Personal Data Sovereignty: A Sustainable Interface Layer For A Human Centered Data Ecosystem

Michael Lockwood
The University of Salford
School of Computer Science and Engineering

First Supervisor: Professor Nigel Linge
Second Supervisor: Dr Wilfred Darlington

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

2020

Table of Contents

List of Tables

List of Figures

.

Acknowledgments

I would like to thank my supervisors Professor Nigel Linge and Dr Wilfred Darlington for all their support and encouragement. I have very much enjoyed our meetings throughout this process and have benefited greatly from their knowledge and expertise. I would like to thank all those who have participated in both the expert interviews and focus groups, their insight and domain expertise has been invaluable.

Most of all I would like to thank my family, my beloved Elizabeth and my darling girls, Alexis, Lucia and Elise. I am grateful for their patience and understanding, and I promise never to lock myself in a room again and miss another moment of your growing up, every weekend is yours from now on.

Declaration

I, Michael Lockwood, declare that the work in this thesis was carried out in accordance to the regulations of University of Salford, and is an original piece of research, except where indicated by reference within the text.

Signed: .....

Date: .....

Dedication

To all those who understand that Dyslexia Rules KO !!!

Glossary

An understanding of the following key terms is important to the reader:

Centralised Internet
The centralised Internet refers to an architecture in which the user connects to a centralised server which acts as an agent for all communications. The server records and stores both communications and account information. This type of architecture is argued to be problematic as it empowers centralised entities to control loaned or federated identity, bind participants into contracts of adhesion, surveil Internet activity, and gather and repurpose personal data and information.

Decentralised Internet
A Decentralised Internet in the context of this research is an architecture that rejects the current model of centralised control and asymmetric relationships, and instead rebalances the landscape through peer to peer relationship and participant control. Individuals are afforded agency over their identity and personal data through tools that allow private, secure, censorship-resistant access to information and web services. It can also be interpreted to mean any technology that allows degrees of transparency, partial agency or disruptive capability over the centralised model.

Human-Centred Data Ecosystem
This term describes the full ecosystem required to place the human at the centre of personal data streams, affording agency, transparency, and control. The term represents the required concepts, technology, and digital tools. It may also refer to external physical hardware, independent systems, organisational bodies and legislative mechanisms. The term represents all the component parts of the greater ecosystem.

Human Data Interaction
This term describes an academic field that investigates theory relating to the way humans interact with personal data across the network. The field recognises the imbalances in the current centralised model and states the aim of *'placing the human at the Centre of the flows of data, providing mechanisms for citizens to interact with these systems and data explicitly*' (Mortier, 2015, p. 1). Human Data Interaction is the academic field through which this research will be labelled and disseminated.

Self-Sovereign Identity

The term describes a concept that endeavours to afford participants the same agency in the digital realm as individuals enjoy in the real world. SSI does this by enabling the individual considerable control over both their digital identity and personal data. SSI is not a single technology, rather a wider ecosystem comprising of a number of related components that collectively operate free from any centralised control, ultimately sovereign to the participant. At its core is an identity layer, a means by which an individual can control and manage an online presence, through which relationship, transaction, and personal data can be managed and controlled.

Sovereign Boundary Mechanism

This term represents a set of methods, tools and user representations that allow for an interaction with a Decentralised Internet, through a Human-Centred Data Ecosystem that enables the principles of Self-Sovereign Identity. The concept represents a standardised interaction, which is best described as a metaphorical ring of steel between the participant and the network, were the participant controls identity, relationships, data streams, and access control.

Participant

In the context of this research and indeed the domain, the word *participant* is used to substitute the word *user*. The rational for this is that the individual engaging with the network is sovereign and is not *using* any owned or controlled system, instead they are acting with agency and participating on their own terms.

Abstract

The reality of ubiquitous computing and exponential personal data generation challenges the notion of privacy, as Surveillance Capitalism and Nation State endeavour to record and analyse personal data with the objective of leveraging influence and control. It is argued that this centralised model threatens to stifle the digital economy, destabilise our democracy, and fundamentally change our social norms. Real-time, non-statistical datasets offer huge potential for governance, commerce, and social cohesion. But the positive benefit of the emerging data driven society is threatened by the tensions formed through asymmetric power imbalances that manifest across a narrow band of walled gardened web services.

In recent years work has been undertaken to counter the centralised model, despite these efforts there has been limited change in trajectory or sustained adoption of decentralised technologies. This research is designed to explore and evaluate the Decentralised Internet. Investigating the challenge of designing usable, sustainable tools for the everyday participant. This research engages mixed methods to explore the trajectory of technologies and public attitudes. Domain experts are consulted to explore application and value proposition. Practice extends the decentralised trajectory to consider participant journeys, interaction, and the interface layer.

This research concludes that the core technological infrastructure now exists to facilitate a genuine Decentralised Internet and that an identity layer facilitated through Blockchain technology is progressing the domain towards Self Sovereign Identity (SSI). This research extends this trajectory through Conceptual Modelling to define a Sovereign Boundary Mechanism (SBM), an independent realm of interaction which enables the principles of decentralisation. Analysis suggests that this interaction is high in friction, requiring considerable internalised cognition and prior knowledge in order to engage.

This research concludes that the concept of network privacy is poorly defined and miss-understood, and that participants struggle to see its value across context and cultures. Investigation indicates that the Decentralised Internet cannot be marketed, and instead has to supersede the centralised model through defined innovations. This research argues that a cohesive strategy is required to achieve adoption, one which collectively identifies and develops offerings of value through design thinking while defining a consistent narrative to deliver targeted solutions within cultural contexts.

This research makes a theoretical contribution to knowledge by connecting the domains of Self Sovereign Identity (SSI) and Human Data Interaction (HDI). The research establishes the fundamental spheres of interaction for an analogue SSI system through what is defined as a Sovereign Boundary Mechanism (SBM). The research identifies issues and paradox's relating to an SBM and identifies further required investigation and research. This research makes a practical contribution to knowledge by presenting a framework and resource for further innovation and development, the wider problem space for a Human-Centred Data Ecosystem is defined, and finally the research contributes to a wider adoption strategy through the identification of value proposition.

Chapter 1: Introduction

The following chapter considers the motivational factors driving this academic work, outlining the profile and background of the researcher. The problem space is defined through a historical narrative, identifying the core issues and areas of investigation. The following sections communicate the focus of the enquiry, the research aims, objectives and questions. The scope of the exploration is defined, followed by a mapping of the research components. The final section presents the thesis structure and chapter content.

1.1 Motivation

After graduating from University in 2001, with a strong passion for technology and a degree in Product Design, I was offered a position at *BT Exact Technologies, Radical Multimedia Lab*, Martlesham UK. The department focused on the research and development of emergent technology, and the generation of intellectual property. The department was culturally progressive, championed the application of design thinking, and problem solving to every facet of its operation. My everyday responsibilities centred around creative thinking, concept generation, and moving ideas from the whiteboard through to tangible prototypes. The department's core areas of investigation included virtual worlds, gaming, web and interactive, non-linear media, and location-based experiences. Towards the end of my tenure, *Blast Theory* and Nottingham Trent University's project, *Can You See Me Now* [1], was developed in my office (CYSMN, 2003). My lab's own investigation, and experiments with Location Based Media, and Blast Theories development of augmented experiences, ignited a personal interest in the field. Having left British Telecom to join the University of Salford in 2004, my interest in location driven interaction, and subsequently the User-Centred-Data-Ecosystem, became more prominent after reading *The Headmap Manifesto* [2] (Russell, 1999). This literature describes a

1: *Can You See Me Now* was one of the first location-based media games, in which online players pursued real individuals in the real world. https://www.blasttheory.co.uk/projects/can-you-see-me-now/

2: Ben Russell's *Headmap Manifesto* published in 1999 is an influential example of the discourse around location-aware technologies, exploring the positive and negative aspects of their capabilities, before cloud computing, the Internet of Things, and Web 2.0. http://www.technoccult.net/wp-content/uploads/library/headmap-manifesto.pdf

location-based data system, in which the individual is defined as a node, a central repository for personal data, media and information. Within this model, peer-to-peer data discovery, and exchange is enabled through geographic location and physical encounter. These ideas offered an alternative lens, through which to view the issue of data ownership, relative to location tracking, and location marking as a derived work.

Late in 2004, I became involved with the *Future Sonic Festival* in Manchester UK (Future Everything, 2019) working with academic Drew Hemmit, and together with my colleague Rodrick Martin, we developed a mapping application utilising for the first time globally, what was to become *Open Street Map*[3] (OSM, 2019). Raw trace data was compiled, and through a protracted pipeline, cartographised into a functional digital map. This initial development of a front and backend information mapping system, acted as the catalyst for multiple location-based applications, developed over a number of years (Lockwood, Martin, 2010), (Lockwood, Lin & Martin, 2011).

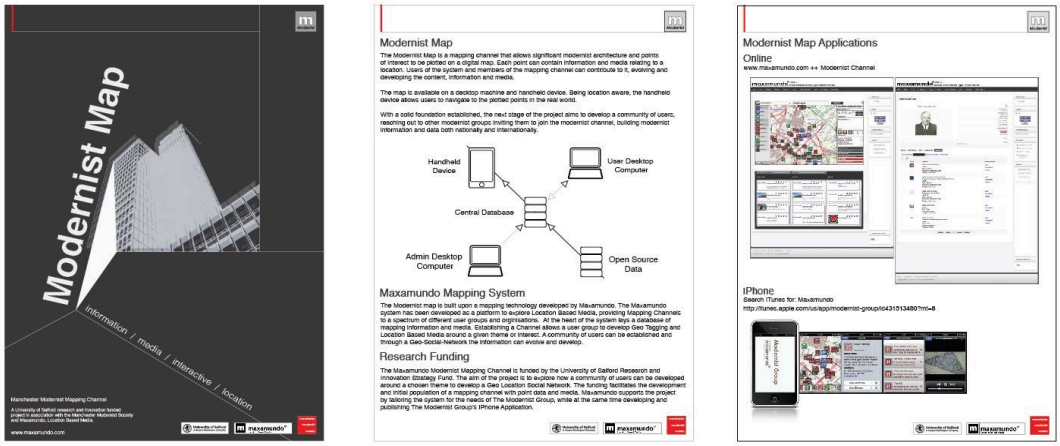


Figure 1: Maxamundo Project Example: Modernist Map Application 2007

This early work provided valuable empirical experience, through which I encountered a myriad of issues relating to usability and adoption. During this period, a start-up company was established, and hundreds of hours where spent designing, building and realising technology that ultimately failed to achieve significant traction and meaningful adoption

3: *Open Street Map* is a non-proprietary, publicly available global mapping data resource, developed through a participatory model of amateur cartographers. https://www.openstreetmap.org

I learnt that new ideas can be difficult to communicate, that new interaction paradigms can lack a suitable mental model and metaphor, that participants require friction free interactions, and that relatively simple tasks can become complex user experiences. Through focus group and user feedback, it became clear that sustained adoption requires clear value proposition, and that this value needs to be balanced against the cognitive load required to engage.

Following my Location Based Media work, I became increasingly interested in the grass roots development of concepts and technologies to decentralise the Internet. Although this resistance to centralisation lacks the dominant mechanic of physical location found in my previous work, it does embrace the concept of decentralised personal data ownership and control, and it is this underlying parallel that bridges my previous work to this current area of research.

During my early investigation into decentralised concepts, technologies and prototypes, I observed many of the issues I had encountered during my previous practice. There seemed to be an imbalance within the decentralised community towards pure computer science, with a notable absence of discourse concerning the design challenges posed by the development and deployment of novel and unfamiliar technologies and user interactions. Many of the prototypes observed replicated the services provided by centralised operators, offering little innovation or obvious value. It was at this point I became focused on investigating how new decentralised concepts and technologies could be transposed into usable sustainable tools that can be understood and adopted by a broad user base. Hence, this experience forms the key motivation for carrying out the research that is detailed in this thesis.

It is important at this point to state that my personal motivation does not come from an activist position, and I do not consider myself a privacy advocate. If asked to express a position on the centralised Internet, I would argue that the current situation is not conducive to a free and open society, and that it has the potential to limit and stifle our digital economy. I would

suggest it impinges on our right to be left alone, and potentially inhibits the many benefits we may derive from real-time *Big Data*. I am aware of what I consider to be a moderate bias, and I have made every conscious effort to prevent it from influencing the design of this research, the methods of data gathering, and the analysis and interpretation of results.

The personal motivation that drives this research, comes from the challenge of understanding the decentralised problem space. My motivation is grounded in the design thinking required to deploy complex, multi-dimensional technology. The guiding word throughout this research, is '*can'*, rather than '*should'*.

My background at *BT Exact Technologies*[4], experience of personal project development, and observation of the work of others, has defined the methodology within this study. I understand from first-hand experience the reality of technology adoption, the importance of value proposition, user experience and usability. I understand that the last step within any software development process is the creation of digital artefact. There is a great deal of design investigation and iterative development, which needs to be engaged before that stage. If the goal of this study is to develop a framework of knowledge to assist in the development of decentralised tools, I see the complete PhD as the preliminary stage of such a design process. It is the prospect of a design process informed by in-depth research, together with a passion for the bridge between human beings and data, mediated by technology that drives my motivation.

1.2 Statement of the Problem

The original architecture of the Internet was designed as a peer-to-peer system, a web of individual computers communicating and exchanging packets of data directly with one another, without the need for hierarchical or structural control. With the advent of early server software, centralised hosting of the growing Internet emerged. This led to a client server model, the client on the edge of the network representing the human, and the server at the centre representing the organisation, holding and subsequently controlling the data. These models of data centralisation led to the growing network being hosted by a relatively small number of service providers. As humans engaged with these centralised systems, activity logs were generated. Initially designed for managing and debugging systems, these logs soon became vast databases representing personal engagement, and activity on the network. Eben Moglen (2010) argues that this evolution in computing *'was a recipe for disaster'*. This trend towards centralisation continued with the advent of the Web 2.0 (Reilly, 2009), the second phase of the Internet, which is characterised by a shift away from static web pages towards dynamic interactions and user-generated content. Web 2.0 is established and sustained by a *Participatory Culture* (Jenkins, 2009), and *Produsage* (Burns, 2006), where the individual participates in a two-way data exchange, and is both the consumer and producer of media, information, and raw data. From the first Web 2.0 web pages, facilitating fluid feedback and interactions, the trend progressed towards the social web, with wiki's, blogs, instant messaging, social media and social networks. All of these centralised participatory web services operate under *Contracts of Adhesion*[5], also known as *Boilerplate* and *Standard Form Contracts*. In order to access services, users are required to accept terms and conditions, through arguably meaningless consent. This type of agreement allows the service provider to claim ownership of all data and metadata uploaded and generated, by the user as they traverse through a service.

5: An *Adhesion Contract* represents an agreement were one party has substantially more power over the other and is accepted as part of a transaction with no opportunity for negotiation.

The term meaningless consent relates in this context to a step that for many may become mundane and subconscious, at first individuals may pause for thought and consider the consequences of granting consent to third parties, but over time in line with the theories of *Immediate Gratification Bias* (Acquisti, 2004) and the *Privacy Paradox* (Norberg, 2007), the need to gain rapid access to online services out weights the friction posed to protect one's personal information. Recent research argues that applications of the required consent under GDPR, is poorly implemented and prone to patterns of user coercion, manipulation and nudging (Utz, 2019).

The humble cookie, a technology designed to make the user experience better, by retaining preferences and populating simple personalised data within webpages, became a technology for tracking individuals across the Internet. Search providers began to harvest the data generated by the search terms which individuals entered. These organisations morphed from companies primarily focused on search, to businesses that focused on data collection, analytics, forecasting and targeted marketing. Popular web destinations have developed second income streams, as they sell the personal data of their customers, through embedded third-party technologies. Mobile smart technologies have added to this inexorable growth in personal data capture, as users now carry constantly connected network devices, continually generating location, application and communications data. Data Brokers[6], who for many years gathered personal information from public records and financial exchanges, now deal in personal digital data, combining data footprints from numerous sources, to build dossiers and additional inferred connections. As Anthes (2015) argues *'You would be surprised how much they know about you, and what they are doing with your information'* (p. 28). It would seem that the Internet has now become a battle ground for a small number of service providers, whose primary objective is to gather as much personal data and subsequent market intelligence as possible. As Max Van Kleek explains

6 Data Brokers represent companies that collect, analysis and redistribute personal data to deliver market intelligence the most dominate of which is Acxiom. https://www.acxiom.co.uk/

> *Currently a handful of dominant platform vendors and application service providers, are grappling for control over individuals' personal information, trying to accumulate as many users as possible, in order to maximize understanding of every nook and corner of social interaction* (Van Kleek, 2014, p. 125).

The development of data science and analytical techniques have evolved to allow inferences to be drawn, generating intelligence to reveal sexual orientation, political affiliation, viewing habits, buying habits, social groups, insurance risk and much more. Dr Michal Kosinski, a prominent researcher in the field and developer of inference techniques, sums it up with the statement, *'this is not my fault. I did not build the bomb. I only showed that it exists'* (Grassegger, 2017). The combination of metadata through graph analysis, can reveal more about an individual, than they know about themselves. Dr Kosinski's work, has indirectly found its way into our collective consciousness, through its application by the company formally known as Cambridge Analytica. Eben Moglin (2013) argues that, *'we are on the verge of eliminating forever, the fundamental right to be alone in our thoughts'* (p. 20).

It would appear that the developing landscape, is far from the Utopian ideal defined within *'A Declaration of the Independence of Cyberspace'* authored by John Perry Barlow, founder of the Electronic Frontier Foundation[7]. The declaration called for a free and open resource of shared knowledge and personal communication, (Barlow, 1996). More recently, significant figures have again called for action, to reverse the trend of personal data collection and exploitation. Tim Berners-Lee, the credited inventor of the World Wide Web, has proposed an *Internet Bill of Rights*, *A Magna-Carta for the Web*, in which individual privacy is recognised and protected (Berners-Lee, 2014). The state has capitalised on this evolution in personal data generation and collection. Landau (2013) explains how disclosures by Edward Snowden[8] have revealed the magnitude of the issue, as security agencies record and retain the

7: The Electronic Frontier Foundation is a non-profit organisation with a mission to defend digital privacy, free speech and innovation. https://www.eff.org/

8: Whistle Blower Edward Snowden revealed US NSA Data Collection activity. Story by Greenwald, Poitras and MacAskill in the Guardian in June 2013. https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-

vast majority of our digital communications transfer. In the United Kingdom the position of the state has become more transparent, with the passing into law of the *Investigatory Powers Act 2016*, named by the media as the *Snoopers Charter* (UK-Gov, 2016). This legislation requires service providers to retain data and grant the state, and its various agencies, unparalleled access to personal information and data. This developing paradigm has huge implications for individual privacy. It remains to be seen what effects the centralisation, exchange and exploitation of personal data will have on our culture, politics, economy and social norms. Richard Mason (1986) in his work entitled *Four Ethical Issues of the Information Age*, asked, *'the question before us now is whether the kind of society being created, is the one we want?'* (p. 5). It would seem for many, that the answer to that question is 'No'.

There are compelling arguments that suggest that the current trajectory is unsustainable, and that the monopoly of the Big Data oligarchs, threatens the potential of our data driven society and digital economies. These arguments include the stifling of the digital economy, the missed opportunity of real-time big data analysis, the negative impact on our democracy, the erosions of social norms, and a complete breakdown of personal privacy. Some of these arguments are very powerful and are compounded further by the advent of the *Internet of Things* and *Artificial Intelligence*.

In recent years, we have witnessed a growing movement to push back against the centralised trend, and instead, re-decentralise the Internet. The primary objective of this resistance is to give the individual control over personal data and address the power imbalances between the large Internet service providers and the individual. The global organisation *MyData*, now has hundreds of active members, defining its declaration as:

To empower individuals with their personal data, thus helping them and their communities develop knowledge, make informed decisions and interact more consciously and efficiently with each other as well as with organisations (MyData, 2019)

CHAPTER 1: INTRODUCTION

Project VRM at Harvard University maintains a list of hundreds of projects, start-up companies, developing technologies and standards, related to the development of commerce for the Decentralised Internet. (ProjectVRM, 2019). Dedicated Academic fields have been established to support the development of theory concerning decentralisation. Human Data Interaction is the prominent example, with a primary objective to place *'the human at the Centre of the flows of data, providing mechanisms for citizens to interact with these systems and data explicitly: Human-Data Interaction (HDI)'* (Mortier, 2015, p. 5). There are a number of prominent global groups who meet biannually to debate, investigate, design and build technology, combining notions of identity and personal data, in order to turn the tide on the centralised Internet, these include *Rebooting The Web of Trust*[9] (WOT, 2017), and the *Internet Identity Workshop*[10] (IIW, 2019). Despite a clear problem space, a desire for change, considerable collective intelligence, and significant continued efforts, little impact would seem to have been made in the mainstream, to address the asymmetric power imbalances described.

The problem can be surmised, as one of imbalance, the rapid growth of dominate forces made possible by a centralised architecture, and private enterprise being the first to market. This imbalance empowers a small number of Internet service providers, with great intelligence and influence. It is argued, that this power monopolises what was envisaged as a free and open domain, and for many this imbalance needs to be addressed. Despite great efforts, the problem remains. The question is, can this rebalancing be achieved, and what are the mechanisms required? This research explores this problem space, and endeavours to make a contribution towards the decentralised community's objectives.

It is important to clarify the term decentralisation in the context of this research in addition to the definition found within the glossary section of this thesis (see:xii). There has been criticism that the term is used widely among advocates of blockchain technology, and

also within economic and political arenas to describe a vague solution to the perceived harms of centralisation without specificity. It is claimed that many conceptual models of decentralisation lead to economically and structurally centralised outcomes. (Schneider, 2019). This research does not start from the position that all centralisation in the context of the digital realm automatically leads to privacy violation or surveillance capitalism. There are clearly centralised systems that rest on traditional monetisation models, that hold privacy aware values or that are commons-based for example Wikipedia/ Wikidata. The objective of this research is not to explore technology interactions at solely aim to move to a model of complete decentralisation. Rather this research considers proposals for a decentralised architecture that facilities an identity layer for the internet. This prospect affords participants the agency to control personal data streams and disrupts asymmetric relationships. Once established this model or ecosystem potentially allows participants to engage in purely decentralised activities, but also engage in new or emergent fully or semi- centralised models, with a greater degree of agency, transparency and control.

1.3 Research Focus

When considering why we are yet to see significant mainstream adoption of decentralised technologies, despite existing solutions being available (ProjectVRM, 2019) and published data suggesting significant public demand for change (Catapult, 2016), this research begins with a number of assumptions: That low levels of adoption are due to the lack of consideration, for the end user by a mainly technology focused community. That there are issues relating to the complexity of the interaction required to fulfil the principles of decentralisation. That there is a missing, weak or poorly communicated value proposition within decentralised applications. A value proposition is an argument or justification made to a new user of a product or service when they are considering adoption. It conveys the clear advantage of using one

product over another, the benefits or how it might solve an existing problem. The user understands the value in the proposition presented, which in terms raises the probability of adoption. Recent research suggests that current decentralised technologies, incorporating concepts of identity and Blockchain technology, have failed to consider usability. In the paper entitled, *A First Look at Identity Management Schemes on the Blockchain*, the authors argue that none of the reviewed projects currently consider human integration and, in summary Dunphy and Petitcolas (2018) state that *'there is a noticeable lack of contextual understanding relating to the user experience within elements of the schemes we encountered'* (p. 13). Insight into usability issues within this general domain, have been explored before. The seminal paper *Why Johnny Can't Encrypt,* Whitten and Tygar (2005) describe issues of usability concerning email encryption software, and argue that despite a well-designed interface, a lack of fundamental mental model, and understanding, prevents Johnny from successfully encrypting his communications. It could be argued that this work draws striking contemporary resemblance to what is being witnessed today. In terms of adoption, there is a real sense that existing decentralised artefacts, and indeed forerunners of current developments, are failing to consider and address some of the fundamental principles of the *Diffusion of Innovation* (Rogers, 2012)[11]. The understanding of a new product offering the value found within concepts, and the progression of a participant when making a decision to adopt, are critical considerations, alongside technological capability.

At a surface level, these assumptions may seem obvious, however as a designer, there is a requirement to investigate the detail to understand better the problem space, and subsequently attempt to develop a framework to support addressing these issues.

The impetus underpinning this research can then be defined in a broader sense with the following statement:

11: The *Diffusion of Innovation* relates to the work of Rogers (see: 2.4.4) but also in this context the general required conditions for the adoption of decentralised technology.

CHAPTER 1: INTRODUCTION

'In order for genuine decentralised technologies to achieve mainstream adoption, significant design driven solutions need to be found, to balance the cognitive load required for engagement, with the value proposition decentralised tools and services may present'

If this assumption can be explored and evidenced, it may explain why decentralised technologies have yet to find significant traction, and this has merit. But an investigation into the related multidisciplinary problems, and a deeper deconstruction of the issues and corresponding arguments, is considered to be of much greater value. The prospect of a better understanding of the interrelated variables, will make a valuable contribution to knowledge, and potentially increase the probability of the realisation of a Human-Centred Data Ecosystem[12]. As such, the central focus of this study is to investigate the probability, of the decentralised movement achieving the principles desired by its advocates and working groups. This research considers the projects and technologies currently in the public domain, and where accessible, those that are under development. The work investigates the capability of technology to deliver the required functionality, subsequent conceptual models, and the defined interactions. It considers usability, and the barriers and obstacles of adoption, the users' mental model, metaphor and collective understanding. The research surveys the general public, to understand attitudes towards privacy, personal data and the way technology is used. Expert opinion has been sourced to understand the trajectory of the domain space, its limitations, and its value. From a theoretical prospective this work investigates *Privacy, Surveillance Capitalism*, the *Diffusion of Innovation*, and classical and contemporary interpretations of *Human Computer Interaction*. Ultimately this research endeavours, to understand better both the theoretical and practical issues involved in the design and implementation of decentralised tools and services. It acts as a framework, and utility for the design community, by defining the empirical connections between the theory surrounding our data driven society, a Decentralised Internet, and the design and adoption of decentralised technologies and tools.

12: For an explanation of the term Human Centred Data Ecosystem see Glossary xiii.

1.4 Research Aims, Questions and Objectives

Research Aims

Derived from the aforementioned problems, the research aims are defined as:

- To investigate network privacy and the rational for a Decentralised Internet.
- To explore the interface between the human and Decentralised Internet technology.
- To investigate the potential adoption of Decentralised Internet technology and tools.

Research Objectives

This research has the following Objectives:

- To make a contribution to knowledge, and expand the academic field surrounding the Decentralised Internet'
- To establish a body of work to inform the design and development of decentralised technologies.

1.5 Scope of Research

It is important to clarify the scope of this research, given the multidisciplinary nature of the decentralised domain, and the depth of the problem space. As Richard Mortier (2009) explains '*the domain is made up of a number of related fields, Computer Science, Statistics, Sociology, Psychology and Behavioural Economics'* (p. 4). It can be argued that this can be expanded further, to include classical and contemporary *Human Computer Interaction, User Experience Design, User Interface Design*, and *The Diffusion of Innovation*.

It is important to state that this study will not endeavour to design and build a decentralised system or interface layer. Early exploration into existing projects and technologies, would suggest that there is a great deal of preliminary work to be done, in understanding and analysing the relevant variables, before any further attempt is made to design and build product. There is also a great deal to be learned by conducting critical investigation, and practice led extension of proposed, existing, and past decentralised projects and technologies.

The research questions have been designed to explore the possible variables of the decentralised equation and can be thought of as follows. With regards the main research question for this study: ***Can a sustainable technology be established to allow for individual agency within a Decentralised Internet?*** It can be dissected further, first, a clear understanding of a Decentralised Internet and Human-Centred Data Ecosystem needs to be established, together with the principles and required functionality to achieve the prescribed agency. An investigation of existing and emergent technologies and standards needs to be considered against the defined requirements, to ascertain whether or not the defined functionality can be achieved.

The words, sustainable and established, within the main research question, relate to longevity, initial adoption, product value, accessibility, and a critical mass for participation.

The additional questions within this study are designed to address these substantive considerations.

The second question: **How can an interface layer for a Decentralised Internet be designed to allow for accessible interaction?** This explores the notion of interface in a broad context, investigating how an interface layer can be developed to be both accessible to the everyday participant, while providing the required functionality for a decentralised engagement.

The third question: **How might a Decentralised Internet emerge and be adopted?** It is proposed to investigate where and how, decentralised technologies might break through into the mainstream, where they might find traction and adoption, and what value decentralised systems might offer in order to satisfy the requirements for a *Diffusion of Innovation.*

In line with the aims of the study and the defined research questions, three primary areas of investigation have been identified: Privacy, Technology and Adoption. As such, the dynamic intersections between these areas where identified, and can be described as: Decentralised Models, Value Proposition and Usability.

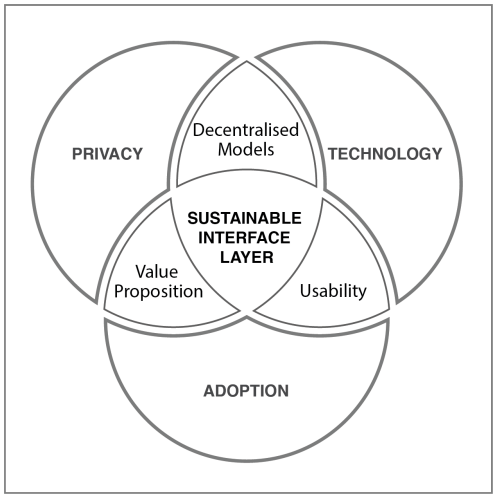


Figure 2: Intersection between the main areas covered within this research

In combination, the knowledge generated from the exploration of these intersecting areas, will contribute to the overarching objective to explore, and understand better a *Sustainable Interface Layer for a Human-Centred Data Ecosystem*.

By investigating the intersection between privacy and technology, a greater knowledge of existing and potential decentralised models can be understood. This means the principles, mechanisms and wider ecosystem. It is worth considering that though this study's central focus is the point at which individuals interact with a decentralised system, the system itself is more than a single application. It is a collection of interdependent variables, which need to be understood better.

By investigating the intersection between privacy and adoption, this study is designed to generate meaningful knowledge in the context of value proposition. The value proposition is a core component of any sustainable technology, and without meaningful value, there is little motivation for a participant to adopt. This study considers value proposition in several contexts, for the individual, for commerce, and for the benefits of wider society.

By investigating the intersection between technology and adoption, this research is designed to generate an understanding of usability issues and challenges. Initial investigation suggests that agency within a Human-Centred Data Ecosystem, represents a complex user experience and usability problem space. By investigating the interactions required in the context of adoption, this research endeavours to understand better the complexities, challenges and potential solutions.

1.6 Research Overview

Figure 3 represents a visual mapping of the components within this research.

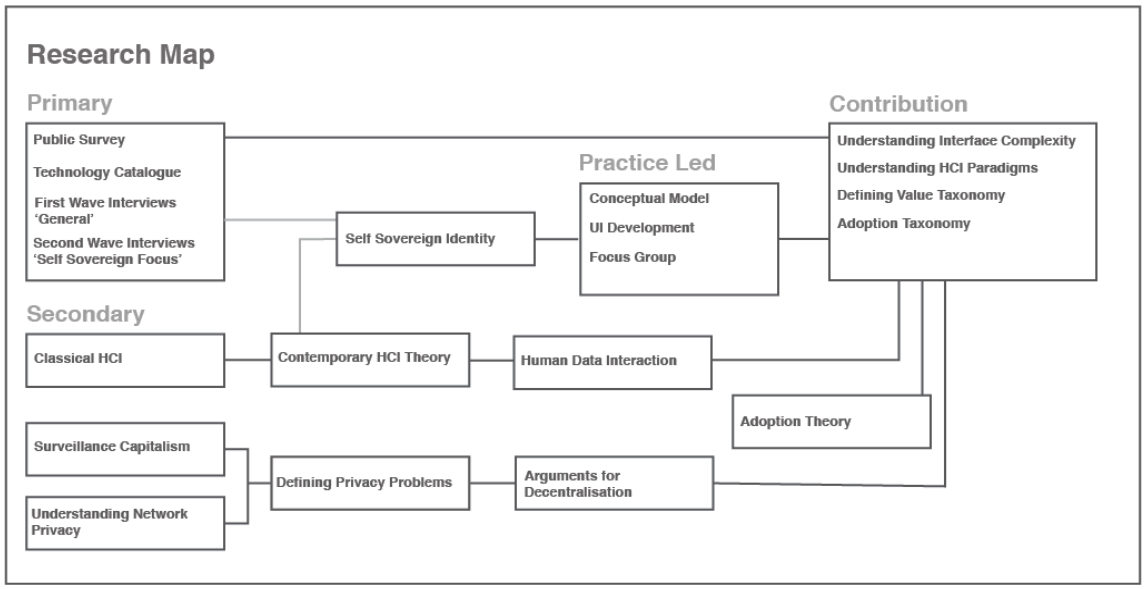


Figure 3: Research Map

The research is conducted with a pragmatic philosophy, using a mixture of methods. The journey began with the identification of relative core academic themes, derived from initial investigation of the literature, relating to the Decentralised Internet. These fields included *Surveillance Capitalism*, *Network Privacy*, and led to an investigation of privacy problems, and supporting arguments for a decentralised ideology. Classical and contemporary HCI theories have been investigated, within the bounds of the domain, progressing to an investigation of *Human Data Interaction* as an emergent theory, together with a practice led dissection, extension, and critical evaluation of Self-Sovereign Identity. Both strands of theory, which in essence encapsulate interaction and value, are then considered in the light of adoption theory, and the *Diffusion of Innovation*.

