In: volume 22. 1998 (cited on pages 85, 103).
- [130] Roel Vertegaal. “Catching the eye: management of joint attention in cooperative work”. In: 1997 (cited on pages 85, 103).
- [131] D I Fels and P L Weiss. “Toward determining an attention-getting device for improving interaction during video-mediated communication”. In: *Comput. Human Behav.* 16.2 (#mar# 2000), pages 189–198 (cited on page 85).
- [132] Amy Vold et al. “Cross-cutting faultlines of location and shared identity in the intergroup cooperation of partially distributed groups”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’12. Austin, Texas, USA: Association for Computing Machinery, #may# 2012, pages 3101–3110 (cited on pages 85, 103).
- [133] Edward Ellsworth Jones and Richard E Nisbett. “The Actor And The Observer: Divergent Perceptions Of The Causes Of Behavior”. In: General Learning Press Morristown, NJ, 1971 (cited on page 85).
- [134] Alaeddine Mihoub et al. “Learning multimodal behavioral models for face-to-face social interaction”. In: *Journal on Multimodal User Interfaces* 9.3 (#sep# 2015), pages 195–210 (cited on page 85).
- [135] T L Chartrand and J A Bargh. “The chameleon effect: the perception-behavior link and social interaction”. en. In: *J. Pers. Soc. Psychol.* 76.6 (#jun# 1999), pages 893–910 (cited on page 85).

- [136] Peter Kollock. “Social Dilemmas: The Anatomy of Cooperation”. In: *Annu. Rev. Sociol.* 24.1 (#aug# 1998), pages 183–214 (cited on page 86).
- [137] Elisabeth Cuddihy and Deborah Walters. “Embodied interaction in social virtual environments”. In: *Proceedings of the third international conference on Collaborative virtual environments*. CVE ’00. San Francisco, California, USA: Association for Computing Machinery, #sep# 2000, pages 181–188 (cited on page 86).
- [138] Jeffrey T Hancock, Jennifer Thom-Santelli, and Thompson Ritchie. “Deception and design: the impact of communication technology on lying behavior”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’04. Vienna, Austria: Association for Computing Machinery, #apr# 2004, pages 129–134 (cited on page 86).
- [139] Håkan J Holm and Toshiji Kawagoe. “Face-to-face lying – An experimental study in Sweden and Japan”. In: *J. Econ. Psychol.* 31.3 (#jun# 2010), pages 310–321 (cited on page 86).
- [140] D J Roberts et al. “withyou—An Experimental End-to-End Telepresence System Using Video-Based Reconstruction”. In: *IEEE J. Sel. Top. Signal Process.* 9.3 (#apr# 2015), pages 562–574 (cited on pages 86, 96, 98, 99, 101, 104).
- [141] Paul M Aoki et al. “The mad hatter’s cocktail party: a social mobile audio space supporting multiple simultaneous conversations”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’03. Ft. Lauderdale, Florida, USA: Association for Computing Machinery, #apr# 2003, pages 425–432 (cited on pages 86, 107).
- [142] Criminisi et al. “Gaze manipulation for one-to-one teleconferencing”. In: *Proceedings Ninth IEEE International Conference on Computer Vision*. Nice, #oct# 2003, 191–198 vol.1 (cited on page 86).
- [143] David J Roberts et al. “Estimating the gaze of a virtuality human”. en. In: *IEEE Trans. Vis. Comput. Graph.* 19.4 (#apr# 2013), pages 681–690 (cited on pages 87, 96, 98).
- [144] J Gemmell et al. “Gaze awareness for video-conferencing: a software approach”. In: *IEEE Multimedia* 7.4 (#oct# 2000), pages 26–35 (cited on page 87).
- [145] Ederyn Williams. “Experimental comparisons of face-to-face and mediated communication: A review”. In: *Psychol. Bull.* 84.5 (1977), page 963 (cited on page 87).
- [146] C Edigo. “Videoconferencing as a technology to support group work: A review of its failure”. In: *Proceedings of the ACM conf. on Computer-Supported Cooperative Work*. 1988 (cited on page 87).
- [147] Tomislav Pejsa et al. “Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment”. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW ’16. San Francisco, California, USA: Association for Computing Machinery, #feb# 2016, pages 1716–1725 (cited on pages 87, 101, 102).
- [148] Dennis Beck et al. “Synthesizing presence: A multidisciplinary review of the literature”. In: volume 3. 2011 (cited on page 88).