In line with the research aims, objectives, and research questions, a number of Primary data gathering activities where engaged:

Quantitative data has been gathered through a Public Survey. The survey investigated understanding, awareness and attitudes towards personal data online.

CHAPTER 1: INTRODUCTION

A systematic Artefact Catalogue of existing and emergent decentralised technologies and projects was established. A critical analysis was then engaged in order to understand their technological capability, functionality, usability, value and trajectory.

A Thematic Content method has been used to analyse two phases of interviews with domain experts. The first investigated broad questions regarding the Decentralised Internet. The second focused on the concept of Self-Sovereign Identity, and explored issues around usability, cognition, value proposition and adoption.

A Practice Led Component has been developed to understand better the required elements for an individual interaction with a decentralised system. In this case the practical application of Self-Sovereign Identity. The first phase engaged a Conceptual Modelling method, to understand more fully the objects and functionality required. Once a conceptual model had been established, an initial interpretation of a User Interface (UI) was developed through a number of iterative cycles. Both the conceptual model and preliminary UI have been evaluated through a critical reflection, and collectively through demonstration and focus group. The data generated by the focus group, has been evaluated as a continuation of the Thematic Content Analysis within this research, using existing and emerging codes.

Data and understanding derived from these research activities, has been analysed and discussed in order to draw conclusions relating to interface layer complexity, HCI paradigms, and an improved understanding of value proposition and adoption.

1.7 Thesis Outline

This thesis is organised according to the following structure:

Chapter 2 presents a literature review of the topics considered relevant to this research. The review first investigates the literature relating to *Surveillance Capitalism*, outlining a narrative progression from the foundation of surveillance theoretical frameworks through to the contemporary understanding of human data gathering and utilisation. The review then moves to consider the field of *Network Privacy*, the foundation of privacy from an anthropomorphic perspective, through to the arguments surrounding the digital sphere. The review then considers the field of *Human Computer Interaction*, investigating the core theories and historical paradigms, with the objective of understanding how this informs the research questions and defined problem space. Finally, the review focuses on the evolving domain of *Human Data Interaction*, and its relationship to a Decentralised Internet.

Chapter 3 considers contemporary endeavours to develop technology, to allow an individual to take control of personal data, through a Sovereign identity layer. The concept of Self-Sovereign Identity is dissected in detail, describing its principles, core concepts, capabilities and function. The chapter investigates the underlying technology and developing standards. It lays out the rational for further investigation within the practice-led element of this research.

Chapter 4 presents the underlaying research philosophy, design, methods and means of analysis. This research takes a pragmatic position and uses a mixture of methods. The methods include secondary data gathering through standard literature search and review. Primary data is gathered in a number of ways. Through a statistically analysed Public Survey, designed to gather quantitative data relating to the opinions and attitudes towards privacy and the use of personal data. Expert interviews have been conducted, before being subjected to a thematic content analysis. A first phase explores the board decentralised space, with the

objective of understanding the trend and direction of travel, the technological usability barriers and opinions and views around sustainable adoption. A second phase focused on individuals from the decentralised space, with an active interest in Self-Sovereign Technology. These interviews are narrower in scope and focus specifically on user interaction and adoption. A technology Artefact Catalogue has been developed and subjected to critical analysis and Evaluation Matrix. The objective is to understand current trends and the trajectory in the domain. The final component involves a practice led endeavour, in which both a conceptual model and preliminary user interface for a decentralised system are defined. The direction and final outcome of this practice are realised by extending current decentralised trends to scale.

Chapter 5 presents the results arising from the research described in chapter four. First, the results of the Artefact Catalogue are outlined. A mapping of the collated Artefacts incorporating an Evaluation Matrix is illustrated, together with a description of defined categories, and specific artefact observations. The section concludes with a list of general findings, in the form of themes. The chapter progresses to present the findings of the Public Survey. Its context and implementation are described, followed by an overview of data and subsequent descriptive and correlative statistical analysis. Next, the findings of both phase one and two of the expert interviews are presented, together with the thematic analysis conducted, a visual mapping, and a detailed description of the derived themes. Finally, the results of the conceptual model of a decentralised system, and subsequent user interface are presented, from both the standpoint of critical reflection and focus group evaluation.

Chapter 6 is a discussion section, synthesizing the results arising from this research. The chapter begins with a summary of the work undertaken. It then proceeds to consider the meaning of the results for each element of data gathering, aligning interpretations with existing theory.

Chapter 7 presents the conclusions. The chapter draws together the results and discussion sections relative to the research questions. Initially, the research questions are explored and dissected before being answered in turn. The chapter then moves to summarise the main findings. The chapter continues to present a clear description of the contributions to knowledge, a reflection of the limitations of this research are communicated, recommendations are made for future research and investigation, and finally the thesis resolves with a personal reflection of the research journey.

Chapter 2: Literature Review

The core objective within this research is to investigate the potential for a sustainable interface layer for a Human-Centred Data Ecosystem. The primary areas of investigation have been defined as Privacy, Technology, and Adoption, there are subsequent intersecting areas of interest which include: Decentralised Models, Usability, and Value Proposition (see: 1.5). This review considers the pertinent arguments concerning decentralisation, and endeavours to investigate the relevant areas through the available literature.

The Literature review is divided into four sections, Surveillance Capitalism, Network Privacy, Human Computer Interaction, and Human Data Interaction. As such, the review undertakes a foundation investigation of surveillance, exploring the theoretical framework, before investigating arguments concerning personal data gathering, aggregation and secondary use. It investigates the historical narrative that has led to the status quo, and the relationship between large-scale data collection, and our digital economy. The review considers the notion of privacy, exploring the fundamental theory and framework, cultural differences and social norms. It explores the economic, social and cultural value of personal data. It investigates the legal landscape, and the arguments for the granting and restriction of privacy rights. The review considers privacy in the digital realm, investigates the positive aspects, and potential harms of big data collection. The review considers Human Computer Interaction, exploring the domain's progression, with a focus on cognition, investigating the theory most associated with individual interaction with both system and interface. Finally, the review considers the emergent domain of Human Data Interaction, charting its evolution, arguments for its realisation, and underlaying principles and trajectory.

In order to initiate the review, a systematic search of academic works has been conducted; the first stage involved a search of key phrases within the Scopus database. These phrases included, Surveillance Capitalism, Network Privacy, Human Computer Interaction and

Human Data Interaction. (Scopus, 2017). From these results, abstracts were read to ascertain the relevance of the literature; the papers were then read in detail. A second stage identified further reading from the references of the papers identified in the first phase.

2.1 Surveillance Capitalism

'*The harms from mass surveillance are many, and the costs to individuals and society as a whole disproportionately outweigh the benefits*' (Schneier, 2015, p. 112).

A dominant theoretical framework within surveillance studies is built on the concept of Panopticon. The Panopticon, conceived by social reformist Jeremy Bentham in 1791, is an architectural design for a penal institution whereby isolated accommodation cells are turned inwards towards a centre. A keeper, shielded from view at the centre, controls inmate's behaviour through a perception of constantly being visible and potentially observed. Gertrude Himmelfard (1968) explored the concept further within her book *Victorian Minds*, *The Haunted House of Jeremy Bentham.* Within this text Bentham's ideas are exposed as more than a method of reform, but one of social control by the capitalist. That saw Bentham himself as the contractor providing a cost-effective penal solution for the state, while mobilising the occupants to generate profit. Bentham not only envisaged the Panopticon model in a penal institution, but also one that might be applied to education, hospitals, mental asylums and factories. The model aimed to exploit and control the powerless, potentially to extract profit. The theory was re-examined and established as a prominent theoretical surveillance framework by French philosopher and social theorist, Michel Foucault (1975) within his text *Discipline and Punish*. This text explores the changes in western social control, where discipline has evolved to be metered within the mind, as opposed to the body. Foucault underlines the significance of the Panopticon. He argues for the transferability of the panoptic model, and the prospect of the Panopticon being a laboratory environment. He explains: *'it could be used as a machine to carry out experiments, to alter behaviour, to train or to correct individuals'*(p. 203).

Cybernetic Capitalism, Robins and Webster (1988) takes the Panopticon analogy into the digital landscape, as they argue a seismic shift far beyond ideology and economy, as technology penetrates deeply into the very fabric of everyday life. They refer to the: *'pervasive and intrusive nature of the information revolution.'* And the way in which: *'the rhythm, texture and experience of social life are being transformed and informed by Capital'* (p. 71). They argue that the information revolution marks the dawn of a new social era, a *'Panoptic Society'*, which as a whole begins to function as a:

> *Giant Panoptic Mechanism, Automatic and continuous surveillance, along with centralised power and peripheral isolation, conspire to create a climate in which the inmates of society not only suspect, but are assured that whatever they do is known* (p. 72).

Oscar Gandy (1996) defines the '*Panoptic Sort'* as a '*Complex and discriminatory technology, which produces intelligence on an individual's economic value'* (p. 113). He identifies the elements of the Panoptic Sort, as the *Identification* and *Classification* of the individual or group, and the eventual assessment in terms of the inclusion or exclusion from communication and marketing materials. He argues that the Panoptic Sort is a predictive mechanism, that offers significant reduction in uncertainty to the marketer. David Lyon (1993) defines the: *'Social Sort'* as the same mechanism extending the marketing focus into a wider context. He explains:

> *The surveillance system obtains personal and group data, in order to classify people and populations according to a varying criterion, to determine who should be targeted for special treatment, suspicion, eligibility, inclusion and access* (p. 20).

Within *The Mode Of Observation*, Mark Poster (1990) introduces the concept of *The Electronic Superpanopticon: 'a system of surveillance without walls, windows, towers or guards'* (p. 93). It is a system which can bring a deeper social context to written language, and where the so-called free individual has a second observed existence within the database, He

argues: *'we suggest that the free individual requires a repressed other, a sort of external super-ego'* (p. 23).

Campbell and Carson (2010) argue that our privacy is being commoditised through online surveillance, as we engage with Web 2.0 technologies, and that we are willing subjects to that end. They argue that: *'self-surveillance exists in cyberspace, in that individuals cooperate in the online gathering of data, about themselves as economic subjects'* (p. 586). They argue a shift in our perception of privacy, a re-conceptualisation of privacy, away from civil liberties towards that of a commodity exchange for perceived benefits. They argue that we perceive that we are missing out by not engaging, that there is a huge imbalance in power relations, and that the corporation wields disproportionate power in the marketplace. They explain: '*it is the corporation that appears to dictate the conditions of the marketplace, and, correspondingly, constructs and maintains are Participatory Panopticon*' (p. 589).

Alex Bruns (2006) has introduced the term *Produsage*. He argues: *'a paradigm shift away from industrial-style production models towards Produsage'* (p. 9). Web 2.0 technologies allow the traditional user of media and knowledge to become both user and producer. Bruns defines five common characteristics of Produsage and the ways in which the Prosumer can engage as an individual or group to generate or develop new content.

1. User-Led Content Production
2. Collaborative Engagement
3. Palimpsestic, Iterative, Evolutionary Development
4. Alternative Approaches to Intellectual Property
5. Heterarchical, Permeable Community Structures (p. 3)

Christian Fuchs (2012) argues that the Prosumer becomes part of *Marxist Class Theory* as they become productive labourers who produce surplus value. He explains that:

> *'If Internet users become productive Web 2.0 Prosumers, then in terms of Marxian class theory, this means that they become productive labourers who produce surplus value and are exploited by capital'* (p. 54).

He illustrates how the Prosumer Proletariat can be aligned to Marx's *Cycle of Capital Accumulation* within *The Sphere of Production* as they generate both content and data. He argues that Web 2.0 is *'not a participatory system, rather a system best described in more negative critical terms such as class, exploitation and surplus value'* (p. 53).

Shoshana Zuboff (2015) introduces the term *Surveillance Capitalism* and argues that each phase of capital requires a reinvention of a *Logic of Accumulation*. She argues that each technology company, of which Google is the pioneer, understands the great wealth of data extraction and analysis, and that *'this emerging logic is not only shared by Facebook and many other large Internet-based firms, it also appears to have become the default model for most online start-ups and applications'* (p. 77). She continues to argue that these new institutional facts are deliberately hidden from the user, and only a small number of experts understand their practices, and that tech companies are trusted as they are emissaries of the future. She argues that we have seen nothing of these methods in the past, that we have little defence against them in place, and that the rapid ascendance of the digital tools Surveillance Capitalism has provided, are now a requirement in our struggle for an effective life, and new paradigms of social inclusion. She asserts:

> *The rapid build-up of institutionalised facts – data brokerage, data analytics, data mining, professional specialisations, unimaginable cash flows, powerful network effects, state collaboration, hyper scale material assets, and unprecedented concentrations of information power produced an overwhelming sense of inevitability* (p. 85).

In summary Shoshana Zuboff does not paint a very optimistic picture for a positive network future, where the individual human is far from respected and valued within a data-driven society.

Jacob Silverman (2017) argues that: *'we are all entangled in these networks of information consumption and production'* (p. 162) and that apart from the rebel, the eccentric,

and the destitute, living off the network is all but impossible. He suggests that increased awareness of Surveillance Capitalism is leading to clandestine behaviour, a kind of thinking that is usually limited to *'dissidents, journalists, diplomats, artists—but it is largely an indulgent form of spy talk, one that reflects underlying principles of secrecy, vigilance, self-reliance, and suspicion of others'* (p. 160). He argues this is not a positive trajectory, and that we need change and should reject the *'current paradigm, which fails to acknowledge privacy as a shared, social good, one that benefits everyone, particularly the most vulnerable'* (p. 161).

Eben Moglen (2014) argues that we are at a crisis point, where *'we are on the verge of eliminating forever, the fundamental right to be alone in our thoughts'* (p. 20). He argues that the Internet is not engineered properly to protect privacy, and that this needs to be readdressed. He continues to suggest that our personal and collective data is *'intensively and instantaneously mined for its value, to those who would want to sell us something'* (p. 20), and that education needs to be vastly improved, to inform individuals of the real consequences of the centralised Internet.

There are counter arguments that try to look past the Panopticon, as a negative model of Web 2.0 participatory culture, and suggest that a new model is required. David Lyon (1993) argues a contradiction, in that the status-quo allows for horizontal observation and the Panopticon does not. Anders Albrechtslund (2008) introduces the concept of social networking as participatory surveillance and argues that once you switch the view of the surveillance hierarchy from the vertical prospective to the horizontal, the act of online social networking can be seen as empowering. He states that it is 'important not to assume automatically that information and communications becomes a commodity for trading' (p. 7), and that there are benefits from sharing your own constructive identities online with others.

The application of solely the Panopticon model, and its subsequent theories within the context of this research are over simplistic. It does, though, offer a solid foundation on which we can further develop theories and understanding. As Lyall King (2001) explains:

> *Even if some of the other facets of the Panopticon no longer wholly apply to the present-day situation, the inspection principle of the Panopticon remains alive and well, and with the ever-increasing advancements in technology, is only getting better and more efficient* (p. 49).

An important focus within this research, is the on-going technological development of software and hardware, by both the open source development community and activist groups. These organisations are pushing back against both state and capital, in the struggle for control of personal data, within our emerging digital landscape. It can be argued that a land grab has been made by a handful of now massively politically powerful technology companies, being first to market with a number of digital services, which engage the Participatory Panopticon, to mobilise the leisure time of the masses, to generate invisibly capital and data models of reality. The State it would seem is on the back foot, as it struggles to keep pace with this new reality, with a series of thinly veiled attempts to regain a legitimate foothold, through new legislation based on national security. While at the same time it would seem to be covertly, desperately or even recklessly disregarding any sense of personal privacy, as it intercepts and stockpiles communication traffic on mass. Privacy issues aside, it is also argued that our interaction and our digital leisure activities are being commoditised, and that by engaging with Web 2.0, defines us as a labour force, with all the ramifications and consequences laid down in the theories of Karl Marx (Fuchs, 2012).

2.2 Network Privacy

The following section explores the notion of privacy, its evolution and cultural context, and the surrounding arguments that are relative to this research.

2.2.1 Privacy as a Complex Social Issue

Allen Western (1967) defines privacy as: *'the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent his or her personal information is communicated to others'* (p. 7). This definition is clear, but when applied to the complexity of the real world, it becomes evident that privacy as a concept is not only incredibly complex, but also poorly defined and misunderstood. Robert Post (2001) explains: *'Privacy is a value so complex, so entangled in competing and contradictory dimensions, so engorged with various and distinct meanings that I sometimes despair whether it can be usefully addressed at all'* (p. 2087). Judith Thomas (1975) observed of privacy that *'perhaps the most striking thing about the right to privacy, is that nobody seems to have any clear idea of what it is'* (p. 272). Privacy in a cultural and political context, adds another dimension to its complexity, as the social norms established by different cultural groups, across different nation states vary considerably. The United Nations (1948) recognises privacy as a human right with the following definition: *'No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor attacks upon his honour and reputation'* (art. 12). Within the European Convention of Human Rights (1950), it is stated that: *'everybody has the right to respect, for his private and family life, his home and his correspondence'* (art. 8). Daniel Solove (2008) describes privacy as a sweeping concept including *'freedom of thought, control over one's body, solitude in one's home, control over personal information, freedom from surveillance, protection of one's reputation, and protections from searches and interrogation'* (p. 1). It is perhaps the fact that the concept is so thinly spread that the articulation of actual privacy harms seems so difficult. In law quantifying privacy harms that are not physically obvious, directly economically impacting, or that manifest at a future time can be difficult, e.g. causing future reputational damage, embarrassment or emotional distress. Julie Cohen (2012) argues that:

> *This difficulty combined with a perception of Privacy as old fashioned at best, and downright harmful at worst, puts the notion of Privacy on the back foot when pitched against the cutting-edge imperatives of national security, efficiency, and entrepreneurship* (p. 1904).

Within the seminal article, *The Right to Be Let Alone* by Samuel Warren and Louis Brandeisis, privacy is classified into six general categories':

1. The right to be let alone – an individual's right to solitude.
2. Limited access to the self – to be able to shield one's self from others.
3. Secrecy – to be able to conceal matters from others.
4. Control over personal information.
5. Personhood – the ability to protect one's personality, individuality and dignity.
6. Intimacy – control over access to one's intimate relationships.

(Warren and Brandeis, 1890).

The majority of US privacy policy has been built upon these founding principles. Over time specifically through the *Records Computers and The Rights of Citizens* report produced in 1973 (US-Gov, 1973). The UK and European Union have defined their directives with regard to privacy as concerning digital records, the definition of digital records and subsequent rights of citizens in this regard. It is interesting to note that the current EU regulation and new GDPR regulations have their roots in US policy, given the fact that the US is often considered to hold personal privacy generally in a lesser regard.

There are many arguments for the need for privacy: It is a fundamental part of our social structure. To have a society without a degree of non-disclosure of private thought, action, property or information would be impossible to achieve. Privacy is fundamental to our notion of self, to our independence and sense of dignity. It is part of our cognitive development, as we first understand that those around us do not have access to our inner thoughts and ideas. In choosing to disclose our emotions, our desires, our motivations or political positions, we

develop complex social structures and intimate relationships. Privacy is a critical component of our democracy, and our western liberal society.

'Privacy has a role to play in our psychological well-being and the health of our democracies.' (Ohara, 2016, p. 86).

'There is widespread belief that privacy is essential to our well-being, physically, psychologically, socially and morally' (US-Gov, 1973, p. 33).

'Privacy enables people to create, explore, and experiment. It provides moments for intellectual and spiritual contemplation.' (Solove, 2008, p. 79).

'Privacy enables people to grow, maintain their mental health and autonomy, create and maintain human relations, and lead meaningful lives.' (Gavison, 1984, p. 423).

It can be argued that an equilibrium needs to be achieved in order to maintain a balanced society. If we take the classic liberal position, where the individual precedes society, and that an individual, if so inclined, has the right to withdraw from society and give nothing of himself. The consequence can be detrimental and fails to consider the social value of privacy. It is argued that Privacy is important from an individualistic prospective, but in turn protecting against privacy harms, should be primarily concerned with maintaining the balance and function of our society. Daniel Solove (2008) argues that:

> *Privacy harms affect the nature of society and impede individual activities, that contribute to the greater social good.' He goes on to claim that 'privacy should be considered alongside its impact on society,' and that 'the value of privacy should be assessed on the basis of its contributions. Protecting individual privacy need not be at society's expense* (p. 90).

These social concepts of privacy are central considerations within this research. Considering potential value propositions, as evolving asymmetric relationships within the network, not only risk individual privacy harms, but arguably also destabilise the balance required for a functional society.

2.2.2 Electronic Information Privacy

The advent of ubiquitous digital technology, and surveillance economics add increased complexity to the issue of privacy. General advances in technology prior to the digital age, and their impact on privacy issues, have been recognised and commented upon historically. Warren and Brandeis (1890) describe their concerns regarding *'recent inventions and business practices'* and the impact these are likely to have on personal sovereignty (p. 195). Allan Westin (1967) expresses: *'a deep concern over the preservation of privacy under new pressures from surveillance technologies'* (p. 26). Richard Mason (1996) see's Information Privacy, as one of the most serious ethical debates of the information age.

In defining the theoretical boundaries for this research, in the context of privacy, concerns regarding state surveillance, and the broader concepts of privacy infringement outside of the digital network, are considered to be out of scope. Instead focus is given to theory and arguments relating to the generation and secondary usage of data, as individuals traverse the Internet as Prosumer (Fuchs, 2012). To put this into context, Alessandro Acquisti (2016) describes how: *'we live in a consumer data-driven and consumer data-focused commercial revolution, in which individuals are at the same time consumers and producers of a valuable asset: their personal information'* (p. 8).

It can be argued, that our western liberal democracy, functions on a 20th century capitalist model reliant on data, and that until plausible alternative business practice can be defined and adopted, the withdrawal of personal information and subsequent market intelligence, may cause significant economic harm. It is also worth noting, that many of the privacy harms that may occur from the disclosure of personal data, derive primarily from economic activity, and subsequent filtering sideways through resale and secondary use. Alessandro Acquisti (2016) argues that privacy related trade-offs, do not necessarily *'imply that the economic agents are always aware of them, as they take decisions that will impact privacy'* (p. 4). The origin, consent, and secondary use of personal data within the *Economics of Privacy* is complex. It can be argued that personal data has become more than market intelligence, and has itself morphed into an independent asset class, which if left unregulated, drawn through opaque practices, without accessible agency, will ultimately be exploited by capital.

Arguments have long been made against regulation of personal data and privacy relative to economics. The most significant come from the Chicago School[13], where scholars argued that for a perfectly competitive market to emerge, complete information of all market participants is required, which in turn leads to economic efficiency. *'People should not, on economic grounds, in any event, have the right to conceal material facts about themselves'* (Posner, 1977, p. 394). The argument claims that protection of privacy creates inefficiencies in the marketplace, as it conceals potentially relevant information. It is worth noting that this concept of disclosure concerns both marketeer and consumer. It is argued that:

> *The early Chicago School studies of privacy originated in what may be defined as pre-ICT' and that more recently, given the advent of the network, researchers have been led 'to formulate more nuanced and granular views of trade-offs associated with privacy protection and data sharing* (Acquisti, 2010, p. 7).

13: The Chicago School of Economics is an economic school of thought, that originated within the department of economics at the University of Chicago in the 1930's

The Chicago position states that the market can protect privacy, and that regulation is not necessary, because if a piece of personal information is considered sensitive, it will inherently have value, and that if the data subject values that information sufficiently enough, they will act to protect it.

This would seem to be based on the notion that the data subject is:

1. Aware of the data being disclosed.
2. Has the agency to protect the data.
3. Is capable of acting in a rational way to protect their own privacy and appreciates the notion of its value.

With regard point three, Ian Brown (2013) concludes that:

> *Behavioural economics research is improving the understanding of cognitive biases, that can lead to non-optimal privacy decisions by individuals. Bounded rationality, time inconsistent preferences, optimism bias and context dependence have all been demonstrated to apply to information disclosure decisions* (p. 13).

Further evidence of poor bias decisions in the disclosure of personal information, can be found in the exploration of the Privacy Paradox (Norberg, 2007), (see: 2.2.8).With regard to point two and agency, this topic is explored further in the Human Data Interaction section of this document (see: 2.4). In relation to point one, and the understanding of disclosure, it is worth considering the points raised in *The Rights of Citizens* report, (US-Gov, 1973).

Within this report it recommends the following five privacy rights relative to digital information:

- **Notice and Awareness.** Secret records should not exist, individuals should have notice of a collector's information practices, and they should know identity of those holding their information. How the information will be collected, the consequences of declining data collection, and the steps taken by the collector to ensure confidentiality, integrity and quality of the data, also need to be considered.

- **Choice and Consent.** Individuals must have a choice regarding how their information is used. They must have a choice regarding whether data can be used in a way that goes beyond that necessary to complete the original transaction. In both areas genuine informed consent is required. Individuals who provide information under duress or threat of penalty, have not provided informed consent. Neither have individuals who provide personal information as a requirement for receiving necessary or desirable services from monopoly providers.
- **Access and Participation.** Individuals should be able to review in a timely and inexpensive way, the data collected about them. They should be able to contest its accuracy and completeness and have the means to correct it.
- **Integrity and Security.** The data held regarding individuals must be accurate and secure, Data collectors must take both procedural and technical measures to protect against loss and the unauthorised access, destruction, use or disclosure of the data.
- **Enforcement and Redress.** Enforcement mechanisms must exist to ensure fair information practices are observed, and individuals must have a mechanism for redress if these principles are violated.

Current EU regulation and the GDPR 'General Data Protection Regulation' are based on these principles. But the GDPR goes further in requiring data agents to gain meaningful consent, with full disclosure of secondary data usage. It can be argued that if we accept the arguments of *Cognitive Bias* and *Immediate Gratification Bias* (Acquisti, 2004), for the digital native who is locked into the network, no amount of informed consent can protect the participant. If an individual agrees to the Contract of Adhesion laid before them, their data can be used in any way the data collector specifies.

Commentary has been made about the confusion surrounding regulation of data privacy, and that its importance should be considered, or at the very least we should try to understand better what it really is we are trying to protect. Kent Walker (2000) states:

> *That before leaping to establish new information rights, we should carefully review not just the benefits of information privacy, but also the benefits of the information and the costs of regulating privacy* (p. 2).

Jeff Jarvis (2011) sees the status-quo as:

> *A confused web of worries, changing norms, varying cultural moves, complicated relationships, conflicting motives, vague feelings of danger with sporadic specific evidence of harm, and unclear laws and regulations made all the more complex by context* (p.101).

In making this observation, Jarvis is arguing that we should stop pushing against the progress of the network and relinquish our privacy for society's greater good. Many within the technology community have argued that the notion of privacy as a social norm is now changing. Scott McNealy CEO of Sun Micro-Systems said: *'You have zero privacy anyway, get over it'* (Manes, 2000, p. 312), and Mark Zuckerberg in 2010 stated that: *'People have really gotten comfortable not only sharing more information of different kinds, more openly and with more people, that social norms is just something that has evolved over time'* (CES, 2010). Kieran O'Hara (2013) eloquently referred to Zuckerburg's claim as *'Zuckerbollocks'*. He said: *'Now it is really disingenuous to maintain that high-volume social networking sites, have had no effect on behaviour, merely passively reflecting and serving evolving social norms*' (p. 90). Max Van-Kleek (2014) extends the term to include also the notion that this course of direction, is somehow fixed, and in some sense destiny: *'In short, the argument is often made that the technological direction of travel is more or less set, that it serves the public good, that the public is uninterested in any alternative, and so, to coin a phrase, get over it'* (p. 16). Mark Zuckerburg is also on record as suggesting that the norm of multiple identities is over. He claimed: *'you have one identity. The days of you having a different image for your work friends or co-workers, and for the other people you know are probably coming to an end pretty quickly. Having two identities for yourself is an example of lack of integrity'* (Kirkpatrick, 2010, p. 199). Bruce Schneier (2015) describes this as remarkably naive and counters by stating: *'We reveal different facets of ourselves to different people. This is something innately human. Privacy is what allows us to act appropriately in whatever setting we find ourselves'* (p. 147).

There is clearly a three-way tension, between those that see absolute privacy as a given right, those that see any privacy affordance as part of the balance between individual agency and the greater good of society, and those who wish to maintain the capitalist trajectory which has become so reliant on personal data as its means of lubrication. It can be argued that there needs to be a balance, and that the current model has moved too far towards exploitation and profit. Our digital economy needs data, and without any rational alternative, to cut off the flow would be damaging. Equally, to allow the current trajectory to continue, threatens to change our social norms beyond all recognition. The development of a Human-Centred Data Ecosystem may well achieve a much-needed balance, though the complexity in achieving this reality having moved so far in one direction, may prove very challenging.

2.2.3 The Right to Be Forgotten

Bruce Schneier (2015) makes a strong argument for the ephemeral, or the *Right to be forgotten*. He argues that the age of mass surveillance marks the end of prehistory, when a continuous record of transaction, interactions and communication eradicate forever the concept of the ephemeral. It can be argued that by our very nature, our social interactions are reliant on an ability to be ephemeral. We need to be able to speak freely knowing our words, thoughts and expressions are only in the present, and only heard by the people with whom we are communicating. The details and the happenings of the past, our capacity to forget, and for painful memories to fade out of existence, is part of the process of healing. If this is removed, we step into a new paradigm. Forgetting is an important mechanism for forgiveness. Schneier elaborates: *'I'm not convinced that my marriage would be improved by an ability to produce transcripts of old arguments. Losing the ephemeral will be an enormous social and psychological change, and not one that I think our society is prepared for'* (p. 151).

2.2.4 Algorithmic Surveillance

Algorithmic Surveillance is a process of automation, where the computer conducts an analysis of the data without a human being part of the process. It has been used by security agencies and companies like Google to defend its retention and analysis of personal data and communications, with the argument that retained personal data is not surveillance or an infringement of privacy, until a human-being has looked at it, and interpreted the data. In Google's case this referred to the analysis of the content of Gmail messages, and the subsequent targeted marketing received by users. There was a prominent defence by US Security Agencies, James Clapper, US Director of National Intelligence. He argued that the collection of personal data from individuals was like a huge library, and that there was no infringement of privacy until one took '*the book off the shelf and opened it up and read it.*' (Kessler, 2013). This position came as defence to his testimony that the US did not collect the communications records of all of its citizens. It was later contradicted by the information released by Edward Snowden. A counter argument to this position comes from Bruce Schneier (2015), in that at any time an algorithm could flag a pattern and draw it to the attention of a human. The fact is that a human developed the algorithm in the first place, and that you can never be sure exactly what software is designed to do. He suggests that if an individual is opposed to this practice, *'it's because they realise that the privacy harm comes from the automatic collection and the algorithmic analysis, regardless of whether or not a person is directly involved in the process'* (p. 153). He goes on to argue that there are no assurances: *'You have no way of confirming that no person will perceive whatever decision the computer makes, and that you won't be judged or discriminated against on the basis of what the computer sees'* (p. 153).

2.2.5 Anonymised Data

A cornerstone of privacy law and policy is built around the anonymisation of PII, (*Personally Identifiable Information)*. The argument is that companies and organisations can

collect, process and use data as long so it is anonymised, and cannot be directly linked back to an individual. Paul Ohm (2010) argues that: *'nearly every information privacy law, and regulation, grants a get-out-of-jail-free card to those who anonymise their data'* (p. 1704). He goes on to argue that personal data cannot be both useful and anonymous, and that the more a piece of data is anonymised, the less value it has. He suggests that anonymisation of personal data is a myth, and that any dataset can be to a greater or lesser degree be re-identified, once aggregated against additional external data sets. He argues that data scholars are wrong to have a *'deep faith in robust anonymisation',* and that *'regulators should prevent privacy harm by squeezing and reducing the flow of information in society. Even though in doing so they may need to sacrifice important counter values like innovation, free speech, and security'* (p. 1701). He goes on to claim that we have fundamentally *'made a mistake'* in putting our faith in anonymisation.

2.2.6 Nothing to Hide

In 2011, Google CEO, Eric Schmidt is quoted as saying: *'If you have something you don't want anyone to know, maybe you shouldn't be doing it'* (Huffpost, 2010). This argument is voiced in many different guises and is often used in the privacy verses security debate. It is often phrased as, if you have nothing to hide, and are a law-abiding citizen, then you have nothing to worry about. Daniel Solove (2008) argues, that the problem with the *nothing to hide* argument, is that it makes the *'assumption that privacy is about hiding bad things'*, and that this view is too narrow, focusing only on surveillance, and does not consider the broader issues. He uses a Kafka metaphor, alluding to the novel *The Trial*, and the potential harms caused by *'bureaucracies, indifference, errors, abuses, frustration and lack of transparency and accountability'* (p. 766). He goes on to discuss the issue of exclusion, and the way in which individuals are prevented from knowing, what information is being held about them, or how it is being used. This, in turn, he argues, leads to a power imbalance and affects our relationships

with government. *'This issue is not about whether the information gathered is something people want to hide, but rather about the power and the structure of government'* (p. 767).

2.2.7 Data Mining / Aggregation

When discussing the concept of inference, Max Van Kleek (2014) comments:

> *'By providing a route for others to understand what we are, or what we have done, or where we are situated, it can threaten our privacy, or our dignity, or our autonomy, by diluting the privileged first-person access to our own experience. It creates the possibility of our being counted, measured, judged, steered or influenced without our knowledge, by mysterious forces or organisations who may or may not have our best interests at heart'* (p. 5).

Identified as possibly the most intrusive of data privacy issues, processing of personal data through data mining or aggregation, is the secondary use of data, in which inferences can be drawn about an individual, by looking for patterns. Digital dossiers are established and traded by data brokers to third party companies, for a variety of different uses. In what is possibly the most cited case, the New York Times (Duhigg, 2012) published an article about an American man, who complained to the US Target store chain, that his teenage daughter had been inappropriately sent coupons for baby related items. It transpired that unknown to him, his daughter was pregnant, and that the company's algorithm had inferred the fact from the data subject's purchasing history. Combining data sets from different sources can reveal a great deal about the data subject. Danial Solove (2009) explains: *'When analysed, aggregated information can reveal new facts about a person, that she did not expect would be known about her when the original, isolated data was collected'* (p. 118). By establishing a rich collection of data about an individual comprising of demographic, social position, browsing history, media ingestion, patterns of behaviour, location, spending habits and communications, cross referenced with personal affiliations, a very clear understanding of an individual's psyche and predictive behaviour is achievable.

The seminal work in this area has been conducted by Dr Michal Kosinski et al (2013). In their paper entitled, *Private Traits and Attributes are Predictable from Digital Records of Human Behaviour*. Kosinski demonstrates a powerful method to develop accurate individual psycho-demographic profiles through the analysis of Facebook Likes. Within this research Kosinski demonstrated that: *'the model correctly discriminates between homosexual and heterosexual men in 88% of cases, African Americans and Caucasian Americans in 95% of cases, and between Democrat and Republican in 85% of cases'* (p. 5803). This method is broadly accepted to be the one used by Cambridge Analytica (2017) which sparked controversy and accusation of electoral manipulation (Rosenberg, 2018).

Kosinski's work demonstrates that a clear picture of an individual's personality can be drawn from information readily available. More importantly we can now predict with relative certainty how an individual is likely to react when posed with a specific question or situation. It is this vision, that Eben Moglen (2015) alludes to in his argument, that we are about to give up our right to be alone in our private thoughts. *'We are on the verge of eliminating forever the fundamental right to be alone in our thoughts'* (p. 1). To follow this to a truly Orwellian conclusion, it can be argued that given the means to understand the broad social graph, and the power to influence, a handful of once liberal California technology companies, will soon wield power the media moguls of the late 20th century could only dream about. Mortier elaborates:

> *The data is accumulated about us by many different organisations; some competing, some collaborating, and they are processed using increasingly sophisticated algorithms to measure and infer increasingly sensitive features of our lives* (p. 3).

There is clearly a tension in the digital context between those who apply the Silicon Valley mass data collection and target marketing business model, and privacy advocates who see the collection and potential exploitation of such huge personal datasets, far exceeding their primary justified means.

2.2.8 The Privacy Paradox

The Privacy Paradox describes a dichotomy between attitudes and behaviour concerning network privacy. It is claimed that individuals voice concern about their privacy online, only then to act in a way that demonstrates little concern for their private information, often releasing personal private information for very little reward. Patricia Norberg (2007) explains:

> *For all the concern that people express about their personal information, which could be expected to drive one's intended and actual disclosure, our observations of actual marketplace behaviour anecdotally suggest that people are less than selective and often cavalier in the protection of their own data profiles* (p. 101).

The theory of a Privacy Paradox is contested, with several studies claiming to demonstrate the existence of this phenomenon, while a number seem to contradict this theory demonstrating that individuals are concerned and do act accordingly to protect their information. Barry Brown (2001) commented on, *'Something of a privacy paradox'* (p. 1) when observing the way online shoppers tended to voice concern about privacy only to give personal information as long as they felt they were getting something in return. In 2004 Acquisti argued that individuals may not be able to act rationally in an economic transaction when it comes to personal data. He extended behavioural economics literature to describe what he termed *'Immediate Gratification Bias'* (p.2), a term which suggests that individuals place higher value on immediate benefits rather than future risks. He also claimed that sophisticated privacy advocates might understand the futility of trying to protect their data, and as a result become apathetic in attempting to do so.

Susan Barnes (2006) makes a reference to the Privacy Paradox in the context of how young people use social media applications, and their lack of understanding of the public nature of the platforms, and the way in which their personal data might be exploited and abused through secondary use. Norberg (2007) solidifies the term Privacy Paradox with a further

explanation of the phenomenon. She argues that there are two elements to consider in terms of behavioural intention and actual disclosure behaviour: *Risk* and *Trust*. Behavioural intention is influenced by perceived risk, which in turn is heavily influenced by the media and popular opinion. But the actual disclosure of information is influenced by the trust an individual has in the entity engaged with. Alastair Beresford (2012) conducted an experiment in which two identical websites sold DVD's at the same price. One of the websites asked for basic information as part of the transaction process. The other asked for significantly more personal data. Individuals where asked to choose between the two competing stores. The results indicated little preference across the two sites. In an additional experiment, if the price was lowered by one euro on the site requesting more in-depth personal information, the vast majority of participates gravitated to this offering. In a post-experimental questionnaire, 75% of participants voiced concerns about their privacy online.

Despite studies supporting the existence of a Privacy Paradox, there are studies that provide evidence to the contrary. In a study by Giles D'Souza (2009), the perception of the secondary use of data was shown to have an impact on the purchase likelihood. In a similar study by Robin Wakefield (2013), the issue of trust in the online artefact was again demonstrated to impact likelihood in information disclosure. Trust and the explicit communication of privacy policy is demonstrated by Janice Tsai (2011) when the probability of purchase was shown to be more likely if the privacy policy of the website is prominently displayed to the user during the purchase process.

What at first appeared to be a Paradox, is, in fact, a more complex collection of human behaviours. Spyros Kokolakis (2017) explains: *'The dichotomy between privacy attitude and behaviour should not be considered a paradox anymore, since recent literature provides several logical explanations'* (p. 130). To understand how individuals might protect their data,

and their desire to do so across multiple contexts, the Privacy Paradox literature provides a valuable reference.

2.2.9 Bounded Rationality & Optimism bias

Bounded Reality argues that many economic predictions of an individual's behavior and decision making when forming choices are based on a capability to act rationally. Herbert (1955) argues that true rational decision making requires a complete understanding of alternative choices and their consequences and would require an infinite time to deliberate. Instead Herbert suggests that an individual's capability to act in a rational way is bounded by the individual's tractability, the cognitive limitations of the mind and the time available to make any decision. Herbert (1955) comments that an 'organism's simplifications of the real world for purposes of choice introduce discrepancies between the simplified model and the reality' (p. 114). When considering the development of any decentralised system we cannot assume that an individual will action rationally in the classic sense, instead an individual may act in a way that reflects their own reality and understanding of the world.

Optimism Bias is a cognitive process by which an individual believes that they are less likely to experience a negative occurrence then is statistically probable in reality, Sharot (2011) explains, *'humans, exhibit a pervasive and surprising bias: when it comes to predicting what will happen to us tomorrow, next week, or fifty years from now, we overestimate the likelihood of positive events, and underestimate the likelihood of negative events'* (p. 941). She argues that *'because people update their beliefs more in response to positive information about the future than to negative information about the future'* (p. 943). In the development of value proposition for a human centred data ecosystem the argument for adoption needs to overcome this general optimism found in the population, together with the obstacles posed by the *Privacy Paradox* (Norberg, 2007) and *Immediate Gratification Bias* (Acquisti, 2004).

2.2.10 Privacy Law

This research is primarily focused on data protection and privacy law as it stands within North America and the EU. This is because the vast majority of network traffic, data collection and secondary use manifests between these two territories. In addition to the current laws and cultural interpretations, it is important also to consider the impact of the new EU data protection laws, the GDPR 'General Data Protection Regulations', and the implications this may have on the way data is collected, processed and subsequently used by businesses and organisations. This applies both in the US and the EU, and the opportunities or restrictions this new legislation may have on the design and implementation of a data management interface layer.

It is argued that European citizens value their privacy more than those in the US. The European attitude towards privacy is likened to the passion felt by US citizens with regard gun control. Privacy in the EU is argued to be of greater concern to its citizens as a consequence of the cultural shadows laid down by the oppression of recent history. In current EU law there are two elements that protect privacy under the EU chapter of Fundamental Rights 2000. Article 7, which is a broad directive relating to a wide spectrum of privacy concerns, and Article 8 is specific to personal data with the statement: *'Everybody has the right to the protection of personal data concerning him or her.'* These laws are directives, meaning they are interpreted and enforced by each member state independently. In the US there are no specific umbrella like constitutional laws regarding privacy. Instead there are sector specific laws relating to personal information and the way it is handled. These areas include health, financial institutes and credit scoring. There are some federal laws relating to child protection, and there are state specific laws relating to privacy. California is an example of a state that has hundreds of independent laws relating to privacy controls. In the US, infringement of privacy law is termed as *Unfair and Deceptive Practices* and is enforced by the Federal Trade Commission. The consequences of infringement can be very severe, running into fines of tens of millions of

dollars and the agreement to company audits for decades. In the EU, because enforcement is handled at a state level, it can be difficult, and the levels of penalties handed down are very low. In essence, the EU regulations though clearer to understand lack any real consequence.

Consent also differs between the US and the EU. In the US it works on the basis of the user having to opt-out, if a user consents to the terms and conditions of the service provider, all data can be collected, exploited and held indefinitely. In the EU, the consent works on the basis of opt-in. Terms and conditions cannot be buried deep within the text and must instead be made clear to the user. In the EU, data can be also used without the need for consent, so long as it falls under the terms of legitimate interest.

In the US, there are no restrictions on the way in which personal data can be exported outside of the territory. In the EU there are restrictions, and data cannot be exported unless there is an agreement in place with the receiving business or organisation. These agreements are known by the terms Binding Corporate Rules, Model Clauses and Safe Harbor.

In the EU data protection law has been revised in the form of the GDPR 'General Data Protection Regulations' (2016). This law was adopted by the EU in April 2016 and became enforceable in May 2018. The law is designed to improve the way personal data is gathered, processed and used, with significant revision and new legislation to improve the rights of network users. The law applies to everybody along the data processing chain, including the data processor and custodian, and applies to companies dealing with the data of EU citizens regardless of the geographical location.

The regulations centre around eight main points:

- Consent needs to be explicit and clearly communicated to the user.
- Hiding terms and conditions deep within policy documents or relying on contracts of adhesion will no longer be acceptable.
- It must be clear that the user fully understands what data is being gathered and what will then happen to that data.
- Users must also have the capability to withdraw consent as easily as it is granted.
- Breach notification requires a company to communicate the loss or compromising of personal data to data controllers and customer within 72 hours.
- *Right to Access* grants the data subject the right to know if data is being processed and the user can request a copy of the data for free.
- *Right to be forgotten* allows the data subject to request personal data be erased once it is no longer relevant to its original purpose.
- Data Portability means that any company or organisation collecting personal data must make that data available to the data subject in a useable transferable format.

In addition, it is recommended that *Privacy by Design*[14] be implemented. This means companies should establish appropriate technical and structural measures wherever possible and should design data security into their services reducing the possibility of any data breaches. Professionally qualified data protection officers need to be appointed by any organisation that has more than 250 employees, and systematic monitoring of processing of sensitive personal data must be undertaken. Enforcement of this new regulation is supported by significant penalties for non-compliance, allowing for penalties of up to 20 million euros or 4% of gross turn-over.

2.2.11 A Pragmatic Approach to Network Privacy

Daniel Solove (2008) suggests a new theory to define privacy and privacy problems. He argues that on-going efforts to isolate common denominators in all instances of privacy and *'attempts to locate the essential or core characteristics of privacy has led to failure'* (p. 08).

14: *Privacy by Design* relates the principals of Privacy Enhanced Technology and originates from the report, *Privacy-Enhancing Technologies: The Path to Anonymity* (Cavoukian, 1995).

He argues that efforts to locate the essence of privacy are either too broad, or too narrow, and have proved unsatisfactory. He argues that the conceptualisation of privacy is of: *'paramount importance for the information age because we are beset with a number of complex privacy problems that cause great disruption to numerous important activities of high social value'.* He suggests that instead of a top down approach we should come from the bottom up to *'understand privacy as a set of protections against a plurality of distinct but related problems'* (p. 171). The term *'privacy'* then acts as an umbrella term to cover these protections. He argues that we should see privacy issues through the lens of the problem and argues a pragmatic approach that resists universals and embraces specific solutions, and that we should *'understand privacy in specific contextual situations'* (p. 47). This set of defined problems does not rely on a shared common denominator. They instead share characteristics in the spirit of Wittgenstein's[15] notion of family resemblance. Daniel Solove has defined taxonomy of four general types of privacy problems with sixteen different subgroups as identified in figure 4.

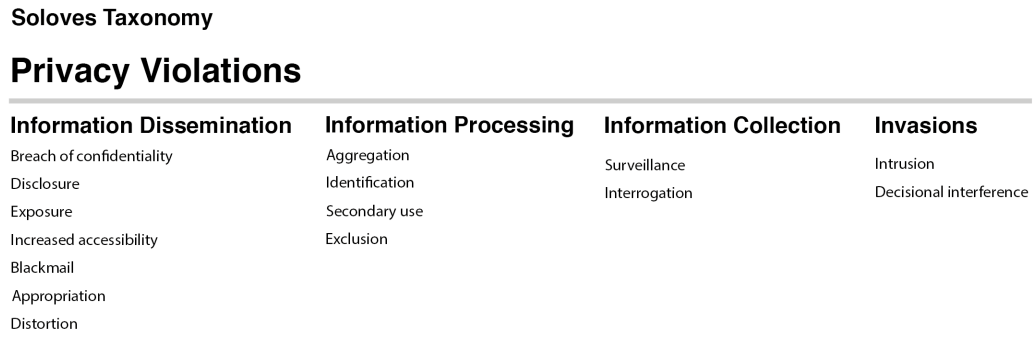


Figure 4: Daniel Solove's: A Taxonomy of Privacy (2008, p. 104)

Danial Solove's taxonomy is important to this research, as it offers a framework from which to explore real world privacy issues, user journeys and potential privacy harms. This research argues that the theory can be extended, not only to support law and policy makers, but also to inform the development of decentralised systems, tools and services, genuine value proposition, and communications strategy.

15: Wittgenstein's Family Resemblance is a philosophical position that argues that things that are thought of as being connected through an essential common feature, are in fact connected by a number of overlapping similarities. (Wittgenstein, 1953)

2.3 Human Computer Interaction

The following literature has been compiled in order to explore the relevant theories of Human Computer Interaction (hereafter HCI) related to this research. The field of HCI is broad and multifaceted, spanning several paradigms, across a number of decades, and it is important to target the pertinent theory. This research is not purely linear in nature, and a clearer understanding of the relevant arguments has evolved by reflecting on the broader literature, the principles of the decentralised domain, evolving technology and contemporary practice.

An important argument within this research, is that some of the earlier HCI theory, which may have fallen out of favour in the contemporary discourse, may still hold value in the light of current developments and proposed decentralised principles. The rationale being that developmental trends, towards a Human-Centred Data Ecosystem through what is termed in this research as a Sovereign Boundary Mechanism, would seem to suggest an interaction which in many respects is cognitively internalised. As such, many of the classical theories of HCI are covered in this section, together with more modern theory, centred around External and Distributed Cognition, and Situated Action.

. Given the broad scope of HCI literature, several prominent themes have been defined in order to focus investigation:

- Internalised Cognitive Theory and Mental Model, relative to the initial adoption of complex technology used in isolation.
- The theory of Externa*l* and Distributed Cognition in the context of general usability, understanding of interface, and reasoning and higher problem solving.
- The theories of Situated Action, in terms of the varied contexts in which decentralised technologies may be utilised.
- The requirement for collaborative engagement and societal interactions in the building of Trust Frameworks.
- The ethical conversation around building software that has a benefit to the individual participant and wider society.

In early work, Harrison (2007) proposes *Three Paradigms of HCI*, a collection of intellectual waves which together define the field. The first establishes the foundations, is concerned with *Man Machine Fit*, and the early endeavours from an engineering and ergonomic perspective, allowing the human to connect and control computer processes and interactions. The second paradigm is related to cognitive psychology, and the modelling of what is occurring internally within the users' minds. Harrison makes the argument for a third paradigm, related to issues arising from more contemporary study, which fit poorly within the confines of the first two paradigms. These include: *'Embodiment, situated meaning, values, and social issues'* (p. 1). Harrison suggests that the way we think about the development of HCI theory and evaluation methods, spans several new approach's which may be applied in various degrees across differing contexts. To understand this emerging collection of theories and methods, Harrison suggests a *Phenomenological Matrix*, which can be used to frame better, to compare and understand the emerging considerations. This is relative to this research, as early investigation suggests that a decentralised model will require a mixture of HCI theory. In more recent times, Yvonne Rogers (2012), again recognises three phases of development, though marginally varied to those proposed by Harrison. Rogers likens the phases of HCI development to those of the arts, classical, modern, and contemporary. Classical deals with theory concerning cognitive psychology. Modern, concerns the theories' related to the interplay between external and internal representations. Finally, contemporary, considers theory's relating to Values, Design, Culture and Embodiment.

This research is concerned with the development of an interface layer for a Decentralised Internet. There are several guidelines for developing usable interfaces, with the seminal concept of affective interface design, surmised in the work of Hutchins et al, through the following heuristics:

(a) *Principle of Visibility:* Users should be able to 'see' the actions that are open to them at every choice point.

(b) Users should receive immediate feedback about the actions they have just taken.

(c) Users should get timely and insightful information about the consequences of their actions. (Hutchins, 1996).

Norman expands the *Principles of Interaction Design* to include mapping, consistency, and affordance (Norman, 2013). There has been a series of heuristics developed to inform and guide the development of interactions between humans and computers. For example: Shneiderman's 8 Golden Rules, (Shneiderman, 1987) and Nielsen's 10 Usability Heuristics (Nielsen, 1994). General heuristics are useful as a guide, and a means of evaluation for designers. However, in the case of this research, in which original interaction concepts are explored, the deeper realm of Cognitive Psychology is considered to be the primary focus.

2.3.1 Applying Cognitive Psychology to HCI

Donald Norman (1986) describes *Cognitive Engineering* as: *'Neither Cognitive Psychology, nor Cognitive Science, nor Human Factors. It is a type of applied Cognitive Science, trying to apply what is known from science to the design and construction of machines'* (p. 31). A recognised contribution drawn from cognitive psychology and transposed into the field of HCI, is the description of the capabilities and limitations of participants, when engaging in computer-based tasks. A number of the fundamental theories include:

- **George Millers Theory of Memory** argues that individuals have a limited capability to remember units of information, numbers, names or places. Also, on average an individual can retain 7, plus or minus 2, at any given time. Millers theories of *Chunking* has been adopted by the design community, and in some respects re-appropriated in ways which are arguably out of context. Practitioners should be aware that the evidence

of Chunking, was derived from controlled experiments, and do not necessarily map on to every conceivable digital interaction (Miller, 1956).

- **Fitts Law** relates to perception and action, and is described as, *'The time taken to acquire a target, is a function of the target, and the distance to it.'* This basic theory underpins the requirements of many graphic user interfaces in the positioning of buttons relative to their importance and hierarchy. The argument being that a well-positioned button or interaction will allow the user to perceive and interact with it more rapidly, making interaction faster (Fitts, 1954).
- **Hicks Law** relates to average reaction times of individuals relative to several given choices. The argument being, that the fastest reaction time is achievable when an individual is given just one choice. For every choice added, the reaction time decreases. This theory in terms of HCI and graphical user interfaces often relates to lists. There is an exception to this rule, that is that if the listed items are listed in sequence, for instance alphabetically, and if the choices are known to the user, the reaction times can be significantly increased (Hick, 1952).
- **Power Law of Practice** relates to the rate at which an individual learns to accomplish a task. The Power Law means that the learning is not linear and is, instead, a learning curve. For an individual with no understanding of a task, a high proportion of the required knowledge to complete that task can be acquired in initial attempts. However, as the task is practised further, the rate of learning decreases, and mastery requires increasing levels of engagement (Newell, 1980).
- **Gestalt Psychology** developed by Max Wertheimer, Wolfgang Kohler, and Kurt Koffka, relates to the way in which individuals visually interpret the world around them, and more importantly build understanding from collections of objects. One element of the Gestalt theory is the way in which individuals interpret additional form,

shape and understanding from a number of visual components. The concept of *Reification* relates to how the mind generates and constructs shape that is not there through a combination of forms. *Multistability* relates to mind switching and the interpretation of more than one form in any given single visual element. *Invariance*, the capability of the mind to recognise a form or shape, even if it is moved, rotated, scaled and rendered from differing prospective. An additional concept within Gestalt, is the understanding of the way humans perceive elements that are grouped, similar or in proximity to one another. Individuals can be influenced to understand relationships of components, relative to their placement.

Although these theories have been greatly popularised through design education and employed to a greater or lesser degree by designers, the impact is considered to be limited. Thomas Landauer (1991) argues: *'I think saving small fractions of a second by optimal button placement, is probably a good illustration of the real but limited impact that traditional psychological theory can have if diligently applied*' (p. 65).

2.3.2 Model Human Processor

The theories surrounding the Model Human Processor 'MHP', (Card et al, 1983) attempted to develop knowledge of capabilities and limitations of the user, in order to evaluate and predict the performance of new interactions. The theory is scientific in nature and aims to make quantitative predictions about user performance. The theory argues that the human has three interacting systems: *Perceptual*, *Cognitive* and *Motor*. Each of these systems is said to have its own memory and processor. To demonstrate how this model could be used to evaluate interactive systems, a set of predictive models was developed. These are collectively referred to as GOMS: *Goals, Operators, Methods* and *Selection Rules*. The resulting suite of methods gave usability engineers descriptive tools, which supported a cognitive engineering approach.

Despite demonstrating great potential, the application of the theory was found to be limited. The theory was criticised for focusing at too low a level, on only highly skilled users, with error free performance, and not accounting for individual differences, and the effects of fatigue and motivation (Olson, 1991). Phil Barnard (1991) critiques the application of basic psychological theory to HCI as follows:

> *The trouble is that the paradigms of basic psychological research, and the bridging representations used to develop and validate theory, typically involve unusually simple and often highly repetitive behavioural requirements, a-typical of those faced outside the laboratory* (p. 103).

2.3.3 Theory of Action

Donald Norman (1986) explained how a user interacts with the environment, and subsequently a computer, through the decision cycle model. It describes a model of interactivity, within which the user establishes a goal based on an understanding, or mental model of the problem and domain space. The user then formulates, and executes an action, before evaluating feedback to determine whether the intended goal has been achieved. If it has not, the user modifies the action, before running through the loop again, until the intended outcome is achieved.

Norman defines seven Stages of User Activities, and are liked as follows:

1 Establishing the goal
2 Forming the intention
3 Specifying the action sequence
4 Executing the action
5 Perceiving the system state
6 Interpreting the state
7 Evaluating the system state relative to the goals and intentions

Norman (1986) describes that: '*Real activity does not progress as a simple sequence of stages. Stages appear out of order, some may be skipped, some repeated. In some situations, the person is reactive, event or data driven - responding to events, as opposed to starting with goals and intention's* (p. 42). David Kirish (1997) revises Norman's decision cycle model (see: 2.3.3) arguing: '*It is essentially incomplete theory, for it says nothing about the dozens of actions that agents perform in their environments which are not concerned with goal achievement actions more connected with improvisation than planning'* (p. 86).

2.3.4 The Gulfs of Evecution and Evaluation

Norman (1986) argues that there is a gulf that needs to be bridged between the user's goals and the way a system works. Execution describes the distance between the user and the physical system. Evaluation describes the distance between the physical system and the user. By closing the gap of execution allowing the user to manipulate the system better, and by improving the means of evaluation of the potential affordances and system state, interaction between human and computer from a task based prospective can be improved.

The gulf of evaluation involves comparing an interpretation of a systems state with the intentions or goals originally defined by the user. This occurs in four segments: the output displays of the interface, the perceptual processing, the interpretation and finally the evaluation. Using a conversational metaphor, it is the systems capability to communicate with the user. The gulf of execution is the degree to which the interaction possibilities of a computer system correspond to the intentions of the user, and what that user perceives is possible to do with the artefact or application: *'The gulf of execution is the difference between the intentions of the users and what the system allows them to do or how well the system supports those actions'* (p. 38).

Further describing this concept, Norman defines the mechanisms from a psychological prospective:

> *There really are three different concepts to be considered: two mental, one physical. First, there is the conceptualisation of the system held by the designer. Second, there is the conceptual model constructed by the user. Third, there is the physical image of the system from which the users develop their conceptual models.* (p. 47).

Both of the conceptual models are what have been called *'mental models'* (p. 47).

2.3.5 Mental Models

Carroll (1998) describes mental models as follows:

> *A rich and elaborate structure, reflecting the users understanding of what the system contains, how it works, and why it works that way. It can be conceived as knowledge about the system sufficient to permit the user to try out actions mentally before choosing one to execute. A key feature of a mental model is that it can be run with trial, exploratory inputs and observed for its resultant behaviour'* (p. 51).

The theory of mental model posits that when a user is confronted with a task, they have an internalised understanding of the problem space, the task needing completion, and the tools and methods at their disposal to solve or complete the task. The individual will draw upon this internalised understanding to define a clear representation of the intended goal, and then plan and internally simulate a sequence of steps to undertake. Once considered and formulated the individual executes the sequence to achieve their goal.

The concept of mental models is attractive to designers, as it provides a clear objective to align the user's mental model of an interactive system, with the conceptual model of the designer. If parity can be achieved, good interaction would follow. If there were inconsistencies in the user's mental model, instruction might be devised to improve alignment, with the probability that the user might achieve their goals.

The theory of the mental model was introduced by psychologist Kenneth Craik in 1967, with his book entitled *The Nature of Explanation*. Craik describes a process by which external realities within the world are translated into internalised models of reality, or Mental Models inside the mind. Craik describes internal constructions of some aspect of the external world, that are manipulated, enabling predictions and inferences to be made (Craik, 1943). He believed that the models we construct of the world serve as resources through which individuals might reason, plan, and execute their behaviours.

Norman (1987) provides a definition:

> *In interacting with the environment, with others, and with artefacts of technology, people form internal, mental models of themselves and of the things with which they are interacting. These models provide predictive and explanatory power for understanding the interaction* (p. 7).

Philip Johnson-Laird (2001) defines that a mental model: *'represents entities and persons, events and processes, and the operations of complex systems'* (p. 187). He goes on to describe four principles that distinguish mental models from linguistic structures and semantic networks:

1 The Principal of Iconicity: A mental model has a structure that corresponds to the known structure that it represents.

2 The Principle of Possibilities: Each mental model represents a possibility.

3 The Principle of Truth: A mental model represents a true possibility, and it represents a clause in the premises only when the clause is true in the possibility.

4 The Principles of Strategic Variation: Given a class of problems, reasoners develop a variety of strategies from exploring manipulations of models (p. 187).

Laird argues that mental models allow individuals to think as perception and discourse allows one to build models of the world. Through deductive and inductive reasoning, through

collections of mental models, Laird describes how *'thinking, which enables us to anticipate the world and to choose a course of action, relies on internal manipulation of these mental models'* (p. 185). Robert Allen (1997) argues that the expectations that a user has about a computer's behaviour, are drawn from their mental model of it. He suggests that mental models are the way that people model processes and that: *'mental models as the combination of the representation and the mechanisms associated with those representations'* (p. 49).

Modern theories of HCI, recognise the limitations of purely internalised cognitive, or mental model theory, to examine, describe, evaluate and design computer interactions. And there is clearly a compelling argument, to suggest that user understanding and action, are driven by a combination of internal and external processes. Indeed, Rogers (2012) argues that:

> *The internal representations that are activated are used in conjunction with the many different forms of external representation, coupled with an array of physical and mental actions, including gesturing, projecting, talking, touching, manipulating, and imagining* (p. 28).

In early work Norman (1983) recognises that: *'Peoples' mental models are often incomplete, easily confused, based on inappropriate analogies and superstition'* (p. 8). This research is focused on investigating a Sovereign Boundary Mechanism, which is a component part of a broader decentralised ecosystem. This mechanism reflects the current trajectory of decentralised development, and, on investigation, suggests a model of interaction that is potentially solitary and internalised. Though modern HCI theory argues against internalised models, the notion that the current trajectory of decentralised technologies is reliant on internalisation, justifies a reappraisal of old theoretical positions if a richer understanding and broader critique of decentralised endeavours is to be made.

2.3.6 The Contemporary Value of Mental Models

Though mental models have fallen out of favour in contemporary discourse, there are still compelling arguments for their continued value. Stephen Payne (2003) argues: *'I contend that cognitive science still has a major, even central role to play. It seems to me that "traditional" user-interface issues, such as what makes an interface easy to learn and use, remain vital'* (p. 136).

The concept of the mental model has been criticised, because of the perception that it is merely a general picture of a system within the user's mind. It is argued by Payne that the concept is misunderstood and generalised: He comments: *'A casual inspection of the HCI and cognitive-science literature reveals that the term is used in so many different ways as to be almost void of any force beyond the general idea of users' knowledge about the system they use'* (Payne, 2003, p. 135). Payne argues that the mental model still holds value, and that the concept should be unpacked to understand better exactly what is meant by the term. As such, Payne offers six domains, through which the concept might be better understood.

- **Mental Content vs. Cognitive architecture: Mental Models as Theories.**Payne (2003) first introduces the concept of the mental model by contrasting it against the more mainstream views of cognitive psychology in HCI. The dominate approach to understanding bounded reality of interactions has been to explore the cognitive architecture of the mind, its capability's and its limitations, with the aim of developing a method for systematic building of interactive systems (Card et al, 1983). Payne argues that to little focus has been given to mental content, what the individual understands and the relationship between beliefs and reality and how that effects behaviour. He argues that singular behaviours driven by individual models of a domain are of critical importance, and that their analysis can lead to the exposure of incidents and the 'diagnoses of misleading features of an interface design' and that *'interesting generalisations'* can be made *'across the content-domains'* (p. 139).

- **Models vs. Methods: Mental Models as Problem Spaces.** This concept concerns the building of internalised mental models of machines, and then using that mental model as a problem space in which to explore new methods for achieving tasks. This allows the process of search to become routine and allows the user to progress from novice to expert overtime. Clearly there is much debate about whether users actively solely hold complete models of domain inside their minds (Kirsh, D, 1997). Indeed, the modern concepts off externalised cognition offer an alternative, but as decentralised trends move towards a internalised cognition, this concept is still considered relative, in either supporting or opposing design decisions in the decentralised domain.

- **Models vs. Descriptions: Mental Models as Homomorphisms.** This concept argues that an individual's mental model, is a special kind of representation. It is what is known as an analogue representation, meaning the model shares a structure with the world it represents. The model is described by Payne (2003) as a *'isomorphic representation of such situations with each object and relation in the represented world being represented by a token and relation in the model'* (p. 144). It is argued that we build analogue models and that these can be combined with others, allowing deductive inferences to be read off the newly combined representation.

- **Models of Representations: Mental Models Can Be Derived from Language, Perception, or Imagination.** This concept draws on work which has established that a mental model can be derived from language or through reading text. The theory suggests that a reader or listener, first constructs a representation of the words or text, and then builds a mental model in order to comprehend the meaning. It is suggested that mental models can also be constructed through interaction and imagination in the same manner.

- **Mental Representations of Representational Artefacts.** This concept considers the relationship between the structure of text and its meaning. The theory argues that we build both a representation of the text in terms of its structure and of its meaning, and that the two are linked. If we want to change our representation of the meaning of the text, we need to work through, and affect the representation of its structure. This is considered to be the same in all representational artefacts including computer systems.

The theory is known as the Yoked State Space (YSS) hypothesis. (Payne, Squibb, & Howes, 1990). The model is described by Payne as:

> *To use such artefacts, requires some representation of the domain of the application of the artefact, the concepts the artefact allows you to the domain of application of the artefact - the concepts the artefact allows you to represent and process. The users' goals are states in this domain, which is therefore called the goal space. However, states in the goal space cannot be manipulated directly. Instead, the user interacts with the artefact, and therefore he or she needs knowledge of the artefact and of the operations that allow states of the artefact to be transformed* (p. 147).

- **Mental Models as Computationally Equivalent to External Representations.** Payne's final ideas centre on the comparison of representations, and whether they are informationally equivalent. If the same information can be extracted from both representations and is computationally equivalent, it means the cost structure in extracting the information within representations is equivalent. Payne argues that this comparison can be reframed around the concept of tasks, meaning representations are informationally equivalent if they allow the same tasks to be accomplished, and if the relative difficultly of the tasks are equal.

Within this research, the concept of mental models is important. Though contemporary theory argues that cognition occurs not only in the mind, but as an interplay between internalised and externalised representations. The development of decentralised technologies and observations which suggest a current design trajectory that is reliant on internalised cognitive processes, warrants investigation of mental models as an independent method of cognition.

2.3.7 Shneiderman's Framework of Direct Manipulation

Direct manipulation is today a core concept of graphical user interface. The interaction allows for the physical moving and manipulation of representational objects on the screen combined with menu-based interaction. The advent of this kind of interface allowed more intuitive interactions and minimal learning. Ben Schneiderman (1983) observes a number of

examples of direct manipulation interfaces outside of the task-based office domain. Early video games, computer aided design, and driving a car, are described as intuitive, informative and easy to learn examples of a task domain and representation of reality being directly manipulated. He argues that interaction that provides continuous representation of the object of interest, physical actions instead of complex syntax, immediately visible incremental reversible operations and a layered and spiral learning approach can make for *'lively, enjoyable interactive systems that reduce learning time, speed performance and increase satisfaction'* (p. 68). The core principles of Direct Manipulation are not without criticism. It is argued that continuous representation can limit action as only a small number of objects can be seen at a given time. Physical action can lead to repetitive strain, continuous feedback is limited to available actions and rapid learning is only ever as good as the design of the interface itself (Sherugar, 2016). In addition, it is noted that Direct Manipulation can be slow. Respective tasks are not well supported, and that some gesture-based actions are more error prone than typing. Direct manipulation was an early concept of interaction that occurred through a combination of internal and external representations, and arguably preceded the later theories of external cognition.