- [149] M J Schuemie et al. “Research on presence in virtual reality: a survey”. en. In: *Cyberpsychol. Behav.* 4.2 (#apr# 2001), pages 183–201 (cited on page 88).
- [150] Mel Slater. “Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire”. In: *Presence: Teleoperators and Virtual Environments* 8.5 (#oct# 1999), pages 560–565 (cited on page 88).
- [151] Frank Biocca. “The Cyborg Dilemma : Embodiment in Virtual Environments”. In: 1997, pages 12–26 (cited on page 88).
- [152] Charlotte N Gunawardena and Frank J Zittle. “Social presence as a predictor of satisfaction within a computer ?mediated conferencing environment”. In: *Am. J. Distance Educ.* 11.3 (#jan# 1997), pages 8–26 (cited on pages 88, 89).
- [153] Kristen Nowak. “Defining and differentiating copresence, social presence and presence as transportation”. In: *Presence 2001 Conference, Philadelphia, PA.* 2001, pages 1–23 (cited on pages 88, 89).
- [154] Saniye Tugba Bulu. “Place presence, social presence, co-presence, and satisfaction in virtual worlds”. In: *Comput. Educ.* 58.1 (#jan# 2012), pages 154–161 (cited on page 88).
- [155] Frank Biocca, Chad Harms, and Judee Burgoon. “Toward a more robust theory and measure of social presence: Review and suggested criteria”. In: volume 12. MIT press, 2003, pages 456–480 (cited on pages 88, 89).
- [156] D Randy Garrison, Terry Anderson, and Walter Archer. “Critical Inquiry In A Text-based Environment: Computer Conferencing In Higher Education”. In: volume 2. Elsevier, 1999, pages 87–105 (cited on page 88).
- [157] Yevgenia Bondareva and Don Bouwhuis. “Determinants of social presence in videoconferencing”. In: *AVI2004 Workshop on Environments for Personalized Information Access.* 2004, pages 1–9 (cited on page 89).
- [158] Mel Slater. “How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments”. In: *Presence: Teleoperators and Virtual Environments* 13.4 (#aug# 2004), pages 484–493 (cited on page 89).
- [159] Chad Harms and Frank Biocca. “Internal consistency and reliability of the networked minds measure of social presence”. In: (2004) (cited on page 89).
- [160] Jörg Hauber et al. “Social presence in two-and three-dimensional videoconferencing”. In: (2005) (cited on page 89).
- [161] Jakob Bardram et al. “ReticularSpaces: activity-based computing support for physically distributed and collaborative smart spaces”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI ’12. Austin, Texas, USA: Association for Computing Machinery, #may# 2012, pages 2845–2854 (cited on pages 89, 90).
- [162] Piotr D Adamczyk and Michael B Twidale. “Supporting multidisciplinary collaboration: requirements from novel HCI education”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI ’07. San Jose, California, USA: Association for Computing Machinery, #apr# 2007, pages 1073–1076 (cited on pages 89, 90).

- [163] Hiroshi Ishii, Minoru Kobayashi, and Jonathan Grudin. “Integration of interpersonal space and shared workspace: ClearBoard design and experiments”. #oct# 1993 (cited on pages 89, 90).
- [164] Holger Regenbrecht and Tobias Langlotz. “Mutual Gaze Support in Videoconferencing Reviewed”. In: *Communications of the Association for Information Systems* 37.1 (2015), page 45 (cited on page 90).
- [165] Anthony Tang et al. “Three’s company: understanding communication channels in three-way distributed collaboration”. In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. CSCW ’10. Savannah, Georgia, USA: Association for Computing Machinery, #feb# 2010, pages 271–280 (cited on page 90).
- [166] Edward Tse et al. “How pairs interact over a multimodal digital table”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’07. San Jose, California, USA: Association for Computing Machinery, #apr# 2007, pages 215–218 (cited on page 90).
- [167] Anthony Tang et al. “Collaborative coupling over tabletop displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’06. Montréal, Québec, Canada: Association for Computing Machinery, #apr# 2006, pages 1181–1190 (cited on page 91).
- [168] James Norris, Holger Schnädelbach, and Guoping Qiu. “CamBlend: an object focused collaboration tool”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’12. Austin, Texas, USA: Association for Computing Machinery, #may# 2012, pages 627–636 (cited on page 91).