2.3.8 A Move Away from Cognitive Psychology

'*The most sustained, focused, and sophisticated attempts to develop explicit extensions to academic psychology for HCI have had no impact on design practice'* (Carroll, 1991, p. 1).

Although cognitive psychology dominated academic thinking in early HCI theoretical development, there were those who began to question its capability to support the development of broad system design. Efforts to inform the design of computer systems and interface through detailed mapping of the human psyche, in all but the most constrained domain, with the most expert of users, where considered inadequate. Thomas Landauer (1991) sums up this notion with the statement: *'My contention is simply that the theory of human cognition is now and*

may forever be too weak to be the main engine driving HCI' (p. 61). Landauer goes on to claim that any design guidance that might be gleaned from the theories of the Model Human Processor (Card, et al, 1983) could be easily exceeded within a real-world design context, by applying crude empirical methods. For example, Jacob Nielsen's rules of thumb heuristics, or testing basic prototypes with a handful of participants (Nielsen 1989, Nielsen & Molich, 1990). The theories of cognitive psychology are considered inadequate in modelling the complexity of human interactions, incapable of interpreting pattern recognition, knowledge-based activities, decision processes, error, motivational drivers, learning and social issues. In the context of cognitive psychology, Phil Barnard (1991) states that: *'There are very few examples where substantive theory per se, has had a major and direct impact on design'* (p. 103). Barnard argues that the first cycle of HCI research, characterised by the theories of cognitive psychology, are *'too low level, they are of restricted scope, as abstractions from behaviour, they fail to deal with the real context of work, and they fail to accommodate fine details of implementations, and interactions that may crucially influence the use of a system'* (p. 112).

2.3.9 The Development of Modern Theories

As HCI theory evolved, ideas began to emerge that focused on the externalised interactivity of users, rather than what was perceived to be happening inside their minds. As Rogers (2009) explains: *'It was recognised that a more appropriate conceptualisation of cognition for HCI, was one that was distributed across people, technologies and the environment, and externalised'*. (p. 31) The central focus of this new lens, became the interplay between external and internal representations.

2.3.10 Ecological Psychology

James Gibson's *Ecological Psychology* argues, that perception itself is interactive. The motion of the head, body and eyes all act in a co-ordinated fashion to control the sampling of

the optic array. Animals can pick up information about the environment directly from their perception of its structure. The structure of the environment has meaning to the animal, and that this perception can guide their actions, without having to deduce cognitively the context of their surroundings. This phenomenon is known as Ecological Constraints. Furthermore, he argued that animals could sense the opportunities to act within the environment, the relationship between the properties of the animal and the perceptual property of an object. Gibson (1979) argues for affordances with the statement: *'The central question for the theory of affordances is not whether they exist and are real, but whether info is available in ambient light to perceive them'* (p. 140). Norman (1988) describes an affordance as giving the participant *'a clue'* as to what actions can be undertaken. The notion of affordance as an alternative to pure cognition is a powerful concept, one which has been embraced by the design community. William Gaver (1991) argues, that affordance can be sensed not only from visual stimulus but also through tactile interactions and through sound. He argues that we should apply the notion of affordance to design, and that interfaces can have affordance designed into them. He describes a framework that *'encourages us to consider devices, technologies and media in terms of the actions they make possible and obvious. It can guide us in designing artefacts which emphasise desired affordances and de-emphasise undesired ones'* (p. 5).

2.3.11 Kirishs Theory of Interactivity

Kirish developed a theory of interactivity, which argues a need to overhaul the decision cycle model (Norman, 1986), (see: 2.3.3) to overcome its limitations. Kirish (1997) argues that two central ideas are missing from Normans model. The first is that *'goals are often not fully formed in the agent's mind'* (p. 86) and that we use the environment to develop, form and shape our intentions, developing our understanding and goals as we interact with it: *'It is essentially an incomplete theory for it says nothing about the dozens of actions that agents perform in their environments, which are not concerned with goal achievement, actions more connected*

with improvisation than planning' (p. 87). In essence, Kirish argues that there is an element of exploration required within the environment, in order for participants to formulate the goals which are not predetermined. The second concerns the missing understanding that *'the environment we confront at each moment is a partial function of our own last action'* (p. 88), meaning we change and manipulate our environment as we work within it to support our understanding. This includes preparing, exploring and maintaining. He argues that interfaces should *'foster this type of coordination between improvisation and planning'* (p. 81) biasing electives within the environment to help the user to understand what to do next. In the context of Gibson's affordances, and the interface being sensitive to the user's goals, protracted reasoning and information processing go way beyond the perception-action paradigm: Kirish concludes:

> *My conclusion is that if dynamic interfaces are to support complex learning activities, they must not only offer the type of perceptual affordances and action effectiveness that Gibson described, they must also facilitate a range of actions for reshaping the environment in cognitively congenial ways (p.80).*

What Kirish is describing in this seminal work, is the notion of and interplay at the interface layer between, internalised and externalised cognition.

2.3.12 External Cognition

The notion of external cognition centres around the argument that: *'when individuals are solving problems, human beings use both internal representations stored in their brains, and external representations, recorded on a paper, on a blackboard, or on some other medium'*, (Larkin and Simons, 1987, p. 66). In the paper entitled *Why A Diagram is (Sometimes) Worth Ten Thousand Words*, Larkin and Simons (1987) compare diagrammatic and sentential representations of information and their relative problem-solving virtues. They argue that if both forms of representation are informationally equal, a well-formed

diagrammatic representation can offer a computational advantage. *'Diagrams can be better representations not because they contain more information, but because the indexing of this information can support extremely useful and efficient computational processes'* (p. 99). They argue that information search and the relationships between elements within a diagram is clearer, and that the visual representation makes understanding and problem solving simpler. Clark & Chalmers (1998) describe, *Active Externalism*, as a coupling between the internal mind and cognitive resources in the environment, an external component of a cognitive systems behavioural competence: *'The human organism is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right'* (p. 8). They argue that a barrier to the understanding of externalised cognition is the rejection of the concept that consciousness exists outside the boundary of skin and scull, before arguing that not all cognition is conscious. They proceed to argue that the environment plays a key role in the way that we think: *'Once we recognise that the crucial role of the environment in constraining the evolution and development of cognition, we see that extended cognition is a core cognitive process, not an add-on extra'* (p. 12).

Scaife and Rogers (1996), in their paper entitled *External Cognition*: *How Do Graphical Representations Work*, ask the question *'what is actually gained cognitively from having more explicit, dynamic and interactive representations of information?'* (p. 186). They argue for the concept of *Computational Offloading*, the way in which external representations change the amount of cognitive effort needed to carry out a task. Kirish and Maglio (1994) demonstrate this with the game Tetras, where expert players engage in *'epistemic action' (p. 513)* to rotate the pieces in the game, and not in the mind, in order to find a solution and best fit. Scaife and Rogers describe graphical constraining, arguing that external representations can be designed in such a way as to limit the possible inferences that can be made in completing a task, reducing the load on internal memory, allowing more space to plan the next move. There

is the notion of re-representation, a consideration of how different external representations with the same abstract structure, make the solving of problems easier or more difficult. They argue for the notion of cognitive tracing, that an external representation should be interactive, and that a user should be able to mark and annotate to aid understanding and to build external memory (Scaife and Rogers, 1996).

O'Malley and Draper (1992) offer an interesting argument through which to consider the interplay between internalised and externalised representations when interacting with systems. They argue that the internalised representation is not only knowledge of a systems function, but also knowledge of where to look, and how to find further information through the externalised representations. The argument goes someway to explain why users with poor overall system understanding can still function at a reasonable level. *'Information representation at the interface, or in other representational artefacts can be seen as external memory aids which fill the gaps in users internalised representations when interacting with a system'* (p. 85)

An analytic framework through which to understand external cognition is offered by Thomas Green (1989). He argues in the context of computer programming, and the generation and editing of notation, for a theory of *Cognitive Dimensions.* The central argument being that the activity of programming is a balance between notation and environment, and that the process is one of *Opportunistic*, or *Design is Redesign*. When developing a program or indeed any other type of HCI, the user changes and redesigns their offering as they progress. This can be interpreted as an externalised mechanism within the application environment. The notion of Cognitive Dimensions are the affordances or concepts within context, that the user may have at their disposal to enact change. Each of these have a degree of interactive friction associated. Green highlights a number of dimensions, *Hidden/Explicit* dependences, *Viscosity/Fluidity*, *Premature Commitment* and the recognition of *Hard Mental Operations*. These dimensions

describe a number of interactive concepts, but also provide the vocabulary to describe and debate. In summary, Green concludes that *'given a particular model of behaviour, such as opportunistic planning, we can state a good deal about the requirements that must be met by the system'* (p. 11). In the context of development of a Sovereign Boundary Mechanism, an exercise in describing the relative cognitive dimensions may prove to be beneficial.

2.3.13 Distributed Cognition

'Cognitive processes may be distributed across the members of a social group, cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure, and processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events' (Edwin Hutchins, 2000, p. 1).

Distributed Cognition (hence forth DCog) extends External Cognition through the inclusion of the social and time. It investigates the cognitive phenomena of an overall system interaction between multiple agents, internal and external representations, the artefacts utilised, and the environments in which they are situated. Mark Perry (2003) describes the goal of the application of DCog as: *'to describe how distributed units are coordinated by analysing the interactions between individuals, the representational media used, and the environment within which the activity takes place*' (p. 169).

The study of DCog is concerned with the way in which information flows through a cognitive system, which in one respect could be an individual interacting with a computational tool. It is more commonly small or large groups of people working together, within the same systems. Originally proposed by Edwin Hutchins (1995), DCog has been used to study the complex interactions of pilots within a plane's cockpit, and the navigational processes on a warship. In these instances, highly complex interactions were observed, involving deep internalised knowledge, complex instrumentation, computational tools, and specialised human

verbal and non-verbal communications. Within each of these systems, no individual is responsible for its function alone. Instead, individuals and artefacts channel data and cognitive process, in highly complex collective information processing, and problem solving. Rogers and Ellis (1992) explain that within the DCog framework:

> *Cognitive activities are viewed as computations which take place via the propagation of representational state across media. The media here refers to both internal (e.g. individual memories) and external representations (including both computer and paper-based displays). The states of the representations refer to how the various information and knowledge resources are transformed during the work activities* (p. 123).

A state of a representations refers to the way information and knowledge are transformed, and the way an individual's understanding changes through interaction with another person. In order to understand how representation states are coordinated during activity, focus is draw to the media through which the representational states flow:

> *These include talk, non-verbal communication, the transformation of information between different modes (e.g. verbal message to a keyed input), the switching between different modes of operation and the construction of a new representation by mental computation in combination with external representations* (Rogers and Ellis, 1992, p. 124).

Rogers and Ellis (1992) also draw attention to the importance of pathways of communication: *'as coordinated sequences of action that are continuously interrupted, by the demands of an ever-changing environment and the development of activities as they are adapted overtime, 'in response to new demands in the environment'* (p. 124).

Distributed cognition is both a theory and a practical framework, through which to analyse and better understand existing situations and working practices. In the context of this research, the theory is important as decentralised technology will ultimately require the collective cooperation between individuals, artefacts and systems. Understanding the ways in

which these mechanisms function, will aid in the conceptualisation and design of a Human-Centred Data Ecosystem.

2.3.14 Situated Action

Within an introductory essay for *Cognitive Science*, entitled *Cognition in the Head and in the World*, Donald Norman (1993), outlines the debate between those in the field, who argue for the traditional view on which cognitive science was founded, symbolic processing, or internal cognition, and those who are for what was then the new theory of *Situated Action* (SA). Norman describes the field of SA as *'emphasising the role of the environment, the context, the social and cultural setting, and the situations in which actors find themselves'* (p.1). Norman communicates the issues encountered when trying to understand human cognition from a symbolic prospective. The enormous amount of information needed to understand a world that has millions or even billions of symbolic representations. The complexity of the world and its unpredictability, the notion that as soon as you plan for something the world changes in front of you, and the issue of observing all the relevant aspects of human behaviour, which Norman describes as an: *'imperfect window into the complexity of the processing activities occurring inside the body'* (p. 3). Norman goes on to describe the raging debate within the HCI community:

"Behaviourist" cries the symbolic proponent, "you think that everything is controlled by the environment, independent of internal processing." "Disembodied intelligence" cry the situated action folks, "you tend to have a person lost in thought, planning all the future actions regardless of the fact that the situation in the world will change faster than thought can keep up, idealistic, over simple, and doom to failure (p. 4).

The theory of Situated Action describes the way in which technology is used by people in different contexts. This is often very different to the way in which a designer envisaged. The theory argues that individuals often approach a situation with a plan of action in mind, only to

amend and change it relative to the situation in which they find themselves. It is argued that designers should not assume a situation or dictate a way in which a system should be used, instead a designer should immerse themselves in as many differing participant contexts as possible.

Lucy Suchman (1987) describes how *'actions are always situated in particular social and physical circumstances'* and that the *'situation is crucial to an actions interpretation'* (p. 178). She continues to argue that this is so obvious that it can sometimes be overlooked. She argues that we should *'explore the relation of knowledge and action to the particular circumstances in which knowing and acting invariably occur'* (p. 179) and that:

> *Action is not adequately explained by either preconceived cognitive schema or institutionalized social norms. Rather, the organisation of situated action is an emergent property of moment-by-moment interactions between actors, and between actors and the environment of their action.* (p. 179)

The theory of Situated Action is relative to this research as any decentralised technology that offered a participant agency over personal data, will be used across a multitude of differing contexts by a wide range of participants. The spectrum of cognitive theory: Internalised, Externalised, Distributed through to Situated Action needs to become part of the discourse within the decentralised community, as it can be argued that from a pragmatic prospective, a full-scale Human-Centred Data Ecosystem will at some point, incorporate components and application of the full gamut of HCI cognitive theoretical positions.

2.4 Human Data Interaction

The following section explores the field of *Human Data Interaction*, a subfield of *Human Computer Interaction*. The literature relating to the field is reviewed, together with related topics and arguments.

2.4.1 The Emerging Field of HDI

The field of *Human Data Interaction* (henceforth HDI), is a subfield of *Human Computer Interaction* (henceforth HCI), as it concerns not only general interaction with computer systems, but predominantly human interaction with data. Richard Mortier (2013) describes the domain as follows: *'HDI concerns interactions generally between humans, datasets and analytics, but not the general study of interaction with computer systems'* (p. 1). HDI is related to the field of PIM's *'Personal Information Management'* (Jones, 2007), as it in part encompasses the way in which individuals manage and organise their personal information. It has parallels with the field of HII *'Human Information Interaction',* though this area is broader in scope as it endeavours to understand human interaction with data of every kind, both digital, physical and metaphorical (Jones, 2007; Fidel, 2015).

The First use of the term HDI can be found in the work of Niklas Elmqvist (2011), describing the concept of HDI as: *'Human manipulation, analysis and sense making of unstructured and complex datasets.'* (p. 104). In this work, the author argues that in order to make sense of complex datasets, embodied interactions should be developed, which enable a deeper emersion than those afforded by a mouse and computer monitor. Francesco Cafaro (2010) extends the field with the concept of whole-body interaction in which he explores *"Embodied Schemata"* and the prospect of exploration, and control of complex datasets using physical gesture (p. 560). His work presents, a study within a museum environment, to understand better human gesture, relative to observed works. Although the proposed process of *'Identification of metaphors, across input and output within HDI systems'* (p. 562) is not

directly related to this research, the concept of metaphor around understanding of data interactions is, and as such, is examined in this research.

Richard Mortier expands the field further, past exploration and sense making, to focus on data relating to the individual, encompassing the interactions between agents that gather, exchange and process, ever-increasing amounts of personal data. He claims that the field needs to draw on related disciplines including computer science, statistics, sociology, psychology and behavioural economics, in order to achieve its objective of developing concepts and arguments, to ensure that the individual human remains the first consideration of the data-driven society. Moritier (2015) proposes, *'placing the human at the centre of the flows of data, providing mechanisms for citizens to interact with these systems and data explicitly: Human-Data Interaction (HDI)'* (p. 1). This extends the positions of Elmqvist and Cafaro, through the argument that HDI needs to consider the individual's ability to:

> *Permit or deny access by third parties to personal data and to enable an understanding of the background and ambient ways in which data is used. There is the need to be able to interact with systems that process and draw inferences from data. And the need to consider data as dynamic, ever generating further data through a process of algorithms* (Mortier, 2015, p. 3).

All of these concepts refer to data outside of embodied interaction and focus instead on data that is defined as being personal in nature. This can be defined as: *'data consciously created by individuals, observable, recorded data concerning individuals, and data that is inferred and created about us by others'* (World Economic Forum and Bain & Company, 2011). At this early stage, Mortier would seem to be pointing towards a richer more transparent understanding of the data ecosystem, related to the individual, and the capability to exercise a degree of agency over it. HDI is illustrated by Richard Mortier (Fig. 5) as a means of disrupting and providing transparency for, the existing feed-back-loop, which is driven by personal data.

The data is analysed, inferences are drawn, which may then facilitate further actions. The data flows back into the loop, from where another opaque cycle can begin.

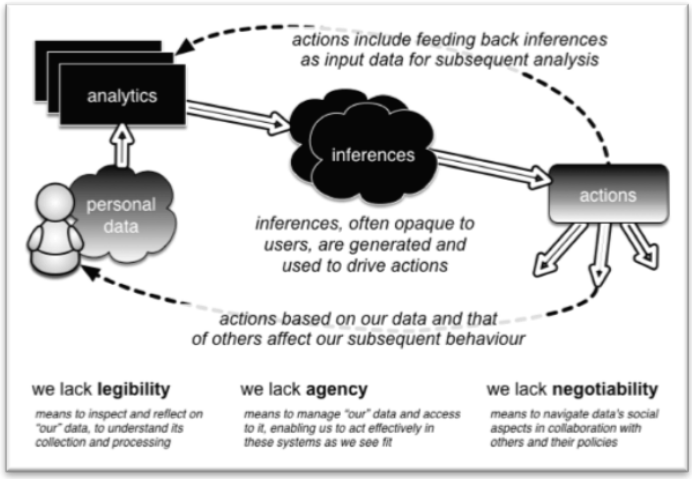


Figure 5: Mortiers Cycle of Human Data Interaction (Mortier, 2015, p. 4)

Richard Mortier (2015) suggests that in order to establish the individual at the centre of personal data flow, and disrupt the existing data model, three areas of communication and interaction are required: Legibility, Agency and Negotiability. Legibility is concerned with *'making data and analytics algorithms both transparent and comprehensible'*. Agency is: *'giving people the capacity to act within these data systems',* and Negotiability is *'concerned with the many dynamic relationships that arise around data and data processing'* (p. 5).

Heiko Hornung (2015) acknowledges the work of Elmqvist and Cafaro, and builds upon the work of Mortier et al, suggesting their concepts of interaction need to go further, in order to achieve a Human-Centred Data Ecosystem. He argues that data production, collection, editing, and use need to be investigated systematically in the context of social impact. Within this work it is suggested that the definition of data should be expanded to include not only personal data but also more generally, *'Data that effects people'* (p. 42). This progresses to suggest that the data lifecycle is not sequential or circular, but arbitrary, for instance when data sets are split or merged. The final working definition proposed is: *'Artefact-mediated representations of phenomena that need to be given meaning by people and that serve some*

purpose' (p. 43). The work goes on to suggest the concepts of Legibility, Agency and Negotiability proposed by Mortier, are insufficient to allow meaningful interaction with data, and that the understanding of data, should be complemented by the consequences of the data life cycle. As such, the main goal of HDI is framed by Hornung as: *'to design human data interactions, that enable stakeholders to promote desired and avoid undesirable consequences of data use'* (p. 41). It is proposed that in order to enact agency, the consequences of data use need to be considered in context, and drawn against the *'Beliefs, Values and Norms'* of the people involved (p. 39). Hornung sums up by suggesting that instead of adding to the topics defined by Mortier, we should adopt a more general theme in HDI, of *'understanding data and the consequences'* (p. 41). This position would seem to be relevant to this research. As previous arguments in the privacy literature suggest any privacy capability afforded to an individual, should be balanced against the overall societal gain it supports (see: 2.2.1). The suggestion by Hornung, that HDI should incorporate interactions which communicate the consequences of the data cycle, together with functionality to allow control, demonstrates alignment and may point towards a value proposition centred around societal benefit.

Haddadi (2015) takes the HDI field closer to the realms of the decentralised community, with the proposal for a Databox, a physical manifestation of the theories previously outlined. *'We propose there is a need for a technical platform enabling people to engage with the collection, management and consumption of personal data'* (p. 1). This work draws motivations from, and comments upon, locked-in and silo service providers stifling a truly competitive market, risks from personal data leakage, the opaque nature of data inference and the trade in personal data. Haddadi states that the concept of a Databox is not to replace established service providers, nor is it solely to focus on privacy issues. Instead it is focused towards enabling *'new applications able to combine data from many silos to draw inferences unavailable in the existing marketplace'* (p. 2). In addition, the system will provide the

components of Legibility, Agency and Negotiability, and go *'some way to redress the highly asymmetrical power relationships that pertain currently in the personal data ecosystems'* (p. 2). The work outlines what functionality the Databox should include: A trusted platform potentially under the control of the user, facilitation of data management, controlled access, and supporting incentives which may be the capability to receive compensation for the release of personal data. The work also comments on the fact that there have been several attempts to achieve such a system, and that these systems have failed to reach critical mass in terms of adoption. *'We believe that this is because there are fundamental barriers, technical and social, that have yet to be successfully addressed'* (p. 4). These issues are broadly related to availability, complexity, usability, and cost. These observations are pertinent, as the issue of adoption and barriers to participation is central to this research.

As Haddadi (2015) argues that the concept of a Databox is not a new idea, and outside the academic domain, a great many projects have been developed to address the asymmetric power relationships experienced by individuals in the context of personal data. *The Cluetrain Manifesto* (Levine, 1999) and *The Intent Economy* (Searls, 2012) have acted as a catalyst for a variety of projects related to business and the emancipation of the consumer. Their argument is that a free customer is better than a locked-in customer, and that the digital economy is operating on out-dated industrial principles, unsuitable for today's digital landscape. Freeing the customer offers huge opportunities and makes our fragile digital economy more sustainable. A further analysis of these concepts can be found in a later section of this thesis (see: 2.4.8). This work has developed into Project VRM 'Vender Relationship Management' at Harvard University (Project VRM, 2017). As part of this project a library of related product concepts and technologies is maintained. Many of these artefacts have contributed to the catalogue formed as part of this research. A similar catalyst for open source development of

decentralised technologies came after Eben Moglen's (2010) *Freedom in the Cloud* speech. In this speech Moglen argued the Internet was fundamentally engineered to exploit the user with a client server model, and that we are throwing away the privacy of future generations. The speech was powerful and very motivational. It communicates the current situation, the problems that we face, the opportunities we have lost, the damage we are suffering, and is a call to arms to change our direction. Eben Moglen founded the *Freedombox Foundation*[16], to develop open source software and hardware that has achieved in part what Haddadi (2015) and Mortier (2015) define as HDI. This gives the user complete agency over personal data and dissemination, an open source operating system running decentralised applications on physical hardware located within the participants home. The fact that data is held entirely within the participant's property, grants substantial privacy law advantages. This is a powerful concept, though it does define an individualist, insular position, and it is difficult to see how one engages with the wider world or finds balance between privacy and greater social responsibility. In essence the project to date, is a decentralised reinvention of services and tools that can be found within the centralised sphere. It lacks the notion of identity and fails to provide mechanism for meaningful social engagement. It has been argued that this type of isolated decentralisation is likely to fail and offer little value to the potential participant. To illustrate another product that achieves a decentralised offering of centralised services, we can turn to ArcOS (2016). ArcOS, developed by Jacob Cook. It is a fully functional product that allows the user to manage personal data and engage with a variety of personal and social applications. The system is hosted on a home server running on a *Raspberry Pi*[17]. This product is more than a prototype. It is functional and stable, but again it can be argued that this is a decentralised version of existing centralised service. It lacks a defined value proposition, and the effort and risk in engaging this model for the average individual, outweigh the perceived personal and social benefit. Mydex (Heath, 2013) is an example of a product that does offer a defined value

16: *Freedom Box Foundation* founded by Eban Moglen in 2010 is an organisation to fund and support the development of decentralised technology. https://freedomboxfoundation.org/

17: A *Raspberry Pi* in small inexpensive microcomputer. https://www.raspberrypi.org/

proposition for both organisation and individual. The system allows the development of trusted relationships through federated identity. Its objective is to make transactions seamless. It promotes efficiencies as it allows personal datasets and proof of claims to be reused. It is established as a social enterprise and is duty bound to serve the community. The proposition is very strong, though many of the concepts proposed highlight potential weaknesses, as it requires significant buy in, and commitment on both sides of the transactional equation. It can be argued that the model, engages new abstract concepts that currently lack understandable mental model or metaphor (Mydex, 2017).

There are many functional examples of what are considered decentralised Web 3.0 applications and technologies in existence. These include Personal Data Stores, Encrypted Decentralised and Distributed Data Storage, Federated Identity Systems, Semantic Based Social Networking, Semantic Otology's, Semantic Search Technology's, Encrypted Communications Systems, Crypto Currencies, Smart Contracting, Experimental AI Driven Digital Assistants, and Intent Casting Solutions. (see: 5.1). Despite a multitude of well developed, open source technologies, with the exception of Blockchain based Cryptocurrencies and Bit-torrent, it can be argued that the majority of these decentralised technologies have failed to gain significant traction.

To extend further the HDI domain, in response to limited adoption of technologies, it can be suggested that there are two additional pragmatic considerations that need to be included in the HDI academic discourse. This research is focused on exploring user interaction and value proposition. The following paragraphs briefly argue for the inclusion of these components.

2.4.2 Value Proposition

If the HDI principle of placing the individual at the centre of personal data streams is to be realised, it will have to break the dominance of well-established service providers. Simply to build decentralised versions of established digital services is not enough. Van Kleek (2014) argues that: *'*

> *While a number of factors are likely responsible, so the lack of interest among users has been attributed to the fact that many of initial PDS platforms have sought to simply re-create existing end-user experiences offered by popular apps and Web platforms, rather than creating new functionality* (p. 15).

A primary focus needs to be the investigation of the value of decentralisation: What can it do better, smarter, faster than its centralised equivalent? And how does this value manifest within practical applications. According to Philip Sheldrake, *'Why would anybody leave Facebook to join a distributed social network, why would they rip themselves away from the convenience the comfort, the connection of Facebook to go and be by themselves over here on a decentralised version'* (P Sheldrake 2016, Personal Communication, 19th October). Whether a decentralised service offers something engaging through its very architecture, such as a marketplace for personal data, benefits to society, democracy, education or financial incentive, there has to be something to engage users. Within the field of HDI, Value Proposition needs to be placed first and foremost in the minds of all concerned.

2.4.3 Usability

In response to a question regarding the main barriers to adoption of decentralised technologies, Jeremy Rushton responds as follows:

> *It's usability, all decentralised systems are hideous to use, and the reason for that hideousness, is really fundamentally about what it means to be a decentralised system, which in some sense is nodes opting to communicate with one another, that's a complex*

> *handshake, it's an inherently complex problem* (J Rushton 2016, Personal Communication, 27th Oct).

In the same vein as Value Proposition, the prospect of HDI and decentralised services need to consider the complexity of what is being suggested. As highlighted later in this document, personal data is complex (see: 2.4.11). The prospect of a user becoming aware of the arguments for HDI, understanding the streams of data, the ways in which it might be managed, the tools for managing, and the engagement with those tools consistently over time, is a formidable user experience, and user interface challenge. The academic and technological community need to step out of their clear understanding of the subject area from a technical perspective, and step into a user centric design domain, and into the mind-set of the common participant, to understand the significant challenges facing the realisation of usable, accessible decentralised tools and services.

2.4.4 Diffusion of Innovation

If a real-world manifestation of HDI is to be realised, adoption is a key consideration that needs to be explored alongside, value proposition, and usability. To understand adoption in a broader context and to deliver any HDI offering in the correct sequence, the Diffusion of Innovation offers a valuable framework. According to Everett Rogers (2012), the adoption of a new product, service or technology happens in stages, known as the *The Innovation Decision Process*. These stages consist of an individual becoming aware of an innovation, forming an attitude towards it, making a decision to adopt or reject, implementation of a product, and re-enforcement of a correct choice.

Rogers defines these stages as follows:

1. **Knowledge:** when the individual is exposed to the innovation's existence and gains an understanding of how it functions.
2. **Persuasion:** the forming of a positive or negative attitude.
3. **Decision:** when an individual engages in activities that lead to a choice to adopt or reject.
4. **Implementation**: when a user commits and begins to use a product or service.
5. **Confirmation:** the user seeks reassurance about a decision to adopt and may reverse that decision if exposed to conflicting messages.

It can be argued that within most decentralised technologies developed up to this point, there are at least three clear areas that challenge adoption within Rogers theory: *Understanding*, *Value*, and *Function.* For an individual at step one in the process, understanding this complex ecosystem of personal information and proposed new paradigms is a challenge. If we move past reinventing existing offerings, the concept of a Human-Centred Data Ecosystem, the communication of privacy infringement, asymmetric relationships in personal data, and the prospect of agency is complex. If an adoptee is to form a positive attitude towards adoption at step one, the *Knowledge Stage*, a clear message needs to be properly considered and communicated. At step two, the *Persuasion Stage*, if the theory of *Discontinuance* is to be triggered, a user has to form a positive opinion at this point. The value proposition needs to be front and centre, whether this is sold as decentralised or hidden from view, a clear advantage needs to be obvious to the adoptee. If it is not, the process will end. If the adoptee is between stages two and four and is actively engaging with a new decentralised offering, there is potentially an issue with the general complexity, usability and interaction. Understanding the methods and the functionality of HCI goes beyond our current understanding of PIM's (Jones, 2007) and *'How People Organise Their Desktops'* (Malone, 1985). The notion of data streams, data types, data management, system state, third parties, access control, identity management and contract negotiation, are arguably beyond our established metaphors and social norms in

this context. If individuals are to adopt these technologies, consideration needs to be given to this obstacle, breaking it down systematically and addressing it within a considered design process.

There are additional arguments to consider within adoption theory laid down by Rogers, one of which is *Re-invention*. The argument is that if an innovation is flexible to some degree of customisation, the rate of adoption is usually faster. This is very relevant to HDI, as any interface layer would need to be customisable, in order to handle differing data streams and situations, user preference and levels of competence.

Another theory is *Discontinuance*, due to an innovation being replaced by something better or because an individual has become disenchanted with an existing offering. The argument is therefore that if a new decentralised technology is to succeed, it cannot simply replace existing offerings. It must do something better than the centralised counterpart if adoption is to be achieved.

Channels of communication also form an important concept. The argument is that mass marketing is key at the *Knowledge Stage* of The Innovation Decision Process, but that this is replaced by a personal or one to one exchange of ideas instruction at the Persuasion and Decision-making stages. The development of any new decentralised offering needs to consider communication of ideas at a broadcast social media level, but also the micro digital communication and support level. Facilitating a balance of clear communication channels across the adoption process is critical.

The graph below (Fig. 6) illustrates the progression of adoption in graph form, with the innovators and early adopters of technology happening at the start of the S-Curve, before the early majority of user adoption as the curve progresses.

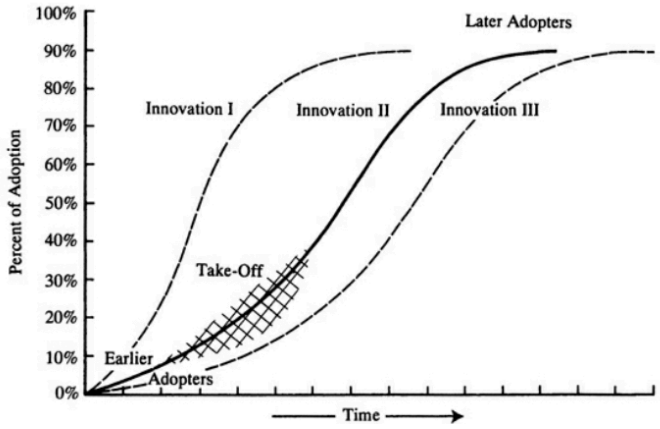


Figure 6: Diffusion of Innovation (Rogers, 2012, p. 11).

2.4.5 Technology Lifecycle Theory

Technology Lifecycle Theory defined by Geoffrey Moore (1991) is a derivative of the adopter categorisation developed by Everett Rogers (1958). The difference in the model are gaps that appear between stages. These gaps in the curve are argued to apply to technology's that are disruptive, and may not be in step with accepted processes, values and participant understanding. This is important as decentralized technologies potentially change the paradigm and these gaps may materialize in this context, posing a problem for adoption. The first of these *'cracks in the bell curve'* are described by Moore as appearing: *'Between the innovators and the early adopters. It is a gap that occurs when a hot technology product cannot be readily translated into a major new benefit. The enthusiast loves it for its architecture, but nobody else can even figure out how to start using it'* (p. 14). The second is described as a *'Chasm*' and Moore explains that *'when a product reaches this point in the market development, it must be made increasingly easier to adopt in order to continue being successful. If this does not occur, the transition to the late majority may well stall or never happen'* (p. 14).