- [169] Paul Luff et al. “Hands on hitchcock: embodied reference to a moving scene”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: Association for Computing Machinery, #may# 2011, pages 43–52 (cited on page 91).
- [170] Izdihar Jamil et al. “The effects of interaction techniques on talk patterns in collaborative peer learning around interactive tables”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: Association for Computing Machinery, #may# 2011, pages 3043–3052 (cited on page 91).
- [171] Jörg Hauber et al. “Spatiality in videoconferencing: trade-offs between efficiency and social presence”. In: *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. CSCW ’06. Banff, Alberta, Canada: Association for Computing Machinery, #nov# 2006, pages 413–422 (cited on page 91).
- [172] Peter Kauff and Oliver Schreer. “An immersive 3D video-conferencing system using shared virtual team user environments”. In: *Proceedings of the 4th international conference on Collaborative virtual environments*. CVE ’02. Bonn, Germany: Association for Computing Machinery, #sep# 2002, pages 105–112 (cited on page 92).

- [173] Yusuke Ichikawa et al. “MAJIC Videoconferencing System: Experiments, Evaluation and Improvement”. In: *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW ’95: 10–14 September, 1995, Stockholm, Sweden*. Edited by Hans Marmolin, Yngve Sundblad, and Kjeld Schmidt. Dordrecht: Springer Netherlands, 1995, pages 279–292 (cited on page 92).
- [174] Alexander Kulik et al. “C1x6: a stereoscopic six-user display for co-located collaboration in shared virtual environments”. In: *Proceedings of the 2011 SIGGRAPH Asia Conference*. Volume 30. SA ’11 Article 188. Hong Kong, China: Association for Computing Machinery, #dec# 2011, pages 1–12 (cited on page 92).
- [175] Kazuhiro Otsuka et al. “MM+Space: n x 4 degree-of-freedom kinetic display for recreating multiparty conversation spaces”. In: *Proceedings of the 15th ACM on International conference on multimodal interaction*. ICMI ’13. Sydney, Australia: Association for Computing Machinery, #dec# 2013, pages 389–396 (cited on page 92).
- [176] Stefan Marti and Chris Schmandt. “Physical embodiments for mobile communication agents”. In: *Proceedings of the 18th annual ACM symposium on User interface software and technology*. UIST ’05. Seattle, WA, USA: Association for Computing Machinery, #oct# 2005, pages 231–240 (cited on page 92).
- [177] Masahiro Mori. “The uncanny valley”. In: *Energy* 7.4 (1970), pages 33–35 (cited on pages 92, 93).
- [178] C Bartneck et al. “Is The Uncanny Valley An Uncanny Cliff?” In: *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*. #aug# 2007, pages 368–373 (cited on page 92).
- [179] Wade J Mitchell et al. “A mismatch in the human realism of face and voice produces an uncanny valley”. en. In: *Iperception* 2.1 (#mar# 2011), pages 10–12 (cited on page 93).
- [180] Christoph Bartneck et al. “My robotic doppelgänger-A critical look at the uncanny valley”. In: *Robot and Human Interactive Communication*, 2009. 2009, pages 269–276 (cited on page 93).
- [181] Chin-Chang Ho, Karl F MacDorman, and Z A D Dwi Pramono. “Human emotion and the uncanny valley: a GLM, MDS, and Isomap analysis of robot video ratings”. In: *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*. HRI ’08. Amsterdam, The Netherlands: Association for Computing Machinery, #mar# 2008, pages 169–176 (cited on page 93).
- [182] M Strait and M Scheutz. “Measuring users’ responses to humans, robots, and human-like robots with functional near infrared spectroscopy”. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. #aug# 2014, pages 1128–1133 (cited on page 93).
- [183] Sigurdur O Adalgeirsson and Cynthia Breazeal. “MeBot: a robotic platform for socially embodied presence”. In: *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. HRI ’10. Osaka, Japan: IEEE Press, #mar# 2010, pages 15–22 (cited on page 93).

- [184] Katherine M Tsui et al. “Exploring use cases for telepresence robots”. In: *Proceedings of the 6th international conference on Human-robot interaction*. HRI ’11. Lausanne, Switzerland: Association for Computing Machinery, #mar# 2011, pages 11–18 (cited on pages 93, 100).
- [185] Eric Paulos and John Canny. “PRoP: personal roving presence”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’98. Los Angeles, California, USA: ACM Press/Addison-Wesley Publishing Co., #jan# 1998, pages 296–303 (cited on page 93).
- [186] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. “A Review of Mobile Robotic Telepresence”. en. In: *Advances in Human-Computer Interaction* 2013 (#apr# 2013), page 3 (cited on pages 93, 94).
- [187] M Desai et al. “Essential features of telepresence robots”. In: *2011 IEEE Conference on Technologies for Practical Robot Applications*. #apr# 2011, pages 15–20 (cited on page 93).
- [188] Katherine M Tsui, Munjal Desai, and Holly A Yanco. “Towards measuring the quality of interaction: communication through telepresence robots”. In: *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*. PerMIS ’12. College Park, Maryland: Association for Computing Machinery, #mar# 2012, pages 101–108 (cited on page 93).
- [189] Annica Kristoffersson et al. “Sense of Presence in a Robotic Telepresence Domain: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part II”. In: *Universal Access in Human-Computer Interaction. Users Diversity*. Edited by Constantine Stephanidis. Volume 6766. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pages 479–487 (cited on page 93).
- [190] Daisuke Sakamoto et al. “Android as a telecommunication medium with a human-like presence”. In: *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. HRI ’07. Arlington, Virginia, USA: Association for Computing Machinery, #mar# 2007, pages 193–200 (cited on pages 93, 100).
- [191] P Lincoln et al. “Animatronic Shader Lamps Avatars”. In: *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. #oct# 2009, pages 27–33 (cited on page 93).
- [192] Oyewole Oyekoya, William Steptoe, and Anthony Steed. “SphereAvatar: a situated display to represent a remote collaborator”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’12. Austin, Texas, USA: Association for Computing Machinery, #may# 2012, pages 2551–2560 (cited on pages 94, 99, 100, 104).
- [193] Norman P Jouppi et al. “Bireality: mutually-immersive telepresence”. In: *Proceedings of the 12th annual ACM international conference on Multimedia*. 2004, pages 860–867 (cited on page 94).
- [194] Kazuko Itoh et al. “Development of face robot to express the individual face by optimizing the facial features”. In: *Humanoid Robots, International Conference on*. 2005, pages 412–417 (cited on page 94).

- [195] Ramesh Raskar et al. “The office of the future: A unified approach to image-based modeling and spatially immersive displays”. In: *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. 1998, pages 179–188 (cited on page 94).
- [196] Diego Rivera-Gutierrez et al. “Shader Lamps Virtual Patients: the physical manifestation of virtual patients”. en. In: *Stud. Health Technol. Inform.* 173 (2012), pages 372–378 (cited on page 94).
- [197] Peter Dalsgaard and Kim Halskov. “3d projection on physical objects: design insights from five real life cases”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: Association for Computing Machinery, #may# 2011, pages 1041–1050 (cited on page 94).
- [198] L O K Benjamin. “Shader lamps virtual patients: the physical manifestation of virtual patients”. In: *Medicine Meets Virtual Reality 19: NextMed* 173 (2012), page 372 (cited on page 94).
- [199] P-A Blanche et al. “Holographic three-dimensional telepresence using large-area photorefractive polymer”. In: volume 468. 2010, pages 80–83 (cited on page 94).
- [200] Tomohiro Yendo et al. “The Seelinder: Cylindrical 3D display viewable from 360 degrees”. In: *J. Vis. Commun. Image Represent.* 21.5 (#jul# 2010), pages 586–594 (cited on page 95).
- [201] Nicole Atzpadin, Peter Kauff, and Oliver Schreer. “Stereo Analysis By Hybrid Recursive Matching For Real-time Immersive Video Conferencing”. In: volume 14. IEEE, 2004, pages 321–334 (cited on page 95).
- [202] Hui Zhang, Kwan-Yee K Wong, and Guoqiang Zhang. “Camera calibration from images of spheres”. en. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 29.3 (#mar# 2007), pages 499–503 (cited on page 95).
- [203] Maximilian Ott, John P Lewis, and Ingemar Cox. “Teleconferencing eye contact using a virtual camera”. In: *INTERACT ’93 and CHI ’93 Conference Companion on Human Factors in Computing Systems*. CHI ’93. Amsterdam, The Netherlands: Association for Computing Machinery, #apr# 1993, pages 109–110 (cited on page 95).
- [204] A Al-Saidi et al. “Distributed Collaborative Visualization Using Light Field Rendering”. In: *2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid*. #may# 2009, pages 609–614 (cited on page 95).