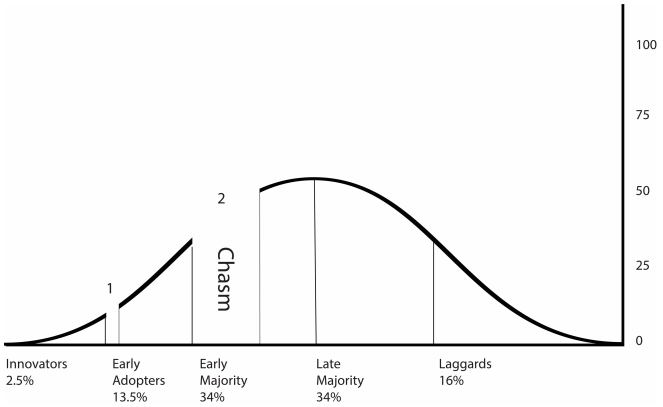


Figure 7: Technology Lifecycle Theory

Moore defines several steps which need to be undertaken in order to Cross the Chasm.

- **Target the Point of Attack:** This step refers to the identification and focus on a specific market niche
- **Assemble an Invasion Force:** This refers to the creating of the whole product, recognizing the problem faced by a participant and providing everything necessary to solve the problem.
- **Define the Battle:** The identification of the competition, the development of a competitive claim, the formulation of the communication of that claim, and the capability to demonstrate its validity.
- **Launch the Invasion:** In the context of traditional sales of technology or product this relates to distribution and pricing. Moore advocates a direct sales approach with a central consultative figure supported by application and technology specialists.

Moore's approach is very much focused on the traditional technology business and not the decentralised Internet. However, there does seem to be several stages and approaches which may be reinterpreted and reframed, in this context to address the potential chasm in the adoption curve decentralised technologies and indeed a wider ecosystem are likely to face.

2.4.6 The Promise of Big Data

Wendy Hall (2016) commented on the value of personal data with the following statement: *'When I say value, I don't simply mean the nation of individuals being able to sell their data for monetary gain. I am talking about how vital the sharing of personal data is in technological, and specifically digital, innovation'* (p. 3).

The term *Big Data* does not solely refer to a vast quantity of data which cannot be processed or made sense of, but rather, to a vast collection of valuable information, that offers great potential to a spectrum of society. Alex Pentland argues that Big Data offers huge opportunities, as it promises to reveal the underlying mechanisms of the world in real-time. We are only just beginning to understand through data science, the potential innovations and benefits to society that this rich knowledge resource can offer. Pentland argues: *'I believe that the power of Big Data, is that it is information about peoples' behaviours, instead of information about their beliefs'* (Pentland, 2012). Planning, health, business, security and personal interactions with the world, can be revolutionised as we move from knowledge based on averages and statistics, to real-time, real-world data at a micro level: *'With Big Data, we can begin to actually look at the details of social interaction, and how those play out and are no longer limited to averages like market indices or election results. This is an astounding change'* (Pentland, 2012). Pentland goes on to argue, that this prospect will only become a reality if people are willing to release their personal data, freely, confidentially, and on their own terms. Without this agency and trust, we run the risk of stifling, restricting or losing altogether this promising capability. Another important issue affecting Big Data, is that since this data is mostly about people, there are enormous issues of privacy, data ownership, and data control. *'We can imagine using Big Data to make a world that is incredibly Big Brother. George Orwell was not nearly creative enough when he wrote 1984'* (Pentland, 2012).

Richards and King (2013) have highlighted three important *Paradoxes of Big Data* and called for a pause and period of reflection to consider their potential impact. That pause would seem unlikely, given the commercial drivers, though the realisation of HDI in some form, might address the concerns and limit the potential harms.

- **The Transparency Paradox** highlights the issue that one of the core claims is that Big Data will create a more transparent and open society. However, to date, the gathering and application of the big data model is shrouded in mystery. Richards argues: *'Big Data promises to use this data to make the world more transparent, but its collection is invisible, and its tools and techniques are opaque, shrouded by layers of physical, legal, and technical privacy by design. If Big Data spells the end of privacy, then why is the Big Data revolution occurring mostly in secret?*' (Richards, 2013, p.42).

- **The Identity Paradox** highlights the issues of our desire to retain sovereignty over our personal identity, but that the current model of Big Data collection threatens identity. Richards argues that even the most basic access to a combination of Big Data pools, can result in; '*I am and I like, becoming, you are, and you will like'* (Richards, 2013, p.43). This argument goes further. It highlights the issue of Big Data feedback loop and echo chamber, that tailors search results, locking an individual into an artificial cycle of knowledge.

- **The Power Paradox** highlights the issue that Big Data aims to empower the individual or the social group. However, the power and the data are gathered, held, processed and utilised by a small number of very powerful organisations. Richards argues, *'If we do not build privacy, transparency, autonomy, and identity protection into big data from the outset, the Power Paradox will diminish Big Data's lofty ambitions. We need a healthier balance of power between those who generate the data and those who make inferences and decisions based on it, so that one doesn't come to unduly revolt or control the other.'* (Richards, 2013, p. 45).

Max Van Kleek (2014) echoes the view of other academics that the risk of losing the prospect of Big Data and personal data flow, jeopardises the development of the next generation of Web 3.0 technologies. Max argues that:

> *This misalignment of incentives between what users want to do with their data, and what platform providers want to do with their data, has the potential to interfere destructively with the development of context sensitive applications, that promise more effective, personalised, behaviourally adaptive interactions, that rely on richer and more sensitive data models, due to either actual or perceived privacy risks entailed* (p. 3).

The prospect of a world being informed by Big Data driven technology, seems to be becoming less of a possibility under a centralised model. We seem to be moving quickly to a point of mistrust, and duopoly where users are beginning to block the disclosure and transmission of high-quality personal data. The issues around Big Data and the risk of losing this valuable resource, form a compelling argument for the justification of HDI theories, while generating powerful arguments about what may support a principle value proposition.

2.4.7 EU Law - General data Protection Regulation 'GDPR'

As highlighted within the privacy section of this thesis (see: 2.2.9), the European Union's GDPR regulation, came into force from May 2018. This legislation is seen by many, as not only a bolstering of privacy law in the EU, but also as a move by the EU, to loosen the strangle hold of the US Silicon Valley technology companies, while fostering technology innovation and business within the European Union. Brexit, adds a further dimension to this situation, and it remains to be seen what the UK will decide in terms of its own data protection regulations. It can be argued that GDPR might offer an opportunity for the UK relative to the US, as we might see the UK adopt laws which essentially establish the island as an offshore intermediary. Regardless of the outcome of Brexit, the GDPR has come into force in the UK, and it remains to be seen if it will be adopted in full or rewritten.

These new regulations threaten to have a considerable impact on established business models, and time will tell if they can be adapted to stay competitive within the new rules and regulations. The tightening of rules around buried consent, and contracts of adhesion are a major obstacle. The concept of meaningless compliance is well established, and how the GDPR enforces clear and reoccurring understanding of terms and conditions, and genuine consent of data usage has yet to emerge. The right to withdraw consent, and the right to be forgotten also potentially cause issues with current business models. The right to access one's personal data free of charge, will also be a difficult issue to deal with. The small charge businesses are currently allowed to levy, for copies of personal data, acts to stave off a floodgate of requests. Once data is made available free of charge, it may require significant investment or automation. Finally, the regulations regarding genuine data portability will also affect the service providers position. If this proves a genuine method of breaking the model of the locked in silo, it may offer a spectrum of new opportunities for innovation and new services. However, there is a view that what the GDPR requires may not be achievable. David Alexander argues that: *'the rules around transparency of use, the right to have a digital version of it, the nature of informed consent, and traditional response regulation, just won't work, the law is a gamble'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

It can be argued that GDPR may offer a genuine window of opportunity for real innovation, and that it could act as a catalyst for a new wave of privacy enhanced digital products and services in the EU. But as research relating to the Privacy Paradox (see: 2.2.8) has demonstrated, user complacency with regards control and distribution of personal data can be easily realised. Unless tools and services are in place to capitalise on the window GDPR presents, it may quickly close, as established technology companies adapt, and individuals become compliant, as they concede to any terms in order to establish their hardwired interactions with the network. As Haddadi (2015) argues: *'fundamentally, imposing constraints*

that ignore the interests of advertisers and analytics providers, in many cases the business models that drive "free" web services and mobile apps, are likely to fail' (p. 20).

2.4.8 Contracts of Adhesion

The large-scale enterprise, comprising mass production and mass distribution, of the 20th century gave birth to the *Contract of Adhesion*. This involved mass standardised contracts that empowered a dominant party with an agreement, where they were free to make changes whenever they wished, and where the submissive party took all the risk and responsibility on a take it or leave it basis. A Contract of Adhesion, Boilerplate or Standard Form Contract, is described by West's Encyclopaedia of American Law as: '*A type of contract, a legally binding agreement between two parties to do a certain thing, in which one side has all the bargaining power, and uses it to write the contract primarily to his or her advantage'* (West's Encyclopaedia of American Law, 2017). The majority of contracts entered into by Internet users, when they accept terms and conditions for web services are these types of contract. There is no personalisation or negotiation. Doc Searls (2015) argues that *'we've lived so long without freedom of contract that we've forgotten what it is and why it's good for everybody. Our heads are now so accustomed to shrink wrap, click wrap, and other one-sided agreements that we can hardly imagine anything else in the network world.*' (p. 51).

As we enter and continue to build an ever more personalised-data-driven-ecosystem, the place of Contracts of Adhesion becomes increasingly questionable. As we move towards *Self-Monitoring*, the *Internet of Things*, and real-time *Personal Biometric Data*, can the big business argument of *'It's the only thing that works'* really stand up? As Doc Searls (2012) puts it: *'Adhesion contracts-as-laws should be deeply offensive to everybody'* (p. 57). As we move towards a time when individuals begin to become offended, we need to start to see the benefits, or indeed the value propositions of a system which can offer personalised terms of contract for the gathering and subsequent use of personal data. The *Blockchain* and the advent of the *Smart*

Contract, offer what can be argued to be a corner stone of the decentralised model, a customisable digital contract between one or multiple parties that is established, recorded and executed in irrefutable code.

2.4.9 A Stifled Digital Economy

'Increasing lock-in and network externalities are preventing formation of a truly competitive market' (Haddadi, 2015, p. 1).

The publishing of the *Cluetrain Manifesto* by Rick Levine et al (1999), communicated to business the profound change the Internet would have on established markets, and mechanisms for doing business. It likens the advent of the Internet, and its ability to facilitate conversation within the market, to that of an ancient bazaar, Levine explains:

> *In sharp contrast to the alienation wrought by homogenized broadcast media, sterilised mass 'Culture', and the enforced anonymity of bureaucratic organisations, the Internet connected people to each other, and provided a space in which the humans voice would be rapidly rediscovered.* (p. 6).

The text argued that business had to adapt to this new reality of two-way conversation or die.

Doc Searls extended his own contribution to the *Cluetrain Manifesto*, with *The Intent Economy* (Searls, 2012). This text incorporates many ideas and concepts derived from the twice-yearly *Internet Identity Workshops* (IIW, 2019) founded by Searls, Kaliya, Hamilin and Windley in 2005, and Project VRM 'Vendor Relationship Management' started by Searls at *Berkman University* (ProjectVRM, 2019). A main argument in *The Intent Economy,* is that in order for Digital Commerce to reach its true potential, the customer must be freed from the silo of *Customer Relationship Management* and is instead emancipated from the current model of *Choice of Captor*. It is argued that the liberation and communication ability that the Internet brings, makes obsolete, or at least inefficient the industrial revolution type business model of mass production, mass marketing and mass media. That the notion of *Contract of Adhesion*, or

Adhesionism, where establishing lopsided contracts is the only option when dealing with large numbers of unknown customers and users, is out-dated. The current models of marketing through the amassing and secondary use of personal data, are unsustainable. It is argued that there are many opportunities, for those who can be first to market, or that empower the user to communicate their intent into the marketplace. We are beginning to see the breakdown of the existing models, and a growing awareness that we have built our digital economy on a foundation that is ethically questionable and unsustainable. In a recent interview Jeremy Rushton argued:

> *I think most of the stuff that Facebook and Google do today will be illegal in 15/20 years' time, and I think that the failure of regulation, politicians and everybody else I have ever worked with, are not understanding the implications of what has been right in front of their face for 20 years* (J Rushton 2016, Personal Communication, 27th Oct).

As individuals become more aware, and begin to employ privacy enhancing technologies, such as Ad and Cookie Blockers, VPN's and Tunnelling, the ability of marketers to gather good quality data and marketing intelligence diminishes. The advent of GDPR in the European Union, has the potential to disrupt the current practices, and it is argued that there needs to be a new approach that recovers the digital economy from a race to the bottom.

Doc Searls (2012) defines the concept of VRM through seven principles:

1. Provide tools for the individual to manage relationships with organisations

2. Make individuals the collection centres of their own data

3. Give individuals the ability to share data selectively

4. Give individuals the ability to control how others use their data

5. Give individuals the ability to assert their own terms of service

6. Give individuals means for expressing demand in the open market

7. Base relationship-managing tools on open standards, open API's and open code (p. 164)

There is clearly value proposition in the VRM theory for both user and business. It remains to be seen whether these ideas will be embraced or rejected by the digital economy. In a recent interview, Philip Sheldrake commented on the challenge of establishing VRM:

> *To corral that join interest, that join benefit, that societal benefit crystallising, getting everybody to contribute to make it happen because they will benefit, but they don't what to pay for it because they are just one of thousands, so why should they be the ones to pay for it so it's a challenge* (P Sheldrake 2016, Personal Communication, 19th October).

2.4.10 The Risk to Our Democracy

'They can target micro-groups of citizens for both votes and money, and appeal to each of them with a meticulously honed message, one that no one else is likely to see. Each one allows candidates to quietly sell multiple versions of themselves and its anyone's guess which version will show up for work after inauguration' (O'Neil, 2016, p.160).

Within any political campaign the objective is to appeal to as many voting groups as possible, spreading your policies widely, while being able to defend each of them in the public domain. If you can profile potential voters and connect with them away from the public sphere, without scrutiny, our model of democracy becomes jeopardised. Monbiot argues that:

> *Our model of democracy is based on public campaigning followed by private voting. These developments threaten to turn this upside down, so that voting intentions are pretty much publicly known but the arguments that influence them are made in secret, concealed from the wider world, where they might be contested* (Monbiot, 2017).

Indeed, one of the most powerful arguments for HDI is the risk posed to the democratic system. Data inference and pattern recognition offer the prospect of micro targeting of an individual's political persuasion, in a very narrow cast and unaccountable manner. Monbiot argues that: '*micro-targeted ad campaigns are by their nature private or narrowcast. They never reach outside their target audience. Thus, they can contain falsehoods or insinuations that are never challenged because they are never brought to light'* (Monbiot, 2017).

A common denominator in both the Brexit leave campaign, and the Trump accession to the White House is Cambridge Analytica (Cambridge Analytica, 2017). This company specialised in targeted campaign intelligence based on establishing psychological profiles through behavioural science and big data analysis. In a recent article entitled *The Data That Turned The World Upside-down* published by Swiss publication Das Magazine, (Grassegger, 2016) it was claimed that by using a profiling technique called: *'OCEAN, an acronym for Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism—we can make a relatively accurate assessment of the kind of person in front of us'* (Grassegger, 2016). Together with Big Data, and Ad Targeting, it generated intelligence that allowed micro targeting of the US voting public in the US 2016 presidential election. This method is said to be a version of that developed by Dr Michal Kosinski and is highlighted in the privacy section of this thesis (see: 2.2.7). The impact of Cambridge Analytica's influence, has been countered and unpicked by Martin Robbins, who disputes the claims based on the numbers presented. He argues that: *'There's no evidence of this voodoo marketing in action, and we have plenty of anecdotes pointing to less than stellar use of data by campaigns'* (Robbins, 2017). Leonid Bershidsky also points out his doubts of the claims made, based on his own experiences of the poorly targeted messages he received during the campaign. Both counter arguments claim that Cambridge Analytica's capabilities have been over-hyped, and that their involvement and media coverage, has more to do with the members of its board, then its actual ability. Whatever the depth of influence, it demonstrates a trajectory that may not be desirable, and that threatens to undermine our democratic systems. As Mondiot explains: *'the Cambridge Analytica story gives us a glimpse of a possible dystopian future, especially in the US, where data protection is weak'* (Mondiot, 2017).

The surveillance, classification and monitoring of individuals and groups to politically profile is nothing new. However, the advent of Big Data analytics allows mass surveillance

and the inference of everybody who engages with the network. The arrival of this capability potentially removes the privacy component that allows democracy to function, allowing shaded micro targeting of a multitude of political messages. It must also be considered that the Cambridge Analytica story involves a third party who received data second hand. Facebook however, who have a vastly larger reservoir of real time data, are arguably one of the world leaders in data analytics. O'Neil (2016) questions *'by tweaking its algorithm and moulding the news we see, can Facebook game the political system?'* (p. 145). Facebook also has the capability to enact the Echo Chamber. A great proportion of current affairs and general news is now ingested by way of the Internet and through social media. The echo chamber metaphor suggests that news and ideas will be tailored for the individual, relative to a profile constructed from personal data. In essence they are telling the individual what they want to hear and reinforcing their expressed view without ever being exposed to the ideas and opinions of others.

It can be argued that the debate around personal data and its threat to democracy, is a significant value proposition in the development of HDI. As David Irvine argues:

> *We need to stop talking about privacy, because folk just don't get it, and they don't care, because they think I'm not a whistle blower, and what we need to start talking about is control. It's not about your privacy being taken off you it's about being controlled* (D Irvine 2016, Personal Communication, 9th Nov).

The message of democracy and control is powerful, and should be woven into the adoption strategy for the Decentralised Internet

2.4.11 The Complexity of Personal Data.

There are compelling arguments for empowering individuals with the agency to control their personal data. Before personal data can be controlled, we first need to understand what it actually is, and how an individual might make sense of it, and subsequently manage it. On investigation it soon becomes clear that this is complex. Hamed Haddadi (2015) explains '*as*

soon as one begins to examine the requirements for a Databox, one thing becomes very clear: data is a dangerous word. In particular, personal data is so complex, and rich that treating it homogeneously is almost always a mistake' (p. 3). The prospect of understanding all of these types of data, and how they might be managed, is an interesting and complex challenge. Van Kleek (2014) comments: *'the task of identifying all of the kinds of data a person might need to keep, manage and use is a complex and not easily scoped'* (p. 8). The afore mentioned field of PIM's, *Personal Information Management* (see: 2.4.1) which is closely related to HDI, *Human Data Interaction*, has much to offer in the classification of personal data and subsequent activities around its management. Below is the taxonomy of personal data outlined by William Jones (2010). The complexity is built from the relevance of the data and its importance, its dynamic or static nature, and whether the data relative to the subject is completely and partly within their control.

Categories of Personal Information - Jones's proposed taxonomy of personal information.

1. Owned/controlled by me	e.g., Email, files on our computers
2. About me	e.g., my credit/medical history, web history
3. Directed towards me	e.g., phone calls, drop ins, adverts, popups
4. Sent (provided) by me	e.g., Emails, tweets, published reports
5. Experienced by me	e.g., Pages, papers, articles I've read
6. Relevant (useful) to me	e.g., Somewhere "out there" is the perfect vacation, house, job, life- long mate.

(Jones, 2010).

Max Van Kleek (2014) goes further, in recognising that although PIMs provides good classification of data types, and a taxonomy for understanding these data categories, it falls short of considering the types of data that exist through inference, and that we aren't fully aware of: *'the categories About me, and Relevant to me, are controversial because these definitions do not require individuals to be aware of the existence of the information; it thus establishes a sphere that goes beyond the scope of information experienced by the user'*

(p. 8). Although PIMs offers a clear starting point in understanding the types of personal data, the dynamic nature of personal data generation, collection, inference and secondary use, in the development of HDI offerings means a broader taxonomy may be required. It can be argued that this should be considered in line with the work of Daniel Solove (2009) (see: 2.2.10) where he suggests we should try to: *'Understand privacy in specific contextual situations'* (p. 47). This will then provide a pragmatic starting point from which a definition of usable personal data types might be derived.

2.4.12 Quantified Self

Eun Choe (2014) defines *Quantified-Selfers* as an *'diverse group of live hackers, data analysts, computer scientists, early adopters, health enthusiasts, productivity gurus and patients, believing in the notion of self-knowledge through numbers'* (p. 1). The movement first started in Silicon Valley, becoming popular through the blog *Quantified.com* authored in 2007 by Wired Editors, Gary Wolf and Kevin Kelly (Quantified.com, 2017). Known in academic terms as *Personal Informatics* and *Personal Analytics*, the practice observes participants' tracking events and actions to build through patterns a better understanding of themselves, while looking for correlations and cause and effect. The practice of self-monitoring dates back to the 1970s, where individuals would monitor their behaviours, thoughts and feelings in order to uncover triggers and context for medical issues and as part of behavioural therapy. The concept of a Quantified Self is relative to this study, and the development of Human Data Interaction, as in essence. It is trying to accomplish what centralised data gathering achieves in the profiling of individuals. The concept aligns with the idea of the *API of Me* proposed by Greg Meyer in 2012. Meyer explains: '*The API of Me, is the name I'd like to propose for a system of capturing, sharing, and limiting information about consumers'* (Meyer, 2012).

Extending the data gathering capability and understanding of the quantified self, to encompass every digital trace, would allow for a rich collection of personal data and subsequent cross-referenced inferred data. It can then be compiled and shared for personal advantage. As such there will be much to learn from those who already capture, process and analyse their personal information. Deborah Estrin (2014) contributes to this area with the notion of Small Data, *n=me*. Within this work, she highlights the importance of Small Data, alongside Big Data, and argues the benefits of collecting personal medical data through technologies such as *Fitbit*[18], and the advantages this can bring to the individuals wellbeing. She argues that current service providers gather a large amount of data about us all but fail to return it back to the user. She argues that they should, because *'this broad but highly personalised data set can be analysed to draw powerful inferences about your health and wellbeing from your digital behaviour'* (p. 32). This is a powerful concept and might add weight to HDI value propositions formed around societal benefits.

2.4.13 A Marketplace for Personal Data

A report published in 2012, by *The Boston Consulting Group*, highlighted the huge current and future value that can be attributed to personal identity and personal data within the EU. They valued the combined total at 8% of the EU-27 GDP. They predicted this to be worth €330 billion annually to organisations, and €670 billion to consumers by 2020 (BCG, 2012). This did though come with one significant caveat. The report explained: *'However, two-thirds of potential value generation, €440 billion in 2020, is at risk if stakeholders fail to establish a trusted flow of data'* (BCG, 2012, p. 3). The report highlights the key areas in which personal identity and data can be employed. These include process automation, user enablement, personalisation, enhanced delivery, personal data driven R&D, and secondary monetisation. Sarah Spiekermann (2009) argues that: '*personal data markets thrive, and personal data is viewed as the oil, of the digital economy. Yet, ordinary people are barely aware of these*

18: A *Fitbit* is a monitoring device warn by the user which gathers data relating to fitness and wellbeing. https://www.fitbit.com/uk

developments' (p. 62). Schwartz (2009) argues that: *'Consumer ignorance leads to a data market in which one set of parties, does not even know that a negotiation is taking place*' (p. 2078). If we are to accept that there is a relationship between awareness and privacy concerns (Nissenbaum, 2004), the prospect of a growing understanding and new legislation by way of GDPR, risks the erosion of trust, the flow of data, with potential economic damage. As such it is argued that establishing a marketplace for personal data in some form is essential if our digital economy is to continue to thrive in the future. Also, *'If people take control over their data through new tools as well as legal enforcement, personal data markets will change and people will play an active role in them'* (Spiekermann, 2017, p. 63). This leads to the question: *Can a sustainable marketplace for personal data be established, to allow the trade in individual or group data sets, for either personal gains or for the gains of wider society?*

A direct trade in personal data holds value for business, as it offers the opportunity to obtain an identifiable, clean, static free, unbiased sample. Indeed Google's, *Option Reward*[19] already pay individuals to engage with questionnaires to gather information for financial reward. There is also a number of companies who invite individuals to trade access to the data streams for direct financial remuneration, or discounts and offers. These include: *Citizenme* (2017) a company that allows the sale of data for cash, but also provides access to personal data intelligence and the ability to donate data to charitable causes. *Datacoup* (2017) is another company allowing users to link service provider accounts for financial rewards. In their marketing materials it would appear that circa six active profiles generate the user up to one dollar per week. This would seem to be a very minimal reward for what can be very personal information, and it remains to be seen if individuals would consider this to be a balanced and fair exchange.

19: Google *Opinion Rewards* is a survey application and audience measurement tool. https://surveys.google.com/google-opinion-rewards/

2.5 Literature Review Summary

This chapter described the literature which defines the underlaying framework for this research. The investigation of Surveillance Capitalism presents the foundation theory and charts its progression through to the digital age. Arguments suggest that individuals and social groups are observed and categorised based on their economic value. It is argued that participants in the network are both producer and consumer of data, within a self-perpetuating cycle of consumption that generates market intelligence for capital. Indeed, alignment is drawn with Marx's theory, with the notion of Prosumer Proletariat in a digital Sphere of Production. Arguments are made that it is very difficult for the digital native to extract themselves from this mechanism, and that an everyday existence, now requires some degree of engagement with the network. The literature surrounding Surveillance Capitalism, provides a lens through which to view todays centralised Internet. It provides arguments to understand better decentralised endeavours and provides a viewpoint through which to consider value propositions.

The investigation of Network Privacy provides an insight into privacy as a concept, its problematic definition, and the challenge faced in trying to understand the essential or core characteristics of the overarching term. The review highlights the work of David Solove and his pragmatic method in defining privacy harms through a bottom up approach which identifies specific problems and solutions in contextual situations. The review considers privacy from an anthropomorphic position, its importance in our individual development, the wider functioning of our society and the cultural context. The review considers privacy in the digital realm, and the way its commoditised leading, arguably, to a change in the social norms. The review considers algorithmic surveillance, methods through which privacy can be infringed through inferred understanding, and the surrounding arguments and controversies such as Anonymised Data, Nothing to Hide, the Privacy Paradox, Instant Gratification Bias and the evolving

landscape of Privacy Law. Privacy is an important theoretical domain within this research, as a Decentralised Internet arguably centres around its protection.

Within the Human Computer Interaction review section, focus has been drawn to human engagement with technology from the prospective of cognition. The application of cognitive theory to HCI, provides a collection of lenses, through which to view the way human thought is coupled through interaction, to achieve the objectives of problem solving and task accomplishment. The progression from an internalised position, through a distributed model, to the broader concepts of Situated Action, and the framework of Distributed Cognition, provides an integral scaffold through which to consider this research. Decentralised interactions would appear to be complex in nature and understanding. How frictions can be identified and reduced is a central consideration, in evaluating current decentralised artefacts and considering how future interactions may manifest or be designed.

The theory of Human Data Interaction is central to this research. It represents a theoretical domain through which to understand, and to argue the principles of decentralisation. As noted within the introductory chapter, new knowledge developed through this research should be disseminated and published under this banner. HDI is a theoretical domain which recognises the value of personal data, the need to place the participant at the centre of data streams, and the requirements for transparency and agency. This review has investigated the origins of the term, its contemporary position and potential development into the future. The review identifies the significant topics surrounding HCI, investigates the arguments for, and obstacles hindering its adoption, while considering the value that might be realised, if HDI can be established through a Human-Centred Data Ecosystem. With a strong conceptual framework now in place, the next chapter of this thesis considers Self-Sovereign Identity. A direct practical attempt to address many of the issues raised in the literature, while satisfying the principles defined within the theory of Human Data Interaction.

Chapter 3: Self-Sovereign Identity

The following section presents the concepts and surrounding standards, collectively known as *Self-Sovereign Identity* (hereafter SSI). It can be argued that SSI, is a concept that endeavours to afford participants the same agency in the digital realm as individuals enjoy in the real world. SSI does this by enabling the individual considerable control over both their digital identity and personal data. SSI is not a single technology, rather a wider ecosystem comprising of a number of related components that collectively operate free from any centralised control, ultimately sovereign to the participant. At its core is an Identity Layer, a means by which an individual can control and manage an online presence, through which relationship, transaction and personal data can be managed and controlled.

At this point, it is important to communicate why SSI has been singled out as significant and why it forms a central part of this research. There are a great number of historical and contemporary innovations which can be considered decentralised. Technologies which allow an individual to store, control, manage, understand better and re-purpose personal data, in a myriad of differing contexts. But many of these concepts are not truly decentralised. An element of the system is still controlled, and this in general relates to a loaned or leased identity. Investigation has demonstrated that the current trajectory of the decentralised movement is towards SSI (see: 6.1.1), a truly decentralised model that fulfils the principles of a sovereign decentralised network engagement. This research aims to consider questions concerning individual agency within a Decentralised Internet, focusing on the edge of current developments in the form of SSI. This is the area most likely to inform the research questions, while generating meaningful new knowledge.

3.1 Stages of Internet Identity Evolution

An identity layer is at the core of current efforts to establish a Decentralised Internet. In his work *The Path to Self-Sovereign Identity*, Christopher Allen describes four phases of Identity evolution on the Internet (Allen, 2016). These phases are shown in figure 7 and are now described.

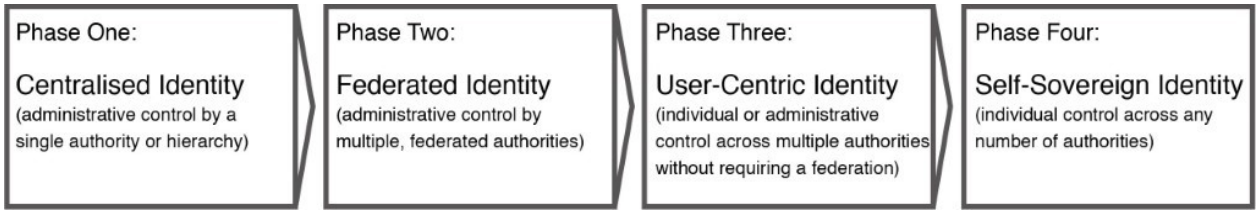


Figure 8: Allen's Stages of Internet Identity

Phase One - Centralised Identity refers to the most common form of identity on the Internet, where the identity is owned by a centralised entity and essentially loaned to the user. This is problematic in that the user has no control over the identifier, or the data associated with it. The user's access to the identifier can be suspended or terminated at any time, and users struggle to remember the details of the vast array of sign-ins they need to utilise.

Phase Two – Federated Identity refers to a model that allows a single identity to be used across multiple services. This is beneficial to the user as it acts as a single sign-in, and to organisations who can quickly allow users to access their services without the friction of registration. High assurance federated identities have emerged, were the verification of individuals is outsourced, and once assured, can be used for important online matters. Gov.UK *Verify*[20] is an example of this. Although federated Identity offers a degree of portability, it is still a centralised concept.

Phase Three – User Centric Identity refers to the capability for a user to decide what information and data associated with an identity, can be disclosed, redistributed or revoked when dealing with third parties. This model is a major step forward, as it allows an identity to be used across multiple entities with the added capability of access control. This goes some

20: Verify is a UK government run online identity system.
https://www.gov.uk/government/publications/introducing-govuk-verify

way towards decentralising network interactions, but at its core, it is still reliant on a centrally controlled identity.

Phase Four – Self-Sovereign Identity: Through the advent of a distributed ledger, together with cryptographic technique, the domain is now at the point where a true identity layer can be established, allowing a participant complete decentralised control over multiple identifiers, without the need for any centralised control. This emancipation acts as a solid foundation, from where a collection of additional concepts and standards can be established, to build a Human-Centred Data Ecosystem.

3.2 Self-Sovereign Identity

The issue of Identity has since its inception posed a problem for the Internet. The difficulty in authenticating an individual or organisation is problematic. Until now the solution has been third party loaned identity or certificates. As Peter Seiners cartoon, shown in figure 9 illustrates *'On the Internet nobody knows you're a dog'*, (Seiners, 1993), it is very difficult to prove you are even human. This has been eloquently expanded upon recently, with the statement '*On the Internet of Things, nobody knows you're a fridge'*. The issue of identity has until recently stifled efforts to decentralise the Internet, but this will potentially now change with the advent and full realisation of SSI.

Figure 9: Peter Seiners – 'On the Internet, nobody knows you're a dog'

SSI solves one of the most challenging issues of the Internet. The capability to establish, own and control a persistent sovereign identity. SSI can be defined as the following:

> *A Digital Identity that is owned and controlled by an individual, company or machine, it has no reliance on any centralised authority. The identity is persistent and can never be taken away from its owner. The identity is part of a wider ecosystem, where relationships can be built, trust can be developed, identity attributes, and data, can be exchanged under the complete control of the Sovereign Identity (Lockwood, 2019).*

The concept of a wider ecosystem is important, once a user can control an identity, the concepts of SSI then allows the user to establish independent unique relationships, and communication channels with peers on the network. It allows for the requesting, issuing and distribution of verifiable credentials. This in turn empowers the user to control their personal data and information in a transparent manner. The ecosystem allows for the development of trust networks that are judged appropriate dependent on the situation. The realisation of SSI is considered to be Web 3.0, and to many is inevitable, (Tobin et al, 2016). Web 3.0 represents the next evolution of the Internet, where the common participant is no longer reliant on a loaned or leased identity, owned and controlled by a third party or centralised authority. The participant is no longer bound by contracts of adhesion, data exploitation, or asymmetric power imbalances.

The genesis of SSI can be traced to the seminal work of Kim Cameron in his *The Laws of Identity* (Cameron, 2005). These principles of digital identity where then evolved further in the context of SSI in the work entitled *The Path to Self-Sovereign Identity* by Christopher Allen (Allen, 2016).

Allen defines that for an online identity to be self-sovereign, it should satisfy the following principles:

- **Existence** - Users must have an independent existence.
- **Control** - Users must control their identities.
- **Access** - Users must have access to their own data.
- **Transparency** - Systems and algorithms must be transparent.
- **Persistence** - Identities must be long-lived.
- **Portability** - Information and services about identity must be transportable.
- **Interoperability** - Identities should be as widely usable as possible.
- **Consent** - Users must agree to the use of their identity.
- **Minimisation** - Disclosure of claims must be minimized.
- **Protection** - The rights of users must be protected.

3.3 The Building Blocks for Self-Sovereign Identity

In the following section the fundamental building blocks of SSI are discussed in order to communicate the underlaying concepts, component parts and interactions required. Figure 10 provides a graphic to support this conversation, illustrating a balanced peer-to-peer relationship between two entities across the network, the exchange of data and credential supported by distributed ledger verification and the current developing standards. The components illustrated in this initial diagram, are explored in detail in the following paragraphs.

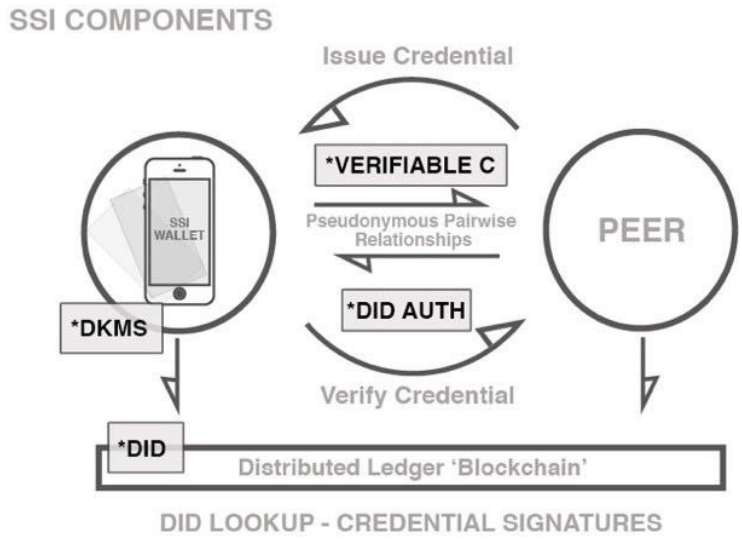


Figure 10: The Components of an SSI Interaction

3.3.1 Distributed Ledger

Recent developments in Self-Sovereign Identity have been made possible by the advent of Blockchain Technology, and its capability to provide a public, immutable, distributed ledger of transactions between peers. In most cases the infrastructure is designed as a public utility and acts as a decentralized foundation, supporting identity development, authentication, verifiable credentials, trust and data management (Wagner, 2018). The system uses encryption and digital signatures to establish a decentralised Public Key Infrastructure, through which

identity communications, credential, trust and authentication can be managed (Hyperledger, 2019).

3.3.2 DID's 'Decentralised Identifiers'

Decentralised Identifiers, (hereafter DIDs)., are digitally verifiable identity's that can be created by an individual, organisation or object. DIDs are fully under the control of the DID subject, and are independent from any centralised authority, identity provider, or certificate authority. A DID is a Key Value Pair. A DID address can be looked up on a Blockchain, which returns a DID document containing information about the DID. That information is composed of Proof Purposes, Verification Methods, A Public Key and Service End Points. Through the public key the identity holder can establish a private communication channel, prove they control the identity and register, resolve, update or revoke details for the DID on the hosting Blockchain. There is no limit to the number of DIDs that can be created and controlled, this capability enables *Privacy by Design* principles, preventing correlation and tracking. SSI promotes the principle of *Synonymous Pairwise Identifiers*, which means that for every peer to peer relationship across the network, a unique pair of DIDs are created, allowing authentication, communication and transaction. In the design of any Self-Sovereign-Identity system, creating, sharing and managing DIDs, is a primary sphere of interaction (DID, 2020).

3.3.3 Verifiable Credentials

We use Verifiable Credentials and subsequent claims in our everyday lives; a driving license, a passport or educational certificate are examples of Verifiable Credentials. These credentials can be used in a spectrum of contexts and combinations, to make claims about ourselves. A Verifiable Credential in the context of SSI, is a digitally signed document, that is produced by an issuer and accepted by a subject or holder. Verification signatures proving the integrity of the document, its issuer's identifier and the acceptance of the credential by the

holder are recorded on the Blockchain. The holder of the credential can then disclose the credential to another peer in full, or in part, and this third party or verifier can confirm the integrity of the credential by looking it up on the Blockchain, confirming origin, signature and integrity. This in turn allows identities to prove claims about themselves to develop webs of trust and engage in transaction. Figure 11 Illustrates the constituent parts of a Verifiable Credential, the stake holders, and the relationship to the underlaying ledger.

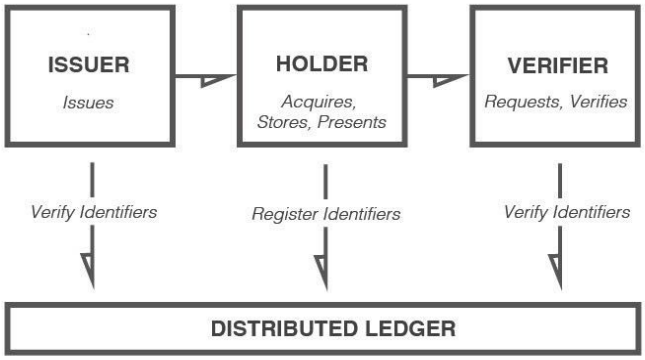


Figure 11: SSI Stakeholders within Verifiable Credentials

Credentials can be shared in part through the concept of *Zero Knowledge Proofs*[21], which allows the holder to make a claim about an identity without revealing the whole of the credential. Credentials or claims can be created and sent by any identity holder. This might be a government, organisation, company, individual or object. The level of trust required is dependent on the circumstance. Credentials and Claims can be combined in different formats, they can be static or dynamic in nature, and they can be rejected or revoked at any time. In the design of any Self-Sovereign Identity system, creating, sharing and managing Verifiable Credentials, derived claims and Zero Knowledge Proofs are primary spheres of interaction (VCWG, 2020)

21: *A Zero-Knowledge Proof* is a cryptographic method allowing one party (the prover) to prove to another party (the verifier) that they know a value, without conveying any information about that value.

3.3.4 DID-Auth 'Decentralised Identifier Authentication'

DID Auth is a method by which the owner of an identity, through a wallet or agent, can prove to a relying party that they are in control of a decentralised identifier 'DID'. The demonstration of control of the DID, is accomplished through a mechanism specified in the DID Document. Depending on context this authentication might be done while establishing a connection, as part of a transaction, or as part of a request. The notion of proving control over the private keys associated with any given DID, is an integral part of an SSI system, within a number of authentication interactions (DID-Auth, 2020).

3.3.5 DKMS 'Decentralised Key Management Structure'

SSI employs the power of distributed ledger technology, to provide security and immutability, scalable key distribution, verification methods and key recovery. This allows for a personal decentralised public key infrastructure, free from any centralised authority or control, and makes the benefit of a public key infrastructure available to everybody. As individuals engage in the practice of SSI, they will be generating a considerable number of encryption keys, that they will have to take responsibility for and manage. When a user manages their relationships, credentials and data through SSI, they will do so with a wallet on the edge of the network, this will typically be on a mobile device; most keys and credentials will be held on that device. The wallet is complemented by an agent, that will automatically manage many of the automated mechanisms needed to engage with SSI. For the identity to be available 24/7 and to also manage backup and key recovery, a limited representation of the edge level Wallet / Agent is available. Figure 12 Illustrates the structural parts of an SSI architecture, including edge and cloud-based wallet and agent and the underlaying ledger.

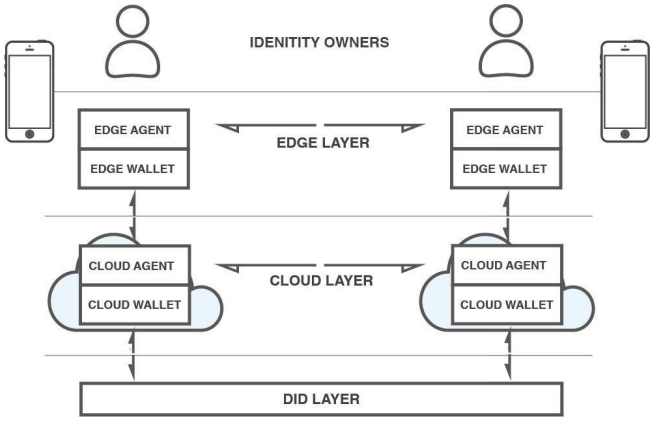


Figure 12: SSI Architecture

DKMS 'Decentralised Key Management System' is an emerging open standard, that deals with the interoperability between both agents and wallets, at both the cloud and edge layers. It is the protocol that stitches all other elements together and provides core SSI functionality and administration (DKMS, 2020).

3.4 SSI Summary

This chapter presented the concept of Self-Sovereign Identity, a collection of technologies supported by a Blockchain driven identity layer. The domain comprises of the component parts necessary for a Human-Centred Data Ecosystem. This innovation represents both the cutting edge and trajectory of the decentralised domain. The concept of SSI has been identified as central to this research, as its principles match those of Human Data Interaction and further investigation is considered relevant to the research question, and the generation of new knowledge. The core concepts, supporting technologies and standards, together with the principles of both SSI and HDI, have been extended, through a conceptual modelling method within the Practice Led Component of this research (see: 5.5).

Chapter 4: Methodology

This chapter outlines the underlying research philosophy and design, the research methods, means of analysis and the stages of data gathering. The final section considers ethical approval and the requirements for data protection.

This research engages four distinct, non-sequential components of primary data collection, in what can be described as a Mixed Methods approach, each component is summarised in Table 01.

Element	Type	Method	Analysis
Artefact Catalogue	Primary	Categorisation / Evaluation Matrix	Critical Analysis
Public Survey	Primary	Semi-Structured Interview	Thematic Content Analysis
Expert Interviews	Primary	Internet Mediated Survey	Descriptive Statistics / Correlation
Practice Lead Component	Primary	Conceptual Model / UI / Focus Group	Critical / Thematic Analysis

Table 1: Research Methods

Component 1 is the development of an Artefact Catalogue, comprising of historical and contemporary projects and technologies, critically analysed through an Evaluation Matrix. This component is designed to build an understanding of decentralised artefact development and trends.

Component 2 is a Public Survey, designed to gather data to ascertain attitudes towards data privacy, secondary use, disclosure and control. The survey results are analysed through both descriptive and statistical analysis, to provide valuable insights related not only to attitudes and behaviour, but also potential strategies for product development and adoption.

Component 3 comprises several phases of consultation with experts. The first considers the overarching decentralised space. The second considers the narrower view of self-sovereign technology and the third considers opinion not from experts within the decentralised field, but from practitioners within the usability and user experience sphere. The data gathered across these phases is analysed using *Thematic Content Analysis*, resulting in the presentation of themes, supported by related quotation.

Component 4 is a practice led element, designed to extend current Self-Sovereign Identity prototypes to scale, through *Conceptual Modelling*, exploring the complexity of interaction through critical analysis, and by engaging expertise through focus group.

4.1 Philosophy and Research Design

This research is conducted through a pragmatic lens, with an ontological position which accepts differing worldviews within context. This research is a component of a design process, that aims to find solutions to a clearly defined problem. This problem space is multifaceted and contains components that are both positivistic and socially constructed. For example, much of the computer science within this study clearly represents a single reality, whereas participant concepts of privacy, value and technology interaction are socially constructed. It is also important to consider that elements of this research are practice led, and the research design requires consideration from an empirical HCI prospective, as methodologies drawn from the social sciences need to be reconsidered (Turnhout, 2017). In essence this research employs a Mixed Methods Research (MMR) approach. Researchers employ MMR as it is considered an appropriate way to explore research questions that are, *'multifaceted, having implicit or explicit interrelated components'* (Tashakkori & Teddlie, 2015, p. 620). Mixed Methods Research involves the *'adoption of both quantitative and qualitative methods, and the mixing and blending of the resulting data to provide a stronger understanding of the problem or question'* (Creswell, 2011, p. 215). MMR has become a popular approach since the 1980's and is

'referred to as the third option, to the either-or dichotomy, of quantitative or qualitative research' (Tashakkori & Teddlie, 2015, p. 618). The approach advocates *'the use of whatever methodological tools are required to answer the research questions'* (Tashakkori & Teddlie, 2009, p. 7). It is argued that within HCI multidisciplinary research, the use of mixed methods is commonplace, although *'authors do not typically refer to their research as mixed-methods'* (Turnhout, 2017, p. 361). Within this research it is important to understand how a traditional mixed methods approach relates to the philosophies, methods and practices of knowledge generation through both design and practice (Zimmerman, 2007; Frayling, 1994). When considering this research from an epistemological perspective the work of Turnhout (2014), *Design Patterns for Mixed-Method Research in HCI* has been influential. Within this work, it is argued that it is difficult to transpose the approach of mixed methods research directly from the social sciences, as there are differences in approach, and what is considered acceptable knowledge. It is argued that *'a core difference between social sciences and HCI is the status of theory. Being a design-oriented field, HCI strives to combine descriptive and prescriptive theory, and recognizes artefacts as a legitimate form of knowledge'* (p. 362).

Turnhout interprets how the methods and patterns employed within HCI research fit across a framework that recognises objective application, *'Application Domain'*, paralleled by theoretical knowledge and showcase, *'Available Work'* (Fig. 12). Within this framework 5 epistemological domains are defined within which traditional methods aligned with the social sciences co-exist alongside those found within HCI. These domains are described in the following section.

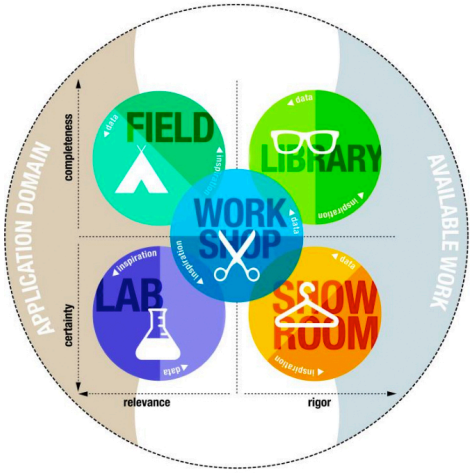


Figure 12: Mapping of Mixed Methods Research Design (Turnhout, 2017, p. 362)

- **Field Methods** are employed to ascertain the context of the design and application domain and take their direction from interpretive processes within the social sciences.
- **Library Methods** are considered both inspirational and data orientated and involve literature study as well as the consideration and bench marking of technology.
- **Workshop** methods aim to conceive or improve a solution through practice, engaging in research- through-design, pragmatically creating artefacts to support and resolve the research questions.
- **Lab Studies** involves the testing or critique of product or prototypes to test against the defined criteria or objective.
- **Showroom** methods are those which make work more reusable by other researchers or designers. These can take the form of showcase, guideline or frameworks.

Turnhout et al (2017) argue that there are 3 trade-offs within domains. The first, *Rigor* and *Relevance*, which recognises the cyclic nature of research. At the Relevance stage, '*the researcher learns about and changes the application domain'*. At the Rigor stage *'researchers learn about and contribute to available work'* (p. 362). The second trade-off is between *Certainty* and *Completeness*, *Certainty* relates to the need for precision of measurement, which may require quantitative methods under lab conditions, and *Completeness*, which is focused on the character of a system and may entail an ethnographic approach. The final trade-off is between *Inspiration* or *Data*. In any given domain the research may generate knowledge

through a subjective approach engaging intuition and creativity or through a data driven objective lens, becoming analytical and positivistic.

Turnhout's framework offers a model through which to chart mixed methods HCI research across abstraction and method domains, as investigations endeavour to explore *Niche's*, develop *Proposition*, contextualize and *Frame* the domain, build *Content,* or develop *Guidance* for further research and development (Turnhout et al, 2017, p. 366). As an exercise to assist in the design of this research, components have been mapped relative to the framework proposed by Turnhout et al.

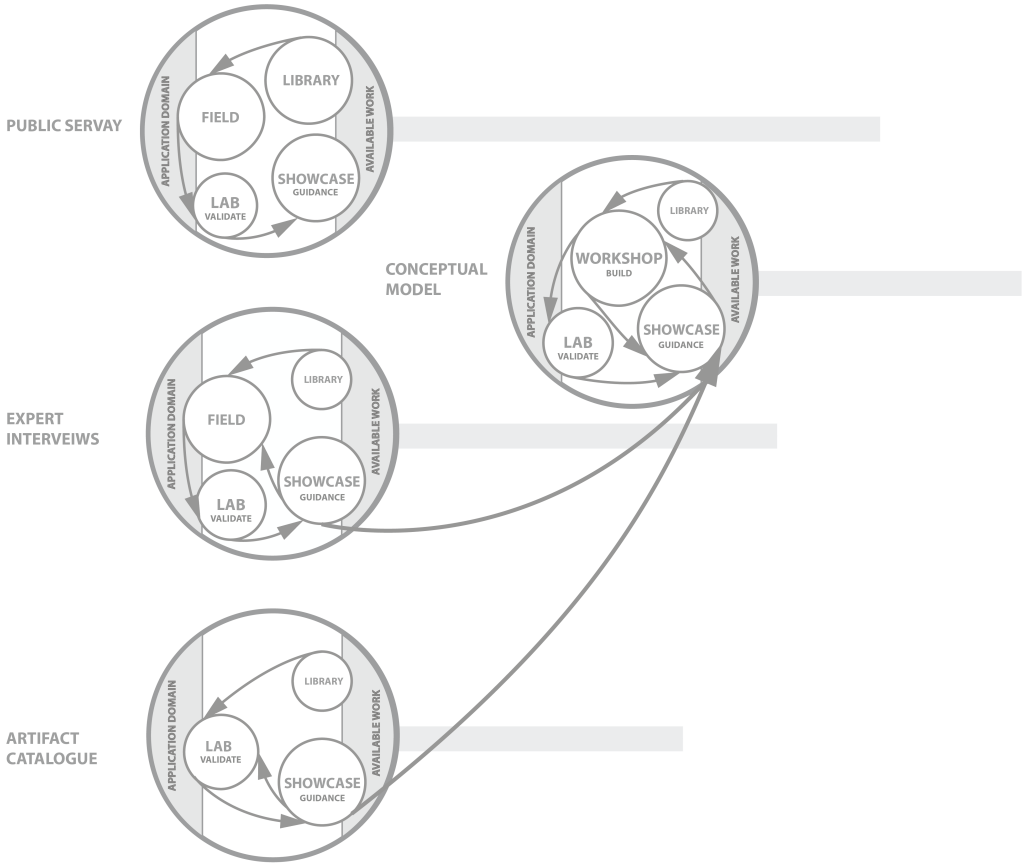


Figure 13: Contemporary Mapping of Mixed Methods HCI Research Design

Below is a traditional mapping of the mixed methods employed within this research. The illustration considers the components of data gathering, the sequence, and their relationship to one another. There is an *Explanatory Sequential* element in which the Expert Interviews and the Artefact Catalogue inform the Conceptual Modelling Component. All elements then stand as individual components and are interpreted in an *Embedded* mixed method design to answer the research questions. (Creswell, 2011, p. 220).

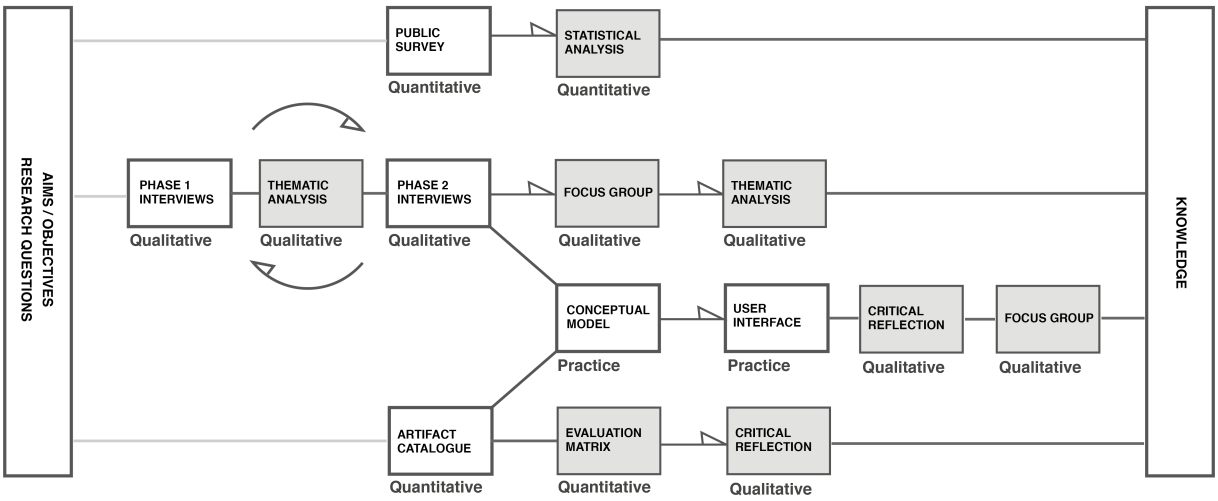


Figure 14: Traditional Mapping of Mixed Methods Research Design

When considering this research from an axiological standpoint, the aims and objectives, the questions and the subsequent research design is considered from a pragmatic position. The research is seen as a preliminary component part of a wider design process, and value is placed on the usefulness, workability, and practicality of ideas. This pragmatic paradigm influences the interpretation of results, as value is placed on the practical application of findings relating primarily to usability and adoption. Awareness of limited bias has already been discussed (see: 1.1). From an ethical standpoint this research recognises the notion of Value-Sensitive Design (Friedman, 2008) in what is, in essence, a domain that is conceived for the emancipation of the Human from asymmetric power imbalances. As Friedman (2008) explains: *'Human values and ethical considerations no longer stand apart from the Human-Computer Interaction (HCI)*

community but are fundamentally part of our practice' (p. 1178). Human-Centred Ethics aside, this research also recognises the potential values offered by the Decentralised Internet in terms of new business opportunities and the general development of commerce. (Levine, et al, 1999, Searls, 2012).

4.2 Artefact Catalogue

The Artefact Catalogue component of this research is designed to inform the primary and intersecting areas of investigation presented in the Scope of Research Section (see: 1.5). It is imperative that this study acts to consider both the empirical as well as theoretical developments within the decentralised field. The objective is to understand the contemporary thinking, trajectory, historical developments, interaction models and technological capability. Within this component, the method is designed to identify, collect and analyse contemporary and historical projects, start-up companies, applications, technologies and technical standards relating to the Decentralised Internet. For the purposes of this section, these entities will collectively be referred to as artefacts.

By considering historical and contemporary artefacts the study investigates the following:

- The explored notions of privacy, the perceived risks and the tools designed to empower the network participant.
- Existing and emergent decentralised models, the principles, concepts, interactive components and the design patterns.
- The underlying technology, hardware, software and standards. By investigating the artefacts, we consider the usability of decentralised technology and understand the complexity, metaphor and mental models.
- The value propositions proposed across a spectrum of innovations to form a judgement against the literature, public consultation and expert interview as to whether these concepts hold meaningful value, either in their current form or when further supported through emerging technological developments.

- Finally, this exercise will assist in considering the assumption that there has been limited up take in decentralised technology.

This component does not aim to build an exhaustive catalogue of artefact analysis. This is considered impractical and outside of the scope of this study. Instead this component aims to build a broad catalogue, comprising a framework of categorisation and valuation to act as a resource for further investigation of variables within this research and beyond.

It is important to communicate why technology standards which may not be considered an artifact, and technologies which are clearly centralised in nature, are included within this catalogue. With regards standards they are considered important to the evolution of the decentralised domain, and act as a supporting layer in the understanding of artefact development and future trends. Centralised artifacts, more specifically early AI driven device controllers are included as it is important to understand developing technology which may be considered centralised, that might be important in a decentralised guise, to the overall function, interaction, and friction reduction of a Human-Centred Data Ecosystem.

4.2.1 Artefact Catalogue Method

The method engaged involves the discovery and cataloguing of artefacts. The artefacts are compiled, and a preliminary analysis considers if a recorded artefact is relevant to, and supporting of, endeavours to decentralise the Internet. Furthermore, artefacts are categorised with several predefined titles based on initial investigation. The categorisation of artifacts was established initially based on those found within Project VRM (Project VRM, 2017). The categories were then evolved further as artifacts where investigated and understood in greater depth. Additional categories were considered as the general trajectory and evolution of artifacts became clearer.

The next stage involves an *Evaluation Matrix*, based on a *Concept Selection* method (Pugh, 1981), A method was originally developed to aid engineers make design choices across different concepts. Along one axis of the Matrix the concepts are listed, on the other a number of predefined criteria. The criteria are scored resulting in an overall score for each concept. The criteria is established in advance and offers a consistent measure for evaluation.

Finally, the artefacts are relisted by category together with the results of the Evaluation Matrix, incorporating visual formatting and colour coding to support the communication of findings.

4.2.2 Rationale for Method Choice

There are many methods for evaluating and testing digital artefacts in the HCI literature, most of which are empirical in nature. These include: *Heuristic Evaluation, Cognitive Walkthroughs, Formal Usability Inspections, Pluralistic Walkthroughs, Feature inspection, Consistency Inspection and Standards Inspections* (Nielsen, 1994, p. 413). Any one of these methods would have generated meaningful information regarding the merits of any given technical artefact across a spectrum of criteria. These methods are focused on usability at the interface layer, and do not account for any wider considerations, such as value or current evolutionary state. In addition, the models require a significant amount of time to undertake. In the light of these considerations, given the volume of artefacts to be considered, employing a Categorisation and Evaluation Matrix method was considered appropriate.

4.2.3 Artefact Catalogue Process

During the first phase a catalogue of contemporary and historical artefacts was established. The artefacts were compiled from two primary sources. The first, Project VRM (2017), (Vendor Relationship Management), is an ongoing project affiliated with Harvard University. The project maintains an exhaustive list of related decentralised initiatives. Its

primary function is to record projects developed to facilitate commerce within a Decentralised Internet, although the resource now spans far beyond its original context. The second source is Re-Decentralised.org, (2017) an organisation comprising privacy advocates, researchers and technologists who maintain an index of primarily technology focused decentralised initiatives. As well as general listings, the resource contains a repository of semi-structured interviews with domain experts. In addition, new entries have been added to the Artefact Catalogue throughout the course of this research as conversation, further reading and conference attendance have highlighted new developments.

Identified artefacts have been compiled into a spreadsheet comprising meta data: listing the name of the project, a description, brief overview and web link to resources. The artefacts have been categorised by type, before being subjected to a Critical Analysis, facilitated through an Evaluation Matrix. The evaluation posed several questions relating to the artefact with the objective of understanding the offering across a number of criteria. These included the success or failure of the concept in terms of adoption, the artefact's technological capability, its functionality, usability and overall value. This evaluation has been conducted in order to gain an overarching understanding of all artefacts identified. The process will conclude with highly significant artefacts being singled out for greater scrutiny and exposure.

4.2.4 Critical Analysis

The first question posed, is simply to consider whether a given artefact is indeed related to the decentralised domain and whether this artefact can be considered to be a decentralised technology or project which would support efforts to establish a Decentralised Internet? The definition of decentralised in this context is taken from Feross Aboukhadijeh. *'The Decentralised Internet is a system of interconnected, independent, privately owned computers that work together to provide private, secure, censorship-resistant access to information and services'* (Aboukhadijeh, 2016).

The next phase involved the categorisation of the artefact against the following predefined and subsequently evolved titles:

- Privacy Protection
- Personal Data Store with Basic Application
- Personal Data Store with Extended Capability
- Collation Personal Data Store
- VRM Principles
- Intent Casting Application
- Personal Assistant
- Personal Data Marketing Tool
- SSI/ Full System Infrastructure
- SSI/ Limited Identity Mechanism
- DAPPS 'Decentralised Applications'
- Decentralised Storage
- Decentralised Stack Software or Standard

Once categorised, each artefact is critically evaluated against predefined criteria. Each element is scored 1-5, and colour coded to allow better visual communication of the final analysis.

The questions developed to critically evaluate artefacts are as follows:

- Is there clear evidence that the artefact is still active? Is there evidence of activity within the last 3 months, *i.e.* updates, news, blog entries.
- Is the offering truly decentralised and open source, or does it have some degree of third-party tie in?
- Is the technology novel or is it a decentralised version of an established centralised interaction?
- Is there a clearly defined value proposition?
- Is establishing a clear mental model of the technology straightforward from its description?
- Does the technology incorporate Identity or Blockchain technology?
- If the project is no longer active, might contemporary technology revitalise the concept?

The Evaluation Matrix is developed as part of the hosting spreadsheet with a list of questions, numbered results and colour coding to aid data visualisation. The resulting matrix offers a valuable data source through which existing and historic artefacts can be better understood. The data is an important element in understanding the decentralised landscape, and acts as a verification tool when considering the questions posed within this research. A sample of the completed Artefact Catalogue can be found in the appendix of this thesis (see: A1).

4.2.5 Pre-empted Artefact Catalogue Limitations

There are limitations to this method given the spectrum of artefacts being considered and the practicality of investigation. Traditional HCI methods of evaluation might have been employed as communicated in the Rational for Method Choice section but given the scope of this research, the limited information required and the restriction of time, this was considered impractical and although this might be considered a limitation, in reality a surface level approach still provided significant data.

There are limitations in a critical analysis conducted by an individual. It is subjective in nature, and reliant on the knowledge and experience of the researcher. The analysis is also potentially vulnerable to bias. In the case of this analysis, the expertise of the researcher is considered sufficient to make valid judgments, and there is awareness and conscious effort made to mitigate the potential for bias.

The following considers the specific limitations within each of the questions posed:

- **Gauging the Success or Longevity of an Artefact.** Within this research, an assumption has been stated that there has been a significant number of attempts to develop technologies to introduce concepts of a Decentralised Internet to the mainstream, but that many of these have failed to attract significant adoption. This is a difficult assumption to validate, as without reliable metrics from the developers involved, it is difficult to gauge usage, adoption or the current position of projects. In

the absence of this developer data, the objective is limited to give an indication of an artefact's position, and not its definite position or quantitative metric. At this stage this is considered acceptable.

- **Is an Artefact Truly Decentralised?** Whether a technology is truly decentralised is an important factor. Is it Open Source? Does it rely on third party for all or part of its architecture? Are there terms and conditions? Where is the data held? Is there portability of data? Is there interoperability? These questions are varied and difficult to validate. At this stage a generalised evaluation of these questions is considered acceptable.
- **Originality and Value of an Artefact** Originality and value of the artefact is related to the probability of adoption. Originality is important as it potentially highlights the emergent interactions that a decentralised ecosystem offers over a centralised model. Value is again an important measure and exploration within this research. The judgement of value might have been conducted through case study and participant questionnaire, but these methods are impractical given the scope of this component. Instead a general critical judgement of value is considered appropriate.
- **Artefact Usability, Complexity and Mental Model** It is important to attempt to gauge the complexity, the usability and, in the case of this study, the mental model associated with the artefact. It is important to consider that the artefact might be a finished application, partly developed or merely conceptual. Usability and user experience evaluation methods (Nielsen, 1993) could be applied to evaluate each artefact in this context but, given the scope of investigation and practicality of accessing and engaging each artefact, the informed judgement of the researcher, a surface level evaluation at this stage is considered to hold sufficient value and offer greater efficiency.
- **Identity, Blockchain and Revitalisation** The final questions posed are related to contemporary developments in decentralisation, and specifically, concepts of identity and distributed ledger technology. This assessment is relatively straightforward to ascertain and has no limitations. The question of revitalisation of historic artefacts relates to contemporary developments, not only related to Identity and Blockchain, but also other developments; for example, distributed storage. The prospect of reinvigoration of historical artefacts through contemporary and emergent technology and developing concepts is, to a degree, subjective in nature, and reliant on the understanding, expertise and creative insight of the researcher.

4.3 Public Survey

The Public Survey component of this research was designed to inform the primary Privacy and Adoption areas of investigation, and subsequently the Value Proposition intersection presented in the Scope of Research section of this thesis (see: 1.5). The Public Survey investigated attitudes towards Internet usage, data privacy, the disclosure and secondary usage of personal data, and engagement with activities and opportunities to protect and control personal information. Analysis of the data gathered provided a detailed picture of public perceptions and attitudes at a descriptive level. Latent considerations were designed into the survey to uncover signifiers relating to Catalyst for Adoption, Value Propositions and potential Development Strategies.

From the research questions posed and the aforementioned areas of investigation, a number of aims have been developed to aid in the design of the survey questions. The Public Survey aims to:

- Investigate awareness and attitudes towards perceived privacy issues online.
- Investigate attitudes towards adoption of decentralised technologies.
- Investigate participation with, and the barriers to entry for, a Decentralised Internet.
- Investigate potential areas of value and development strategies for a Decentralised Internet.

4.3.1 Public Survey Method

The method engaged involved the development of a Public Survey comprising of Likert Items and Forced Binary Questions. The survey was distributed via an Internet mediated artefact in physical locations via a handheld device, and remotely via an Internet URL. The compiled data has been formatted with Microsoft Excel before being Imported into IBM SPSS.

The data has been analysed through descriptive statistics and Likert Scale correlation and comparison.