- [205] Daniel Knoblauch and Falko Kuester. “VirtualizeMe: interactive model reconstruction from stereo video streams”. In: *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*. VRST ’08. Bordeaux, France: Association for Computing Machinery, #oct# 2008, pages 193–196 (cited on pages 95, 97).
- [206] Eddie Cooke, Peter Kauff, and Oliver Schreer. “Image-based rendering for teleconference systems”. In: (2002) (cited on pages 95, 97).
- [207] Amela Sadagic et al. “Tele-immersion portal: Towards an ultimate synthesis of computer graphics and computer vision systems”. In: *4th Annual International Workshop on Presence*. 2001, pages 21–23 (cited on pages 95, 97).

- [208] Rufael Mekuria et al. “A 3D tele-immersion system based on live captured mesh geometry”. In: *Proceedings of the 4th ACM Multimedia Systems Conference*. MMSys ’13. Oslo, Norway: Association for Computing Machinery, #feb# 2013, pages 24–35 (cited on pages 95, 97).
- [209] J Allard et al. “The GrImage Platform: A Mixed Reality Environment for Interactions”. In: *Fourth IEEE International Conference on Computer Vision Systems (ICVS’06)*. #jan# 2006, pages 46–46 (cited on pages 96, 97).
- [210] Wojciech Matusik et al. “Image-based visual hulls”. In: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. SIGGRAPH ’00. USA: ACM Press/Addison-Wesley Publishing Co., #jul# 2000, pages 369–374 (cited on pages 96, 97).
- [211] Jonathan Starck and Adrian Hilton. “Model-based human shape reconstruction from multiple views”. In: *Comput. Vis. Image Underst.* 111.2 (#aug# 2008), pages 179–194 (cited on pages 96, 97).
- [212] J-S Franco and E Boyer. “Exact polyhedral visual hulls”. In: *Proceedings of the British Machine Vision Conference 2003*. Volume 1. Norwich: British Machine Vision Association, 2003, pages 32.1–32.10 (cited on pages 96, 97).
- [213] Bruce G Baumgart. “A polyhedron representation for computer vision”. In: *Proceedings of the May 19-22, 1975, national computer conference and exposition*. AFIPS ’75. Anaheim, California: Association for Computing Machinery, #may# 1975, pages 589–596 (cited on pages 96, 97).
- [214] O Grau et al. “A Robust Free-Viewpoint Video System for Sport Scenes”. In: *2007 3DTV Conference*. #may# 2007, pages 1–4 (cited on pages 96, 97, 99).
- [215] W Waizenegger et al. “Patch-sweeping with robust prior for high precision depth estimation in real-time systems”. In: *2011 18th IEEE International Conference on Image Processing*. #sep# 2011, pages 881–884 (cited on pages 96, 97).
- [216] Jean-Sébastien Franco and Edmond Boyer. “Efficient polyhedral modeling from silhouettes”. en. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 31.3 (#mar# 2009), pages 414–427 (cited on pages 96, 97).
- [217] I Feldmann et al. “Multi-view depth estimation based on visual-hull enhanced Hybrid Recursive Matching for 3D video conference systems”. In: *2009 16th IEEE International Conference on Image Processing (ICIP)*. #nov# 2009, pages 745–748 (cited on pages 96, 97).
- [218] Eddie Cooke et al. “Extension of incomplete 3d for arbitrary multi-view-synthesis”. In: *VMV*. 2000, pages 205–212 (cited on pages 96, 97).
- [219] J R Starck. “Human modelling from multiple views”. PhD thesis. Citeseer, 2003 (cited on pages 96, 97).
- [220] A Maimone and H Fuchs. “Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras”. In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. #oct# 2011, pages 137–146 (cited on pages 96, 97).

- [221] Jakob Zillner et al. “3D-board: a whole-body remote collaborative whiteboard”. In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. UIST ’14. Honolulu, Hawaii, USA: Association for Computing Machinery, #oct# 2014, pages 471–479 (cited on pages 96, 97).
- [222] Mingsong Dou et al. “Fusion4D: real-time performance capture of challenging scenes”. #jul# 2016 (cited on pages 96, 98).
- [223] Markus Gross et al. “blue-c: a spatially immersive display and 3D video portal for telepresence”. In: *ACM SIGGRAPH 2003 Papers*. Volume 22. SIGGRAPH ’03. San Diego, California: Association for Computing Machinery, #jul# 2003, pages 819–827 (cited on pages 96, 98, 99).
- [224] Patrick Klie, Torsten Büschenthal, and Jörn Ostermann. “Vipid-virtual 3d person models for intuitive dialog systems”. In: *IEEE Workshop on Content Generation and Coding for 3D-television*. Volume 1. 2006 (cited on pages 97, 98).
- [225] A Maimone et al. “General-purpose telepresence with head-worn optical see-through displays and projector-based lighting”. In: *2013 IEEE Virtual Reality (VR)*. #mar# 2013, pages 23–26 (cited on pages 97, 98).
- [226] Alois Ferscha and James Johnson. “Distributed interaction in virtual spaces”. In: *Distributed Interactive Simulation and Real-Time Applications, 1999. Proceedings. 3rd IEEE International Workshop on*. 1999, pages 5–13 (cited on page 98).
- [227] Thomas Erickson et al. “Synchronous interaction among hundreds: an evaluation of a conference in an avatar-based virtual environment”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: Association for Computing Machinery, #may# 2011, pages 503–512 (cited on page 98).
- [228] Gernot Goebbel, Vali Lalioti, and Martin Göbel. “Design and evaluation of team work in distributed collaborative virtual environments”. In: *Proceedings of the ACM symposium on Virtual reality software and technology*. VRST ’03. Osaka, Japan: Association for Computing Machinery, #oct# 2003, pages 231–238 (cited on page 99).
- [229] D Roberts et al. “Reducing Fragmentation in Telecollaboration by Using IPT Interfaces”. In: *IPT/EGVE*. 2005, pages 211–216 (cited on page 99).
- [230] David J Roberts, Robin Wolff, and Oliver Otto. “Supporting a Closely Coupled Task between a Distributed Team: Using Immersive Virtual Reality Technology”. en. In: *Comput. Inform.* 24.1 (#mar# 2012), pages 7–29 (cited on page 99).
- [231] Jeremy N Bailenson et al. “The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction”. In: *Presence: Teleoperators and Virtual Environments* 15.4 (#aug# 2006), pages 359–372 (cited on page 99).
- [232] Oliver Schreer et al. “Real-Time Avatar Animation Steered by Live Body Motion”. In: *Image Analysis and Processing – ICIAP 2005*. Edited by Fabio Roli and Sergio Vitulano. Volume 3617. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pages 147–154 (cited on page 99).

- [233] Maia Garau et al. “The impact of eye gaze on communication using humanoid avatars”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’01. Seattle, Washington, USA: Association for Computing Machinery, #mar# 2001, pages 309–316 (cited on page 99).
- [234] Maia Garau et al. “The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, pages 529–536 (cited on page 99).
- [235] David Roberts et al. “Constructing a Gazebo: Supporting Teamwork in a Tightly Coupled, Distributed Task in Virtual Reality”. In: *Presence: Teleoperators and Virtual Environments* 12.6 (#dec# 2003), pages 644–657 (cited on page 99).
- [236] Norman Murray et al. “An assessment of eye-gaze potential within immersive virtual environments”. #dec# 2007 (cited on page 99).
- [237] William Steptoe et al. “Eye-tracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments”. In: *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. CSCW ’08. San Diego, CA, USA: Association for Computing Machinery, #nov# 2008, pages 197–200 (cited on page 99).
- [238] O Divorra et al. “Towards 3D-aware telepresence: Working on technologies behind the scene”. In: 2010 (cited on page 99).
- [239] Stephan Beck et al. “Immersive group-to-group telepresence”. In: volume 19. IEEE, 2013, pages 616–625 (cited on page 99).
- [240] Ilona Heldal et al. “Successes and Failures in Co-Present Situations”. In: *Presence: Teleoperators and Virtual Environments* 14.5 (#oct# 2005), pages 563–579 (cited on page 99).
- [241] Kibum Kim et al. “TeleHuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’12. Austin, Texas, USA: Association for Computing Machinery, #may# 2012, pages 2531–2540 (cited on page 100).
- [242] David M Krum, Evan A Suma, and Mark Bolas. “Augmented reality using personal projection and retroreflection”. In: *Pers. Ubiquit. Comput.* 16.1 (2012), pages 17–26 (cited on pages 100, 104).
- [243] Susumu Tachi. “Telexistence and retro-reflective projection technology (RPT)”. In: *Proceedings of the 5th Virtual Reality International Conference (VRIC2003) pp. Volume 69.* 2003, pages 1–69 (cited on pages 100, 104).