4.3.2 Rationale for Method Choice

A survey method was chosen as it offered a consistent, efficient method to establish statistically reliable quantitative data around defined areas of interest. Observational methods and semi-structured interviews might have been engaged to generate similar data, but the timescales involved would have been impractical and impossibly labour intensive. An Internet Mediated distribution method was selected as it offers advantages and efficiencies over a physical alternative. The digital survey allowed for a broader cast of participant engagement. The survey was conducted via direct email request, and in the field where participants are physically approached and asked to conduct the survey via a supplied handheld device. The Internet mediated option also allowed for quick and easy download of data in a suitable digital format. In this case, following recommendation, cost implication, and information regarding data protection, *So-Go-Survey*[22] was chosen as the digital survey tool.

4.3.3 Public Survey Process

The following paragraphs outline the design, development, piloting and dissemination of the survey.

4.3.3.1 Survey Design

In order to develop the survey in a rigorous manner, a Survey Method Plan has been developed following recommendations established by John Creswell (Creswell, 2003, p. 146). The plan in essence comprised a check list of seven questions which needed to be considered. The full Survey Method Plan can be found in the appendix of this thesis (see: B1). The questions considered the purpose of the survey and rationale for the method choice whether it

22: *So Go Survey* is an online survey tool. https://www.sogosurvey.com

is cross-sectional or longitudinal, the population, sample size, sample randomisation, and stratification.

4.3.3.2 Question Design

In order to develop the survey questions relative to the defined objectives, a Data Requirements Table has been developed (Saunders, 2009, p. 425). The table can be found within the appendix of this thesis (see: B2). The table defines themes of investigation, the variables required, how data will be measured, and finally, the questions.

Within the Data Requirements Table investigation categories are defined under which a collection of questions have been designed. Areas of investigation were established to:

- understand the level of awareness of surveillance capitalism.
- understand the attitude towards personal data collection and use.
- understand concerns of personal data collection.
- ascertain the extent of actions taken to protect privacy.
- understand attitudes towards agency and adoption of decentralised technologies.

A great deal of consideration has been given to the reduction of bias and the consistent interpretation of questions. In this respect, direction has been taken from Floyd J Fowler (1995), who advises that: *'one standard for a good question is that all the people answering it should understand it in a consistent way, and in a way that is consistent with what the researcher expected it to mean'* (p.2). The language within these questions is complex, and could potentially be interpreted in different ways, the preliminary piloting focused on this issue. Another concern in terms of bias is the wording of the questions, so consideration was given to reducing any leading tendencies.

The survey was made up of 52 questions consisting of Likert Items and Forced Binary Questions. The development and mapping of these questions is presented in the Data Requirements Table. Most of the questions are original to this study, but a number have been

re-appropriated from similar recent surveys. (Coles-kemp, 2010), (Rainie, 2013), (Digital Catapult, 2016). The origin of questions is highlighted within the Data Requirements Table.

The questions were designed to function in two forms. Firstly, as individual Likert Elements targeting specific desired information and Second, collections of Likert Elements have been designed to fit within combined Likert Scales (Likert, 1932). Several forced binary questions were also grouped to enable combined scales. The combined scale categories are listed as follows:

- Digital fluency - (Q2,3,4) - Likert Scale
- Understanding of personal data and its value - (Q5,6,7,8,9,10,11) - Likert Scale
- Comfort level with network engagement (Q12,13,14,15,16)- Likert Scale
- Perception of importance regarding personal data - (Q18,a,b,c,d,e,f,g,h,i) - Likert Scale
- Effort made to protect privacy - (Q18,a,b,c,d,e,f,g,h,i,j,k,l) - Forced Binary
- Interest in controlling personal data - (Q21,22,23,26) - Forced Binary

An additional table, illustrating the combined scales, is embedded within the Data Requirements Table (see: B2).

4.3.3.3 Piloting

The survey was first piloted with a small number of academics with knowledge of the chosen field. The objective of this exercise was to seek reassurance that the questions posed were relevant to the research questions and areas of investigation. The survey was then piloted within a wider group of academics online, with feedback being invited with regards to the wording and clarity of the questions, the suitability of responses, and any views on bias. A full transcript of responses received during the piloting process can be found within the appendix of this thesis (see: B3).

4.3.3.4 Distribution and Sampling

The survey was cross-sectional in nature and gathered a representative sample of general computer users. Efforts were made to adopt a stratified sampling technique to avoid over representation. Efforts were also made to engage participants in an Internet Mediated Survey at physically different locations, and also through direct URL. The rationale has been to engage participants across demographics, varied age range, gender, professional, and non-professional. Within the sample of university students, a spectrum of subject disciplines has been targeted to reduce the prospect of participants clustering around a professional discipline. Varied physical locations around the Salford University campus supported this objective.

A target number of 300 participants was derived from Fowlers reference table, which suggests that a confidence range of +/- 5% across any given question can be estimated from a sample size of 300 randomly sampled individuals (Fowler, 1993, p. 31).

Figure 15: Public Survey Data Gathering Session

The photograph in Figure 15, was taken during a survey data gathering session held in the entrance of Salford University's New Adelphi Arts and Media building. Similar events were held in the main campus library and the university's Media City Campus.

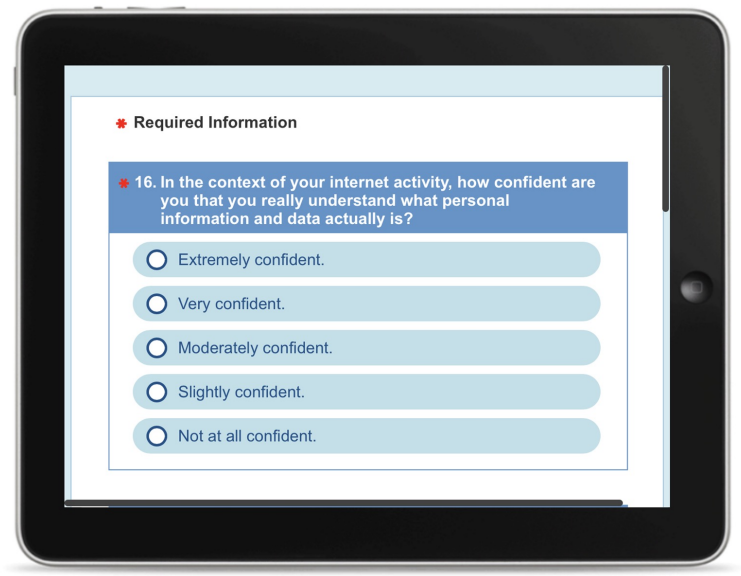


Figure 16: Public Survey Developed as Internet Mediated Artefact

Figure 16 is a direct screen capture from the Internet Mediated Survey, built and distributed through the SoGoSurvay service.

4.3.4 Survey Analysis

The survey data was retrieved from the digital survey tool and coded within Microsoft Excel. It has then been transposed into IBM SPSS and analysed through both descriptive statistics and combined scales through correlation and comparison.

4.3.4.1 Descriptive Analysis

As part of the descriptive analysis the mean is calculated for each ordinal response type question, so the central tendency could be considered. Descriptive analysis was rendered as bar charts; the resulting images offer a strong visual means by which to better understand the data.

4.3.4.2 Statistical Analysis

Within SPSS, data is combined into the defined scales. The scales have been tested for internal consistency by applying a Cronbach's Alpha function. The next phase applied composite measures to illustrate the mean and standard deviation. In the final phase of analysis, correlations between scales have been tested through a Pearson Correlation Coefficient. Scales have also been considered through an Ordinal Dependent - Independent Variable Comparison, using a Independent Sample T Test. Results were then reported as numeric, table and chart as appropriate.

4.3.5 Pre-empted Public Survey Limitations

The most significant potential limitation of the survey is the preconceived opinion and bias of the sample group. The subject area itself is inflammatory. Data privacy is a topical issue that is often sensationalised in the media, and as such could cause validity issues. The survey is not designed to support the Privacy Paradox (Norberg, 2007), and it can be argued that the attitude towards data privacy is, to a degree, predictable. This survey is focusing on broader issues and latent signifiers. If direct privacy violation is left out of the equation, the deeper nuanced areas of investigation should not be affected by preconceived ideas. The questions were designed with this limitation in mind, and the wording is phrased to avoid obvious leading tendency.

The survey was designed to strike a balance between gathering a broad range of views around a number of core themes, and the search for relationship through correlation of combined scales. The broad scope of questioning could be considered a limitation, as without narrow focus the verification of combined scales is limited. The benefit of this broad data set is, however, considered the priority at this specific point within this research. A further study might choose to focus on specific issues and build multiple combined scales to establish validated quantitative results.

The sample size could be considered a limitation, in that the survey engaged the minimum target number of 300 participants. A higher number would have been desirable. The challenge, however, of engaging a larger number of varied individuals in a survey, which on average took 12minutes to complete, is difficult. The strategy of approaching individuals in a physical location with a handheld device proved, in this instance, to be the most productive method.

Every effort was made to stratify the sample and engage participants from a broad demographic. As can be seen in the results, (see: 5.3.1) the effects of this are limited. But efforts have provided a reasonably diverse sample and stratification has improved the validity of the data.

4.4 Expert Interviews

Expert interviews have been conducted to inform all primary areas of investigation and dynamic intersections presented in the Scope of Research section (see: 1.5). It is imperative that this study engaged with a broad range of expertise, directly relating to the Decentralised Internet, as well as the wider field of user experience and user interface design. It is important to note that this component is multi-phased, and that there is a degree of overlap with the Practise Led Component of this research. The first phase explored the broad decentralised domain with the objective of understanding the trend and direction of travel, the technological usability barriers and obstacles, and views around sustainable adoption. A second phase of interviews focused on individuals from the decentralised space with an active interest in Self-Sovereign Technologies. These interviews are narrower in scope and focused specifically on user interaction and adoption. A third phase related to data gathered from a focus group conducted as part of the Practice Led Component of this research. As this data had value in the context of this analysis, it was subjected to, and included in the same analytical process.

4.4.1 Expert Interview Method

Within the first stages of research, expert interviews were considered to be a productive and effective method to obtain useful, insightful information about a given domain. This component engaged an interpretivist approach, deploying semi-structured methods to gather qualitative data across three phases of expert interviews. Ann Blandford (2013) describes the role of semi-structured qualitative study in HCI as having two purposes, *'understanding current needs and practices, and evaluating the effects of new technologies in practice'* (p. 7). In planning and developing this component of research the *PRETAR* framework was applied: *Purpose, Resources, Ethical, Techniques, Analysis, Reporting,* (Blandford, 2008a). In terms of technique, semi-structured interviews have been developed according to the guidelines offered by Arthur and Nazroo (2003) and Legard et al, (2003). Clear themes have been developed

resulting in primary and secondary questions, and participant materials, copies of which can be found in the appendix of this thesis (see: C3, C4). The Analysis and Reporting method used within this component, is *Thematic Content Analysis* (Braun & Clarke, 2006).

4.4.2 Expert Interview Process

Figure 17 illustrates the phases of this research component. The analysis is cyclic and as further data is gathered it feeds back into and informs the subsequent cycle of analysis. For reference the figure also illustrates chronologically how the results of this component informed the Conceptual Modelling and Preliminary UI stages of the Practice Led Component of this research

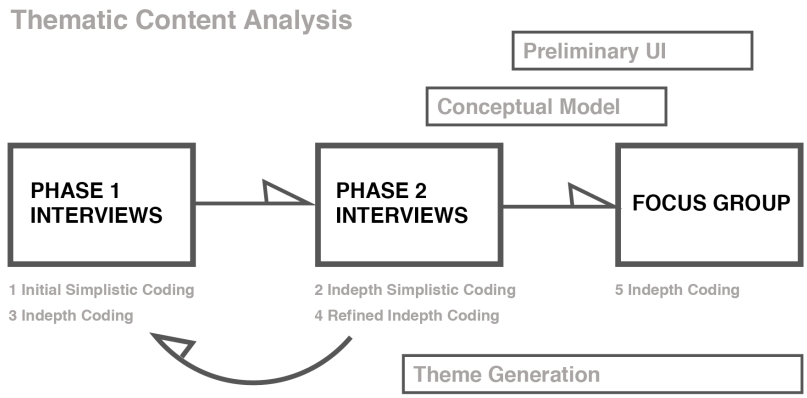


Figure 17: Expert Interview / Focus Group Phases

4.4.3 Phase One, Sample, Objective and Questions

Following attendance at the Re-Decentralised Conference 2015 (Redecentralize.org, 2015) a number of prominent contacts in the decentralised field where established. As a result of this networking, an invitation was received to contribute to the Hi-Project (2017). During this participation and following weekly conference calls, several prominent members of the decentralised community where approached and asked to take part in this research. The objective at that stage was to gain a broad understanding of the decentralised space. The

individuals targeted were suitably placed to offer their expertise. A full list of phase one participants can be found in appendix of this thesis (see: C1).

Semi-structured interviews were conducted through questions developed around five themes. The themes were defined to build a broad understanding of the decentralised domain, the driving principles, developing technologies and expert opinion regarding adoption and usability of artefacts. A full transcript of the phase one questions can be found in the appendix of this thesis (see: C3).

The five themes for investigation were as follows:

- Participant background and work within the decentralised field.
- The evolving trajectory of decentralised technologies and its associated application.
- The perceived barriers to adoption and the challenges ahead.
- User Journeys within a decentralised personal data framework.
- User interaction and interface issues.

Once conducted the interviews were transcribed and analysed through basic thematic analysis. Rudimentary Codes were applied allowing basic patterns and a consensus to be drawn, with the results of this preliminary stage shaping the development of the second phase of questions. It is important to note that this data was subsequently analysed for a second time, through in-depth thematic analysis following the phase two interviews, together with data generated through the Practice Led Component focus group.

4.4.4 Phase Two, Sample, Objective and Questions

Following the analysis of data from the phase one interviews, further reading and attendance at MyData 2018, a decision was made to narrow the focus of investigation to Self-Sovereign Identity. The rationale for this decision is described within the SSI section of this thesis (see: 3). A list of prominent individuals was drawn up with expert understanding of the

decentralised space, including the knowledge of its relationship with identity management, and the supporting concepts and technologies. These individuals where approached directly. A full list of phase two participants can be found in the appendix of this thesis (see: C2).

The existing semi-structured interview questions developed for phase one were used again, in a modified form, and reframed around the narrowed theme of SSI. A full transcript of questions can be found in the appendix of this thesis (see: C4).

The three narrowed investigative themes were as follows:

- The Concept of Self-Sovereign-Identity.
- The Complexity at the Interface Layer of Decentralised Technologies.
- The Adoption and Value Proposition for Decentralised Technologies.

Once conducted, the interviews were transcribed and analysed through an in-depth thematic content analysis (see: 5.4.3).

4.4.5 Phase Three, Sample, Objective and Questions

Phase three, took advantage of the assemblance of user experience and user interface experts during the focus group element of the Practice Led Component of this research. Participants were selected for their reputation, experience and expertise, and approached directly. Prior to the focus groups a full website was constructed and disseminated to communicate the results of a conceptual modelling exercise and subsequent user interface development (see: D4). As part of the focus group agenda (see: D5) a presentation was given regarding the general decentralised space and problem area. The primary objective of this focus group was to gain expert opinion on the developed user interface. During this event, the opportunity was taken to ask this group of experts' questions regarding not only interface layer related issues, but also questions relating to user perceptions, value and adoption of decentralised technologies. Guidance was taken from Finch and Lewis (2003) in the design and planning of the focus group. A full list of topics and themes discussed, and the planned

agenda can be found in the appendix of this thesis (see: D5). The focus group was transcribed in full. The data was combined with the first and the second stages of this component, before being taken through an additional cycle of thematic analysis.

4.4.6 Thematic Content Analysis

A qualitative analytic method was required to make sense of the data gathered through semi-structured interviews. Thematic Content Analysis was selected as it offers an accessible and theoretically flexible approach. Braun & Clarke (2006) explain: *'through its theoretical freedom, thematic analysis provides a flexible and useful tool, which can potentially provide a rich and detailed, yet complex account of the data'* (p. 5). The method generally consists of the *'identifying, analysing, and reporting patterns (themes) within data'* (p. 6), and requires the development and application of codes to the data. The coding develops through convergence and grouping into defined themes. Braun & Clarke describe two levels of themes: *Semantic* and *Latent*. Semantic themes emerge through the analysis of the data without drawing inferences beyond what a participant has said. Latent themes are developed by moving the analysis beyond the surface, examining and interpreting the data at a deeper level. At this stage the researcher: *'starts to identify or examine the underlying ideas, assumptions, and conceptualisations—and ideologies—that are theorised as shaping or informing the semantic content of the data'* (p. 84). Braun and Clarke state the importance of defining the theoretical framework through which the data will be considered. This has been clearly defined within the introduction and literature review sections of this thesis. Arguments surrounding Surveillance Capitalism (see: 2.1), Privacy (see: 2.2) and the concepts or Human Data Interaction (see: 2.4), help to frame the development of themes concerning justification for, and potential propositions of value. Investigation of HCI cognitive theory (see: 2.3) and arguments surrounding technology adoption (see: 2.2.4 / 2.2.5) assist in the development and justification

for themes regarding friction found in potential interactions, while offering a context through which to explore for solutions.

When conducting Thematic Content Analysis there are a number of clearly defined steps to be undertaken. These steps are illustrated in Table 2 below.

Thematic Content Analysis Steps

Familiarising yourself with the data	i) Narrative preparation, i.e. transcribing data ii) (re-reading the data and noting down initial ideas
Generating initial codes	i) Coding interesting features of the data in a systematic fashion across entire data set ii) Collating data relevant to each code
Searching for themes	i) Collating codes into potential themes ii) Gathering all data relevant to each potential theme
Reviewing themes	i) Checking if themes work in relation to the coded extracts ii) Checking if themes work in relation to the entire data set iii) Reviewing data to search for additional themes iv) Generating a thematic "map" of the analysis
Defining and naming themes	i) On-going analysis is refine the specifics of each theme and the overall story the analysis tells ii) Generating clear definitions and names for each theme
Producing the report	i) Selection of vivid, compelling extracts examples ii) Final analysis of selected extracts iii)Relating the analysis back to the research question, objectives and previous literature review

Derived from Braun, V. & Clarke, V. (2006)

Table 2: Thematic Content Analysis Steps

All recorded expert interviews were transcribed and time coded. The transcripts were then loaded into the qualitative research software package Nvivo where codes were developed and applied. Following the coding of the data, printed transcripts were evaluated through a process of memoing, and initial themes were developed. Themes were reviewed and refined through the mapping software Freemind and a full visual mapping was defined (see: 5.4.4). Finally, the themes have been described with supporting verbatim quotation where available.

4.4.7 Pre-empted Expert Interview Component Limitations

Participant choice, access and appropriateness is recognised as a challenge and potential limitation within this study. Identifying individuals with the required knowledge and securing participant engagement requires subtle persuasion and planning. During the interviews the interpersonal nature of the situation required a degree of skill. There is the potential for bias

in unconscious signalling during discussion, and the open-ended nature of the questioning makes comparing outcomes difficult. There is a clear understanding of the limitations and potential issues when developing and conducting semi-structured interviews but, through following guidance defined by Legard et al, there is confidence that any limitations have been minimised (Legard et al, 2003). With respect to the data analysis and the decision to adopt Thematic Content Analyses (Braun & Clarke, 2006), potential limitations rise from the researcher essentially being the instrument of analysis. It is recognised that this analysis is being undertaken by an embryonic researcher, and as such, emphasis was placed on process and rigor. To support this objective, Lorelli et al's work entitled *Thematic Analysis: Striving to meet the Trustworthiness Criteria* has been consulted and followed. (Lorelli et al, 2017).

4.4.8 Rationale for Method Choice

Conducting semi-structured interviews with domain professionals is considered an appropriate way of gathering domain specific information. As Bogner (2009) explains: *'talking to experts in the exploratory phase of a project is a more efficient and concentrated method of gathering data than, for instance, participatory observation or systematic quantitative surveys'* (p .4). When considering which means of qualitative analysis to employ, a number of methods were considered. *Phenomenology, Discourse Analysis*, and *Narrative Analysis* were among the theories considered with the choice ultimately being made between *Grounded Theory* (Glaser, 1992, Strass & Corbin, 1998) and *Thematic Content Analyses* (Braun & Clarke, 2006). Thematic Analysis was selected over Grounded Theory for several reasons. The researcher has prior knowledge of the domain and approaches the questions with a number of pre-assumptions. Though there is room for emergence, the analysis leans towards deductive rather than inductive reasoning. By contrast, Grounded Theory is prescriptive in its process, and theoretical sampling and axial coding did not seem appropriate. There is the consideration that this analysis does not necessarily aim to develop theory directly. Instead the intention is to

generate broad themes to be considered alongside other data to answer the research questions. When these points are considered against the inexperience of the researcher, the flexibility and the methods entry level credentials, Thematic Content Analysis, on balance, is considered the appropriate choice.

4.5 Practice Led Component

The Practice Led Component of this research has been developed to consider the Technology and Adoption areas of investigation, and the intersection of Usability as presented in the Scope of Research section. (see: 1.5). This component focuses on the decentralised model of Self-Sovereign Identity as a framework for exploration. The rationale for the decision to focus on Self-Sovereign Identity can be found within Chapter 3, which is dedicated to this decentralised model. It has previously been stated that this research will not endeavour to design and build a decentralised system or interface layer, and that initial exploration of existing projects and technologies suggested that there is a great deal of preliminary work to be done in understanding and analysing the relevant variables before any further attempt is made to design and build product. This is indeed the case, but it is also important to evaluate a user's interaction with a decentralised technology to better understand the complexities and potential friction. As such, the decentralised model of Self-Sovereign Identity, its established concepts, and preliminary prototypes offer the opportunity to extend existing artefacts to scale, without the need to develop a decentralised system and interface layer independently from the beginning.

The process of engaging in *'practice as research'* within this component is critical. In order to understand and develop knowledge regarding the interactions required for decentralised tools, we need to develop practically, or in this case extend a tangible artefact. As Archer (1995) argues: *'there are circumstances where the best or only way to shed light on*

a proposition, a principle, a material, a process or a function is to attempt to construct something, or to enact something, calculated to explore, embody or test it' (p. 11).

4.5.1 Practice Led Component Method

The applied research method within this component consists of three distinct phases, each of which is now described in detail.

4.5.2 Phase One: Defining A Conceptual Model

The current position of SSI comprises a clear description of principles, defined concepts, developing standards and basic prototypes. The objective of this component was to extend this current work to a scalable analogue interaction, and then critically analyse the result. As such, phase one of this component engaged a conceptual modelling methodology. Johnson and Henderson (2002), define a method for developing a *Conceptual Model*, arguing that designers should *'begin by designing what to design'*. The method describes steps which result in a structured text and table-based outcome. The first step is to define the application's purpose and *High-Level Functionality*. Once this is established the method moves to define the *Major Concepts and Vocabulary*. The next stage considers the conceptual objects visible to the user through what is termed as an *Objects and Operations Analysis*. This process investigates the objects users manipulate, the objects attributes and operations, and any relationships between them. The method then progresses to a *Task-to-Tool* mapping exercise. This considers how a user can use the application to carry out tasks combining the visible objects, attributes and operations, within this practice, the development of a preliminary user interface reflects this stage.

An important understanding within the Conceptual Modelling process, is that it is agile and flexible. As the designer progresses to develop a user interface, and evaluation methods are employed (Nielsen, 2005) the Conceptual Model is continually updated and refined.

4.5.3 Phase Two: Developing a Preliminary User Interface

With a defined conceptual model in place, this method progresses to establish a preliminary user interface in a wireframe format. The purpose of the wireframe is to visualise the defined objects and their relationships, while developing the minimal interface touch points to enable the required interactions. The mantra in developing the user interface is simplicity. There are no radical design choices and the process utilises existing interface design patterns.

4.5.4 Phase Three: Evaluation Focus Group

Within this research, the Preliminary User Interface and Conceptual Model are used to communicate the concept of the domain, its components, the required objects and the scope and scale of user interaction. During this final phase the developed model was subjected to a critical analysis, first by the researcher and subsequently through consultation with experts. The objective of this evaluation was to consider this first cycle of development, investigate the integrity of the fundamental concepts and offer a perspective from a user's standpoint. In order to facilitate this evaluation, a workshop was designed incorporating a focus group made up of usability and user interface professionals. Guidance was taken from Finch and Lewis (2003) in the design and planning of the focus group. Participants were selected locally, in the Manchester UK vicinity, for their reputation, experience and expertise. A website was authored together with supporting materials which were then distributed to participants ahead of the workshop (see: D4). The workshop started with a detailed presentation of the research to date, the concept of the Decentralised Internet, and its principles and objectives. An overview was given of existing technologies as well as Self-Sovereign-Identity. A detailed explanation was given regarding the practice-based element of this research, including conceptual model and initial UI development. Following the presentation and Q&A, a focus group was conducted based on the following topics of discussion.

Putting this technology in the hands of the masses from a task and UI perspective.

- How understandable are the concepts?
- What are your views on the complexity of these interactions?
- What are your views on the UI in front of you?
- How do we build something that's accessible?
- Thoughts on the mental models, and metaphors?
- Can any of this be automated?

Where is the value proposition that may drive adoption?

- Balancing Value Proposition against the Cognitive Load.
- Where's the value?
- Views of privacy around the world and across cultures.
- Bottom up – Top down?
- How much of the complexity does the user need to see?
- Trust Frameworks?
- Opportunity for business.
- Compliance.
- Quality data, society and Big Data...

The Focus Group recordings where transcribed and time coded. They were then considered first through a basic Thematic Analysis relative to the Practice Led Component and specific questions regarding usability and interaction. The data was then combined and analysed again through an in-depth Thematic content Analysis, considering broader themes established within the expert interview component of this research.

4.5.5 Pre-empted Practice Led Component Limitations

The fact that the interpretation of concepts, the development of the conceptual model, and the development of a preliminary user interface, was undertaken solely by the researcher could be seen as a limitation. It is only one interpretation of a design solution that might have manifested in a number of different forms. In defence of this process, the researcher has

experience and expertise in this area. The process has been undertaken through existing design patterns, and there has been a focus on minimalism and simplicity. There is confidence that the conceptual model is optimal, and the subsequent user interface is well crafted. A design of this kind would normally go through several design iterations. In this case, given the complexity, time constraint and access to professionals, there has only been one full cycle. The resulting outputs, however, are considered sufficiently developed to generate meaningful data.

4.5.6 Rationale for Method Choice

The utilisation of practice as a method of investigation within this research is critical. Throughout this research, it has become apparent that despite a sizeable development community, there is no publicly available conceptual model or representation of a complete user interaction or user interface for Self-Sovereign Identity. As a designer, in order to evaluate the usability, complexity, friction and user journeys, there has to be a tangible representation of a considered user experience. This requires design practice. The conceptual modelling process is a recognised HCI method for the initial development of concepts, interactions and specification, and the user interface development is a direct continuation from the conceptual modelling stage. As such, practice, engaging these methods, is considered the most appropriate choice.

4.6 Ethical Approval

In order to conduct this research, ethical approval has been granted by the University of Salford. The process involved submitting an application which covered all aspects of the proposed investigation. This included the research field and the topic of study, the ethical considerations of participants, risk assessment, participant information sheets and participant consent. The instruments were considered, including the interview and survey questions. Data protection is also a requirement with consideration given to data storage and anonymisation.

Ethical approval was granted for all aspects of this research, and examples of participant information and consent forms can be found within the appendix of this thesis (see: E).

4.7 Research Methodology Summary

This chapter has described the research philosophy underpinning this study and the associated methodological approach. It presented the pragmatic lens through which this research has been considered, the utilisation of mixed methods, and a strategy which incorporated four distinct phases of data collection, each related to the research questions and subsequent areas of investigation. Component one described an Artefact Catalogue comprising historical and contemporary projects and technologies, critically analysed to develop a better understanding of decentralised artefact development in practice. Component two described a Public Survey developed to gain a better understanding of attitudes towards data privacy, disclosure and control. The survey results are analysed through both descriptive and statistical analysis to provide valuable data relative to not only attitudes and behaviour, but also insights into potential strategies for product development and adoption. Component three outlined a number of phases of consultation with experts in the field. The first considered the overarching decentralised space. The second considered the narrower view of Self-Sovereign Technology, and finally, the third considered opinion, not from experts within the decentralised field, but with individuals from within the usability and user experience sphere. The data gathered

throughout these phases was analysed using a Thematic Content Analysis, resulting in a presentation of themes supported by related quotation. The final component considered a practice led element, designed to extend current Self-Sovereign Identity prototypes to scale, exploring the complexity of interaction through critical analysis by engaging expertise by way of focus group. The chapter concluded with a brief explanation of the ethical approval required and obtained for this research.

In the following chapter, the results of the four elements of data gathering and subsequent analysis are presented.

Chapter 5: Results

This chapter presents the results of the four components of primary data gathering described in the previous chapter. The first section presents findings from the Artefact Catalogue. The classification of artefacts is communicated through detailed categories. The development of the Evaluation Matrix is highlighted, describing the measures used. A visual sample of the overall catalogue is illustrated to demonstrate the colour coding and recorded structure. A visual mapping of artefact categories charts decentralised development chronologically. The results highlight the trajectory of the domain, significant contributions, issues of complexity and product value. A sample of the Artefact Catalogue can be found in the appendix of this thesis (see: A1).

The next component presented is the Public Survey, an overview of the sample outcome is communicated, followed by a description of the demographic, and a table of descriptive statistical outcomes on a per question basis. The validity of Likert scales is outlined, followed by an overview of the longitudinal component. Finally, significant descriptive statistics are noted, alongside relative correlations and comparisons. A full illustration of detailed survey results and statistical analysis can be found in the appendix of this document (see: B4).

The next component presented is the Expert Interviews. The staged process is reiterated together with an overview of participants. Initial phase themes are presented, and their impact on further stages is communicated. Combined stages are then considered with an overview of the Thematic Content Analysis conducted. Defined thematic categories are then presented, supported by visual mapping and verbatim comment.

Finally, the Practice Led Component is presented. The Conceptual Modelling process is illustrated in stages supported through simplified tables. The development of the preliminary user interface is illustrated through a complete graphical representation. A critical analysis

follows, together with information of participation and findings from the evaluation Focus Group. Full illustrations of the Conceptual Modelling Tables, together with pointers to full-scale User Interface and purpose-built web artefact, can be found in the appendix of this thesis (see: D).

5.1 Artefact Catalogue

This section presents the results of the Artefact Catalogue which considers historical projects, start-up companies, applications, technologies and technical standards relating to the Decentralised Internet. These elements are collectively referred to as artefacts. Artefacts have been sourced and collated within a structured table. Following an assessment of validity relative to the Decentralised Internet, the artefacts are categorised. The artefacts have then been subjected to a critical analysis through an Evaluation Matrix, to determine a number of factors: the maturity, state, originality, value, complexity and inclusion of identity. The artefacts have also been considered in a historical context, with consideration given to their future value as concepts, and their relevance to various domains. In considering these factors, this component generates a valuable, cross sectional view of decentralised technologies over time, and importantly indicates the capability and trajectory of the domain.

It is important to reiterate that this component of research does not aim to build an exhaustive catalogue of artefact analysis. This is considered impractical and outside of the scope of this study. Instead this component aims to build a broad catalogue, comprising a framework of categorisation and valuation, to act as a resource for further investigation of variables. There now follows a summary of the artefacts considered, the categorisation of entities, and their relevance to the field. This is supported by a chronological visual representation charting decentralised development, and an example of the Artefact Catalogue in table form. The results highlight the trajectory of development, significant contributions to the field, issues of complexity, and product value.

5.2.1 Artefact Catalogue Results

A total of 189 projects and technologies have been reviewed, of which 125 are considered to align with the principles of a Decentralised Internet. The inception of artefacts considered span a period from 1998 through until the present day. The artefacts have been compiled into a structured table, through which analysis has been undertaken utilising an Evaluation Matrix. The Matrix considers the following elements and scores the artefact from 0-5: Maturity, State, Originality, Value Proposition, Mental Model, Identity, Blockchain, Technology Rejuvenation. In addition, each artefact is considered relative to its value in the following domains: Individual, Commercial, Societal.

Table 3 is a sectional example of the catalogue highlighting artefacts relating to SSI technologies. A full listing of the catalogue can be found in appendix of this thesis (see: A1).