- [244] Susumu Tachi et al. “MUTUAL TELEXISTENCE SYSTEM USING RETRO-REFLECTIVE PROJECTION TECHNOLOGY”. In: *Int. J. Humanoid Rob.* 01.01 (#mar# 2004), pages 45–64 (cited on pages 100, 104).
- [245] Hong Hua, Leonard D Brown, and Chunyu Gao. “Scape: supporting stereoscopic collaboration in augmented and projective environments”. en. In: *IEEE Comput. Graph. Appl.* 24.1 (#jan# 2004), pages 66–75 (cited on page 100).

- [246] Rudolph P Darken, Joseph A Sullivan, and Mark Lennerton. “A chromakey augmented virtual environment for deployable training”. In: (2003) (cited on page 100).
- [247] Paul Sermon. “Telematic Dreaming”. In: *Leonardo* 33.2 (#apr# 2000), pages 90–90 (cited on pages 101, 104).
- [248] Brett Jones et al. “RoomAlive: magical experiences enabled by scalable, adaptive projector-camera units”. In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. UIST ’14. Honolulu, Hawaii, USA: Association for Computing Machinery, #oct# 2014, pages 637–644 (cited on page 101).
- [249] N H Lehment, D Merget, and G Rigoll. “Creating automatically aligned consensus realities for AR videoconferencing”. In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. #sep# 2014, pages 201–206 (cited on page 101).
- [250] Daniel Avrahami, James Fogarty, and Scott E Hudson. “Biases in human estimation of interruptibility: effects and implications for practice”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’07. San Jose, California, USA: Association for Computing Machinery, #apr# 2007, pages 50–60 (cited on page 101).
- [251] Darren Gergle, Robert E Kraut, and Susan R Fussell. “The impact of delayed visual feedback on collaborative performance”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’06. Montréal, Québec, Canada: Association for Computing Machinery, #apr# 2006, pages 1303–1312 (cited on page 101).
- [252] Chris Greenhalgh et al. “Making Networked Virtual Environments Work”. In: *Presence: Teleoperators and Virtual Environments* 10.2 (#apr# 2001), pages 142–159 (cited on page 101).
- [253] Petr Slovák, Pavel Troubil, and Petr Holub. “GColl: A Flexible Videoconferencing Environment for Group-to-Group Interaction”. In: *Human-Computer Interaction – INTERACT 2009*. Edited by Tom Gross et al. Volume 5727. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pages 165–168 (cited on page 102).
- [254] Tejinder K Judge and Carman Neustaedter. “Sharing conversation and sharing life: video conferencing in the home”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’10. Atlanta, Georgia, USA: Association for Computing Machinery, #apr# 2010, pages 655–658 (cited on page 102).
- [255] Danielle Lottridge, Nicolas Masson, and Wendy Mackay. “Sharing empty moments: design for remote couples”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’09. Boston, MA, USA: Association for Computing Machinery, #apr# 2009, pages 2329–2338 (cited on page 102).
- [256] John C Tang et al. “HomeProxy: exploring a physical proxy for video communication in the home”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’13. Paris, France: Association for Computing Machinery, #apr# 2013, pages 1339–1342 (cited on page 102).

- [257] William W Gaver. “The affordances of media spaces for collaboration”. In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. CSCW ’92. Toronto, Ontario, Canada: Association for Computing Machinery, #dec# 1992, pages 17–24 (cited on page 102).
- [258] Robert S Fish et al. “Video as a technology for informal communication”. #jan# 1993 (cited on page 102).
- [259] Alan R Dennis et al. “Ad hoc versus established groups in an electronic meeting system environment”. In: *System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on*. Volume 3. 1990, pages 23–29 (cited on page 102).
- [260] Anne-Laure Fayard and John Weeks. “Photocopiers and Water-coolers: The Affordances of Informal Interaction”. In: *Organization Studies* 28.5 (#may# 2007), pages 605–634 (cited on page 103).
- [261] Ralph Schroeder et al. “Collaborating in networked immersive spaces: as good as being there together?” In: *Comput. Graph.* 25.5 (#oct# 2001), pages 781–788 (cited on page 103).
- [262] Joanna McGrenere and Wayne Ho. “Affordances: Clarifying and evolving a concept”. In: *Graphics interface*. Volume 2000. 2000, pages 179–186 (cited on page 104).