			Inception	Position	Location	Employees	Decentralised	Maturity	Alive	Original	Value Proposition	Mental Model	Identity	Blockchain	Technology Rejuvenation	TOTAL	Individual	Commercial	Societal		
								0	1	2	3	4	5								
SSI TECHNOLOGY																					
CORE	SSI / FULL SYSTEM / INFRASTRUCTURE	SOVRIN	8-Jul-05	Active	US	No data	5	3	5	5	3	2	5	5	NA		5	5	5	The Sovrin Network is a new standard for digital identity – designed to bring the trust, personal control, and ease-of-use of analog IDs – like driver's licenses and ID cards – to the Internet.	https://sovrin.org/
CORE	SSI / FULL SSI SYSTEM / INFRASTRUCTURE	VERES ONE	No data	Active	US	No data	5	3	5	3	2	2	5	5	N/A		5	4	4	Veres One is a fit-for-purpose blockchain optimized for identity on the Web	http://www.veres.one
CORE	SSI / FULL SSI SYSTEM / INFRASTRUCTURE	UPORT	2016	Active	US	1-10	4	3	4	4	1	2	5	5	N/A		3	3	3	Uport is a full spectrum SSI solution built around Ethereum	http://www.uport.me
CORE	SSI/CENTRALISED HUB	BLOCKSTACK	2013	Active	US	11-50	3	3	5	4	4	4	4	5	NA		5	4	4	Blockstack is a decentralized computing platform that puts users in control of their data and identity. Apps built on Blockstack make data breaches and trust violations an antiquated notion.	https://blockstack.org
CORE	SSI/LIMITED IDENTITY MECHANISUM	Blockpass	2017	Active	No data	No data	2	3	5	3	2	2	5	2	3		3	3	2	Participant can verify their identity by entering details and the submitting password and utility bill information > Once this has been verified I assume by a human being looking at the photgraph and	https://www.blockpass.org/2018/07/06/step-by-step-user-guide-to-activating-pass-tokens/
OI	IDENITY / CRYPTO / SOCIAL	Pillar Project	2017	Active	UK	11-50	3	3	3	4	3	3	4	4	N/A		4	2	4	This is an interesting project in which Idenity is built into a cryto wallet that includes a Social Element.	https://pillarproject.io/wallet
CORE	SSI / COMPONENT	Connect Me / Evernym	2013	Active	US	51-100	4	2	4	3	3	2	5	5	N/A		3	3	3	Evernyms Wallet application based around Sover	https://www.connect.me
OI	SSI/LIMITED IDENTITY MECHANISUM	Blockcerts	2016	Active	Canada	101-250	3	2			1	2	5	5	N/A					Seems to be a development community for now… With an explination of SSI for what seems like qualifications.	https://www.blockcerts.org
OI	SSI/LIMITED IDENTITY MECHANISUM	Lifeid	2017	Active	US	1-10	2	1												Very early, mjust seems to be a website…	https://lifeid.io
OI	SSI/LIMITED IDENTITY MECHANISUM	Securekey verified.me	No data	Active	Canada	No data	3	2	4	3	1	2	5	5	N/A		3	3	3	Embrionic SSI system that promises all triongs, back end system and application, poor reveiws.	https://securekey.com/digital-id-ecosystem/
CORE	SSI/LIMITED IDENTITY MECHANISUM	Civic	No data	No data	No data	No data	3	2	4	3	1	2	5	5	N/A		3	3	3	Our Secure ID product allows you to securely provide authenticated identity to partner websites for account sign up and access, including technologies that no longer require legacy usernames and passwords.	https://www.civic.com/app/
CORE	SSI/LIMITED IDENTITY MECHANISUM	Shocard	2015	Active	No data	No data	3	2	4	3	1	2	5	5	N/A		3	4	3	Seems to be another SSI system for limited idenity but aimed more at the commercial application.	https://shocard.com
CORE	SSI/LIMITED IDENTITY MECHANISUM	Dominode	2017	Active	No data	11-50	3	2	4	3	1	2	5	5	N/A		3	3	3	Seems to be another SSI system for limited idenity at a very early stage, aimed it would seem at gaming, gambling…	https://dominode.com

Table 3: Sectional example of Artefact Catalogue

5.2.2 Artefact Categorisation and Critical Analysis

Through the examination of artefacts, the following categories have been identified:

- Privacy Protection
- Personal Data Store with Basic Applications
- Personal Data Store with Extended Capabilities
- Collation Personal Data Stores
- VRM Principles 'Vendor Relationship Manager'
- Decentralised Principles 'Applications'
- Intent Casting Applications 'Innovative, Functional Trade, Collective'
- Personal Assistant / Agent
- Personal Data Marketing Tool
- SSI / Full System Infrastructure
- SSI / Limited Identity Mechanism
- SSI / Centralised Hub
- SSI / Component
- DAPPS 'Decentralised Applications'
- Decentralised / Distributed Storage
- Decentralised Stack Software Component or Standard

There follows a detailed description of each category together with examples and analysis:

5.2.2.1 Privacy Protection

Privacy protection relates to stand alone browser applications, browser plugins or Virtual Personal Networks, that allow a participant to block cookies, prevent tracking, disrupt data gathering, and traverse the Internet with relative anonymity. Although these technologies are at the peripheral of what would be considered decentralised offerings, they disrupt and protect, and potentially prevent many of the centralised surveillance mechanisms which decentralised technologies are conceived to counter. Exemplifiers within this category include Ad and Tracker blockers *Adblock* (2009) and *Ghostery* (2017), tunnelled browsers including

Tor (2019) and *Brave* (2016), and VPN providers such as *CyberGhost VPN* (2004) and *Express VPN* (2009).

5.2.2.2 Personal Data Store with Basic Applications

This category relates in the most part, to early efforts to provide decentralised storage and applications in the home. Such as a personal server, in some cases running on a small low-cost device or custom Blackbox microcomputer. This server runs an operating system that supports decentralised applications. Exemplifiers within this category include: *ArcOs* 'Now Inactive', *Personal* 'Now Inactive' and *OwnCloud* (2010).

5.2.2.3 Personal Data Store with Extended Capabilities

Extended Capabilities within Personal Data Stores relate to offerings that afford secure storage within the home, on a secure cloud service or in a distributed manner. They also include technologies that engage other more contemporary forms of personal data formats, or those technologies that offer functionality or capabilities that are progressing towards what could be considered Self-Sovereign Identity. Exemplifiers within this category include *FreedomBox* (Freedombox Foundation, 2011) and *Databox* (2018), that allow the user to gather, and disseminate their personal data on their own terms, and *Mydex* (2007), which integrates the concept of verifiable credential pre-blockchain.

5.2.2.4 Collation Personal Data Stores

Collation relates to a personal data store that allows participants to bring together data from different sources. By understanding more about their lives and activities, the data can be disseminated for their own gains to third parties from what is considered to be an ethical source. This may be data created through centralised engagement with service providers, or data generated independently and from government sources. Exemplifiers within this category include: *Digi.me*, (2013), *Meeco*, (2012), *Cozycloud*, (2012).

5.2.2.5 VRM Principles 'Vendor Relationship Manager'

This category relates to artefacts and organisations that are designed to align broadly with the principles of Vender Relationship Management (VRM), (see: 2.4.9). VRM is the opposite to customer relationship management. It empowers the participant with the capability to control and leverage personal data for commercial transactions. This category highlights artefacts that innovate existing business models and mechanisms established by contemporary ecommerce. Exemplifiers of organisation and application within this category include: *OpenBazaar*, (2016), *HiProject*, (2017), *QIY*, (2017), *Digital Me*, (2018).

5.2.2.6 Intent Casting Applications

Intent Casting is in essence a sub-category of VRM. It represents a specific mechanism by which a single individual or collective, can signify their desire for a product or service. The participant broadcasts their intention to the market. Vendors then respond to this signalling and a negotiation can commence.

5.2.2.7 Innovative Intent Casting

There are attempts at what could be considered to be a more complex in-depth development of intent casting applications. They allow a participant to declare an interest in a product line or make an offer or highest bid for a service. Though models are interesting from a VRM prospective, many seem rather centralised and convoluted, and in many cases would seem to have expired. Exemplifiers within this category include: *Greentoe*, (2012), *Get Magic*, (2016), *Intently*, (2012). Although in many cases Intent Casting artefacts no longer appear to be active, a great deal of consideration and work would seem to have gone into many of these projects. There is much to learn from them, in terms of concept and interactions. Many of these projects are considered valued and viable for further investigation, when considered in the light of contemporary technology developments.

5.2.2.8 Intent Casting for Local Trades and Services

The intent casting applications connecting local trades and services across North America and Europe, would in many cases seem to be active, having found a sustainable business model. Although centralised in most respects, the notion of Intent Casting has been achieved, and has proven an accessible value proposition in this context. An exemplifier within this category is: *Home Advisor*, (2017).

5.2.2.9 Collective Intent Casting

A relevant and valuable concept is that of Collective Intent Casting, a group of individuals act collectively to negotiate the best deal for a product, service or utility. Although the example cited below would seem to have expired, investigation of now redundant projects suggests that developers have dealt with issues of complex collective interactions, hierarchy, collective decision making, contracts and automated negotiation. For this reason, they hold great value for further research. An exemplifier within this category is: *Crowd Spending,* (2014).

5.2.2.10 Personal Assistant

This category relates to artefacts that provide assistance in order to get a task completed, or to streamline transactions. To date, many of the artefacts have a human interacting with data, as part of the workflow, assessing an individual's requirements before identifying a solution and supposedly negotiating on the participants behalf. In many respects this is a branch of intent casting. It is centralised in nature and would seem a steppingstone towards future AI driven systems. An exemplifier within this category is: *MyTime*, (2014).

5.2.2.11 Personal Assistant AI

Although the AI assistants recognised in this category are not in any way decentralised. It is recognised that in contemporary decentralised models, the complexity of interaction and decision making may require a degree of automation. As such, attention is paid to the peripheral development of pure AI based personal assistants and technologies. Exemplifiers within this category include: *Alexa*, (2014), *Watson,* (2011), *Siri,* (2011).

5.2.2.12 Personal Data Marketing Tool

Artefacts within this category are designed to allow participants to monetise their personal data. This can be achieved directly by simply providing access to data. Or more recently, through more complex mechanisms that allow individuals to collate data from a spectrum of centralised sources, before agreeing terms to allow third party access, usually for a small monitory return. Exemplifiers within this category include: *Datacoup,* (2017), *Citizenme*, (2017).

5.2.2.13 SSI / Full System Infrastructure

A Full System Infrastructure, for a Self-Sovereign Identity, relates to a collection of technologies providing a platform for emerging standards. These systems encapsulate the full array of required elements for SSI: Wallet and Agents, Issuer and Verifier code, and distributed Ledger and Blockchain Technology. These artefacts, organisations and companies are considered to be at the cutting edge of the decentralised realm, attempting to build an eco-system from the ground up. Exemplifiers within this category include: *Sovrin,* (2016), *Veres One,* (2019), *Uport,* (2018).

5.2.2.14 SSI / Limited Identity Mechanism

At the time of writing, *November 2019*, many artefacts are focused on using blockchain technologies, that offer a mechanism for proving identity and / or collecting credentials for a limited application. These applications are either aimed at the individual or targeted at corporate business use. These artefacts are very commercially focused, and aim to monetise the existing ecosystem, with most focusing on high value and high friction domains. Exemplifiers within this category include: *Civic,* (2019), *Shocard,* (2019), *Dominode,* (2019).

5.2.2.15 SSI / Centralised Hub

This category considers artefacts that focus on the development of Self-Sovereign Identity Hubs and Decentralised Applications. These examples would seem to be a semi-decentralised offering, where the core identities offered are reliant on a blockchain which is ultimately controlled by a centralised organisation. Equally, the applications that are developed for the ecosystem, and are tied to the APIs made available by the centralised system. These early examples would seem to be an attempt to find a sustainable business model for a semi-decentralised ecosystem. Exemplifiers within this category include: *Hub of All Things,* (2017), and *Blockstack,* (2018).

5.2.2.16 DAPPS 'Decentralised Applications'

Decentralised applications are developed with an architecture which separates the application from the data. In many cases decentralised applications are developed around a system which offers decentralised storage which, in most cases, is tied to an identity mechanism or system. In many cases identified within this category, the value offered is unclear, other than the application being a clone of a centralised service. This in many respects may be due to the limited functionality offered by the contemporary supporting systems. Exemplifiers within this category include: *Graphite,* (2018), *Dmail,* (2018), *SpringRole,* (2018), *Bitpatron,* (2018).

5.2.2.17 Decentralised / Distributed Storage

This category relates to technologies which allow for decentralised, encrypted or distributed storage. These systems do away with the centralised storage of data, and instead distribute encrypted data across the network through a series of nodes. This is very similar to peer to peer file sharing technology. This refined technology is now a reality and is being incorporated into many of the early decentralised models. Exemplifiers within this category include: *Madesafe,* (2016), *IPFS,* (2016), *Swarm,* (2017).

5.2.2.18 Decentralised Stack Software Component or Standard

This category relates to any technology which forms part of a wider decentralised system, stack elements or open standards. Within this domain there is a reliance on established cryptography, public private key, digital signatures and the evolution of the Blockchain, token-based systems that have allowed for federated identity, access control and the semantic web. This considers methods and standards relating to the creation and management of decentralised Identifiers, Verifiable Credentials and Zero Knowledge Proofs. Exemplifiers within this category include: *OpenID,* (2020), *OAuth,* (2020), *UMA,* (2019), *SOLID,* (2017) *DID,* (2020) *DID-Auth,* (2020), *Verifiable Credentials Working Group,* (VCWG, 2020)

5.2.3 Domain Mapping

Figure 18 offers a visual mapping of the Artefact Catalogue, illustrating a loose chronology, defined categories, artefacts and technologies, it endeavours to communicate the decentralised landscape and importantly the domains trajectory.

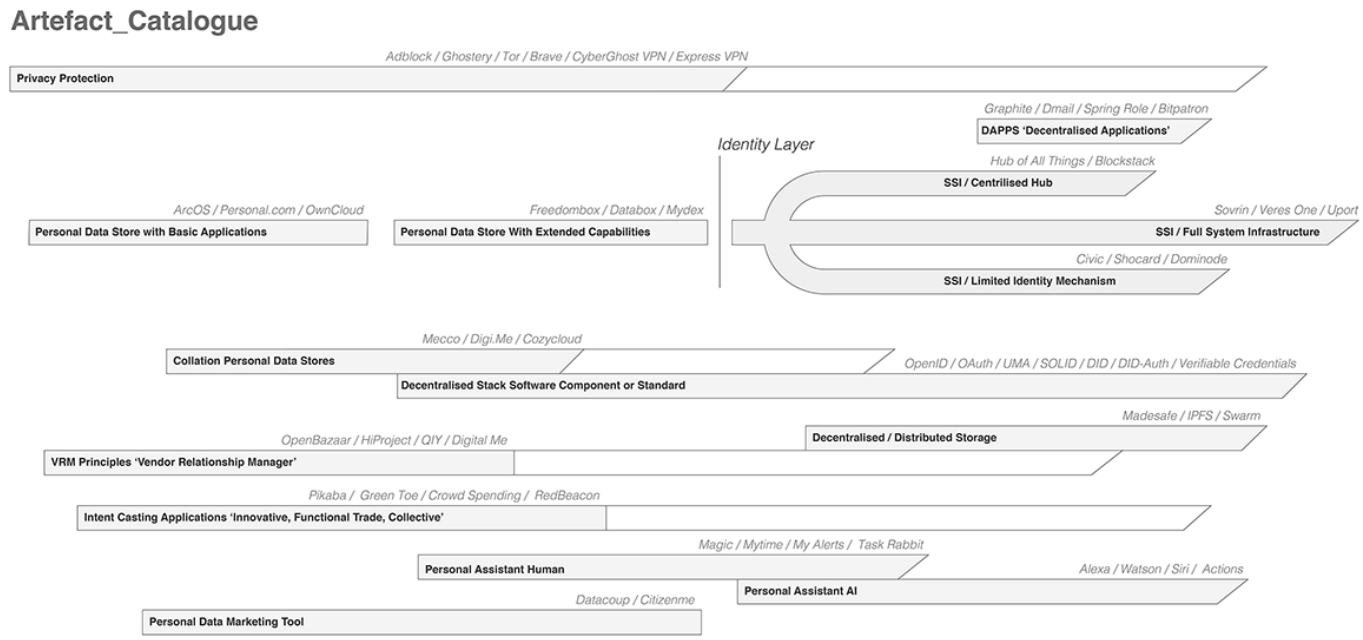


Figure 18: Artefact Catalogue Mapping

The narrative unfolds with passive, privacy protection add-ons and dedicated privacy enhanced browsers leading to VPN. These technologies are relative and contribute building blocks to the overall decentralised picture. The first evolution of Personal Data Stores, though complex to configure, and lacking in value proposition outside of privacy, established a model of data management, control, storage and the use of decentralised applications. The evolution of the data store explores a richer complexity in the way data can be utilised for the benefit of the participant. Also, the means of data storage is expanded, and the advent of credential and verification emerges. Overtime, VRM projects have been conceptualised to allow a participant to cast intention, providing quality data, flipping the relationship between participant and vendor. Although many of these concepts are defunct, they have great value as an R&D

resource for an identity layer rejuvenation. Collation data stores allow for the collection of dynamic personal data from third parties, building a richer view of a participant's digital footprint, providing vendors with ethically sourced data streams. Underneath, technology develops through encryption, federated identity protocols, concepts around linked data, the blockchain, smart contracts, distributed storage and DAPPS (decentralised applications). Progressively the identity layer develops with supporting standards. This has led to the probing of different models of decentralisation with varied levels of control. The most significant of which are listed in the following paragraphs.

5.3.4 Important Findings Drawn from The Artefact Catalogue

5.3.4.1 The Impact of the Identity Layer

It is evident that the Identity Layer has energised the development of a Decentralised Internet and conceptually rebalanced the equation. The prospect of immutable identity under the control of the participant, strips the dominant decentralised forces of the monopoly of identity ownership. It goes someway to restoring genuine peer to peer relationships and transactions. Alongside the Blockchain technology that acts as a foundation layer, there are a number of evolving standards that are designed to support its functionality, the forging of relationships, the authentication of digital credentials, and peripheral tools for the management of the ecosystem.

5.3.4.2 The Driving Trend Towards the Identity Layer

The current trajectory and underlying concept currently driving decentralisation is the Identity Layer, facilitated through a decentralised ledger, the development of standards around verifiable credentials, and a decentralised public key infrastructure. This current trend sees endeavours to establish the foundation components required, which include, according to Kaliya Young: Wallet and Agent applications, Issuer and Verifier code, and Distributed Ledger

and Blockchain Technology (K Young 2018, Personal Communication, 21st Nov). At present, Identity Layer artefacts, combine these elements within three core trends.

1 Self-Sovereign Identity: This is an identity that is controlled and owned by the user. It does not rely on any administrative authority and can never be taken away. This identity is part of an ecosystem, which allows multiple unique relationships, and the exchange and recording of data and verifiable credential exchange, within a web of trust. The concept of SSI satisfies the principles of a Decentralised Internet and offers the participant full network agency.

2 The Centralised Hub: This is a model in which there is a component of the system that is still centralised or controlled by an organisation. This retained identity is either offered through a hub application or supported through infrastructure, that is a Blockchain. However, these artefacts are structured in a manner that retains a degree of control, through which a business model can be constructed.

3 Decentralised Trusted Identity: This is a model that is provided by a centralised organisation. It performs identity proofing of users based upon existing trusted credentials. It records identity information through a distributed ledger for later validation by third parties for specific applications. This model is developed around potentially high value, high friction commercial applications.

5.3.4.3 Probing for Sustainable Business Models

Across the artefacts catalogued, we observe how developers are searching for business models to support sustainable technology. The reality is, there has to be a stream of revenue, and in a decentralised context, this is a difficult problem to solve. Early VRM projects, functional trade intent casting, and personal assistants are all essentially centralised. The model allows the participant to leverage personal data to their advantage, but the underlying business model is that of centralised control. Collation personal data stores are, in fact data, intermediaries disrupting the centralised model, by offering a secure service for data gathering

and dissemination based on the participants terms. The emergence of the decentralised ledger has seen innovative methods of revenue generation, such as decentralised storage. This allows micro payments via crypto currency to be made to the holders of data, and the generators of content based on the demand. As the Identity Layer evolves, innovators are probing for sustainable models with various levels of exposure and control of the underlying system. Centralised Hubs provide the Blockchain, identity generation and storage, and functional agent software for the participant. They are establishing an eco-system for developers, to build and monetise decentralised applications. Decentralised Trusted Identity Providers focus on high value, high friction commercial applications for very specific needs. Self-Sovereign Identity development is made up of organisations who are developing an ecosystem which is truly decentralised, at the same time, in parallel, developing and building applications and technologies to capitalise on a decentralised ecosystem once established. The trajectory is clear, and there is no shortage of attempts to capitalise on decentralised development at every step of the way. It remains to be seen if the goal of a truly decentralised ecosystem can be realised, to a point where it is self-sustaining, where the model switches from surveillance capitalism, to a model of decentralised services based on a sovereign data ecosystem.

5.3.4.4 Value

Across many of the artefacts considered, it is challenging to find significant current value proposition. This is of concern in the context of the *Diffusion of Innovation* and adoption. The potential for value would seem to be centred around several concepts; privacy, economic gains, seamless transaction, transparency, and the quantifiable self. There are some endeavours to explore the social benefits of data sharing. Privacy is currently the dominant value proposition, with the status-quo offering a measure of control and utilisation of personal information. There are many concepts that are considered to hold genuine value in early examples of VRM and these are discussed in the next section. The value in economic gain is

limited at present to the direct sale of personal data, which in many ways is considered to have ethical implications. The potential value found in VRM, with seamless transaction and social benefits of data sharing, is still limited. This may be a consequence of the current limited capability and functionality at the Identity Layer as this is, arguably, something which will be resolved if a full data ecosystem through Self-Sovereign Identity is realised.

5.3.4.5 The Value In VRM

A great deal of innovation can be found in what might be considered the first wave of Vendor Relationship Management artefacts. The notion of intent casting and the power of negotiation envisaged through the development of relationships, Verifiable Credential and the providing of accurate market intelligence to vendors, demonstrate a clear and balanced value proposition. Many of the concepts of VRM which were initially semi-centralised, could potentially flourish within a genuine decentralised data ecosystem. These models can act as a valuable resource for designers and developers as the domain matures.

5.3.4.6 The Mental Model

Across the considered artefacts, the notion of mental model, or internal representation and the potential for user understanding is varied. Understanding the value through the passive application of privacy protection is straight forward. The concept of private encrypted communication is common parlance, and the concept of selling personal data, collectively biding for services, and the collation of personal data to build a better individual understanding of oneself, are arguably, easy to comprehend. A personal data store with privacy respecting applications where the data sits in a box, in the participants property, shielded from observation is again understandable. As we step further into the Identity layer, many of the concepts may become abstract and lack metaphor and existing mapping. Multiple identifiers, pairwise pseudonymous relationships, verifiable credential, a distributed ledger, distributed storage and zero knowledge proof, may in practice pose a problem to the user, and inadvertently exclude,

rather than emancipate. A consistent mental model of decentralised technologies is something that is critical to this research, and is something that will arguably be required, if mainstream adoption off such technologies is to be achieved.

This concludes the results section for the Artifice Catalogue. The following section Presents the Public Survey, the process, outcomes and statistical analysis.

5.3 Public Survey Results

The following paragraphs present the results of the Public Survey. The survey has been developed to investigate attitudes towards Internet usage, data privacy, the disclosure and secondary usage of personal data, engagement with activities and opportunities to protect and control personal information. Details regarding participants and demographics are conveyed followed by descriptive statistics drawn from the results. Later in the chapter correlation and comparison statistics are presented, together with the results from an unexpected longitudinal component.

5.3.1 General Information, Numbers and Demographic

A Public Survey administered through an Internet Mediated Questionnaire, has been undertaken in line with the defined methodology and survey method plan. In total 295 surveys have been completed. In terms of the demographic, 62% of participants were male, while 34.6% were female, with the remainder preferring not to answer. The age of participants resulted in 52.5% aged 21 and under, 20.3% aged 22 to 34, 12.9% ages 35 to 44, and 9.8% being aged 45 to 54, and 3.4% being 55 or above.

5.3.2 Descriptive Statistics

Within the survey there is a total of 51 relative questions comprising Likert Items and Binary Choice questions. In terms of the descriptive statistics, a visual representation within this section is considered inappropriate. Instead each question is listed with a mean outcome

together with the standard deviation. It is recognised that a mean result may not be considered best practice with what can be argued to be non-continuous data. The following section is intended to act as an indicator of general treads within the data. A full charted representation of the descriptive statistics can be found in the appendix of this thesis (see: A1).

The descriptive results are listed by question:

Understanding of Personal Data and its Value
Likert Elements: Scaled 1-5

Q16 In the context of your Internet activity, how confident are you that you really understand what personal data is? (M = 3.45, SD = .963)
Q17 How aware are you that your personal data might be gathered while you use the Internet? (M = 3.99, SD = .844)
Q18 How aware are you that your browsing patterns might be tracked as you move across different parts of the Internet? (M = 3.97, SD = .952)
Q19 Are you aware that your personal data is valuable to many different types of companies and individuals? (M = 4.12, SD = .891)
Q20 How aware are you that collected personal data is actively bought and sold by data brokers on the open market? (M = 3.58, SD = 1.28)
Q21 Are you aware that personal data has great value and potential not only for marketing purposes but also for wider society? (M = 3.59, SD = 1.12)
Q22 How aware are you of the IOT and the implications this technology might have on the amount of data different objects produce? (M = 3.59, SD = 1.12)

Comfort Level with Network Engagement
Likert Elements: Scaled 1-5

Q23 When you visit or use online services that are free to use 'Facebook, Google, Twitter' it can be argued that you pay for this service with the disclosure of your personal data and information. Do you think the disclosure of your personal data and information is a fair exchange for the services you receive? (M = 2.46, SD = 1.09)
Q24 How would you describe the amount of control you feel you have over the amount of data collected about you online? (M = 2.37, SD = 1.00)
Q25 How confident are you that the information you disclose will be kept private and secure? (M = 2.12, SD = 1.05)
Q26 If motivated how difficult do you think it would be for a person or organisation to learn details about your past you would rather keep private? (M = 2.23, SD = 1.12)
Q27 Overall, how happy are you that organisations and companies collect your personal data? (M = 2.24, SD = 1.04)

Specific Concerns
Q28 What most concerns you about sharing your personal data?

As shown in Fig 16, 68.5% report the main concern as: *'That I do not have control over how my data is shared or who it is shared with*.'

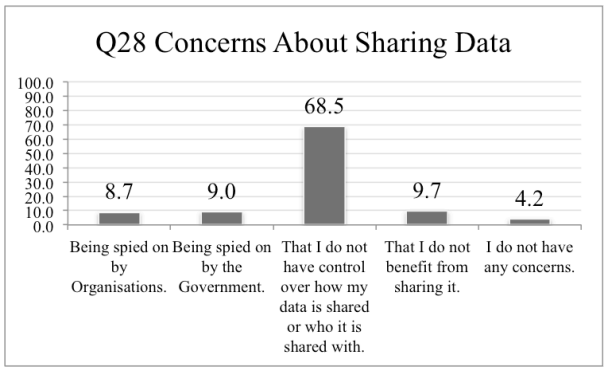


Figure 19: Survey Q28, Concerns About Sharing Data.

Perception of Importance

Q 29 How much do you care that only you, and those you authorise, should have access to the following information?

Likert Elements: Scaled 1-5

a: Content of Your Email (M = 4.42, SD = 0.97)
b: People You Exchange Email With (M = 4.19, SD = 1.02)
c: Content of Files You Download (M = 4.19, SD = 1.05)
d: Located When You Use the Internet (M = 4.07, SD = 1.16)
e: Content of Your Online Chats / Hangout (M = 4.32, SD = 1.01)
f: Websites You Browse (M = 3.79, SD = 1.19)
g: The Searches You Perform (M = 3.73, SD = 1.19)
h: Apps or Programs You Use (M =3.50, SD = 1.22)
i: Times of The Day You Are Online (M =3.35, SD = 1.40)

Effort Made to Protect Privacy
Q 30 Do you engage in any of the strategies below in order to be less visible online?
Forced Binary: Scaled 0-1

a: Clearing Cookies and Browser History (No = 30.7%, Yes = 69.3%)
b: Deleted / Edited Something You Posted in the Past (No = 42.3%, Yes = 69.3%)
c: Set Browser to Disable or Turn Off Cookies (No = 70.3%, Yes = 29.7%)
d: Not Used Website Asked Me for Real Name (No = 75.3%, Yes = 24.7%)
e: Used Temporary Username / Email Address (No = 56.7%, Yes = 43.3%)
f: Post Comments Without Revealing Who You Are (No = 76.5%, Yes = 23.5%)
g: Asked Somebody to Remove Something Posted (No = 75.8%, Yes = 24.2%)
h: Tried to Mask Your Identity (No = 79.2%, Yes = 20.8%)
i: Used Public Computer to Browse Anonymously (No = 87.3%, Yes = 12.7%)
j: Used Fake Name Untraceable Username (No = 72.6%, Yes = 27.4%)
k: Encrypted Your Communications (No = 81.9%, Yes = 18.1%)
l: Used Services to Browse the Web Anonymously (No = 64.5%, Yes = 35.5%)
m: Given Inaccurate Info About Yourself (No = 57.2%, Yes = 42.8%)

Understand What Participants actually do at this time to Safeguard and Protect Their Personal Data
Mixture - Likert Elements: Scaled 1-5 / Forced Binary: Scaled 0-1

Q31 Have you ever considered controlling access to your personal data, information and identity online? (No = 36.9%, Yes = 63.1%)
Q32 If technology was available to allow you to better manage and control access to your own personal data would you be interested in using it? (M =4.06, SD = 1.40)
Q33 Do you think your concerns about your personal data online 'if any' would motivate you sufficiently to manage part or all of your personal data if the capability existed? (No = 68.9.9%, Yes = 31.1%)
Q34 Would you feel comfortable in allowing another person or organisation to manage your personal data for you. (M =2.53, SD = 1.05)
Q35 Would you feel comfortable in allowing an Artificial Intelligence to manage your personal data. (M =2.37, SD = 1.17)
Q36 If a service was available that would manage your personal data, and the benefits of that were clear is this something that you would consider paying for?
(M =2.53, SD = 1.10)
Q37 Which sector do you trust the most with your personal data?
Media / Retail / Telecoms / Public Sector / Financial / Utilities / Transport / Other

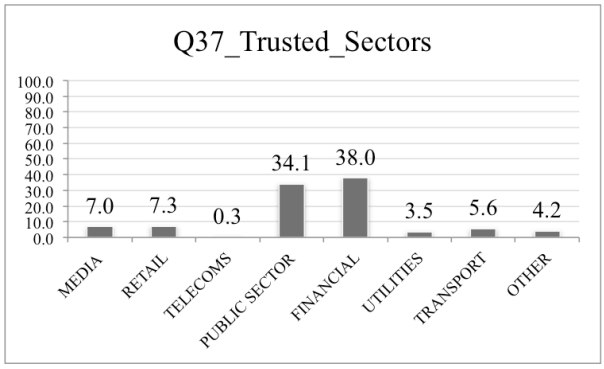


Figure 20: Survey Q37, Which sector do you Trust?

Q 38 Which one of the following options would most convince you to share your personal data? Benefit Economy / Improve Society / Improve Personnel / Paid for It / Other

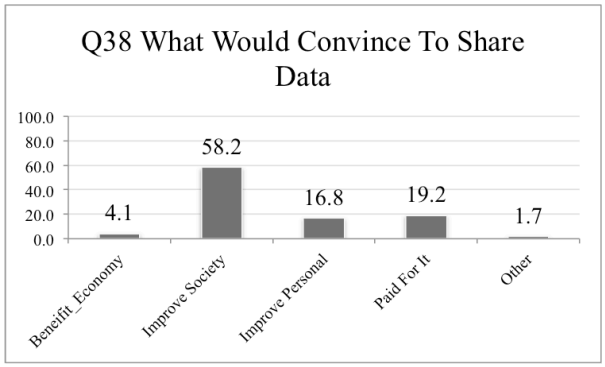


Figure 21: Survey Q38, What would convince you to share data ?

5.3.3 Comparison Information

A number of questions have been designed to act as comparison statistics to be used in further analysis, the following presents the results of these elements.

Mixture - Likert Elements: Scaled 1-5 / Forced Binary: Scaled 0-1

Q39 Do you think people should have the ability to use the Internet with complete anonymity? (No = 35.2%, Yes = 64.8%)
Q41 Do you read the Terms and Conditions when agreeing to a service online? (Yes =9.9%, Sometimes = 54.6%, Never = 35.5%)
Q42 In all honesty, at this moment in time, how concerned are you about the personal data you disclose online? (M =2.77, SD = 1.01)
Q43 Have you ever been the victim of what you would consider a fraud, breach or an abuse of personal data? (No = 71%, Yes = 29%)

5.3.4 Acceptable Reliability Likert Scales

The following scales have been defined from selected Likert items. Before any further analysis was conducted each scale has been checked for reliability in SPSS, using a Cronbach's Alpha measure of internal consistency. Any result higher the 0.7 is considered acceptable. The results of each test are now reported:

Likert Scales

A: Digital Fluency
Scale defined from questions: 13,14,15. Likert Elements Scaled 1-5
Cronbach's alpha showed the combined elements reached an unacceptable reliability, $\alpha = 0.483$. The Digital Fluency Likert Scale has been rejected.

B: Understanding of Personal Data and Its Value
Scale defined from questions: 16,17,18,19,20,21,22. Likert Elements: Scaled 1-5
Cronbach's alpha showed the combined elements reach acceptable reliability, $\alpha = 0.86$. All items appeared to be worthy of retention, resulting in a decrease in the alpha if deleted. The Mean result from this scale is as follows. ($M = 3.64$, $SD = .800$)

C: Comfort Level with Network Engagement
Scale defined from questions: 23,24,25,26,27. Likert Elements: Scaled 1-5
Cronbach's alpha showed the combined elements reach acceptable reliability, $\alpha = 0.77$. All items appeared to be worthy of retention, resulting in a decrease in the alpha if deleted. The Mean result from this scale is as follows. ($M = 2.28$, $SD = .775$)

D: Perception of Importance of Personal Data
Scale defined from questions: 29 a,b,c,d,e,f,g,h,i. Likert Elements: Scaled 1-5
Cronbach's alpha showed the combined elements reach acceptable reliability, $\alpha = 0.90$. All items appeared to be worthy of retention, resulting in a decrease in the alpha if deleted. The Mean result from this scale is as follows. ($M = 3.95$, $SD = .775$)

5.3.5 Force Binary Scales

The following scales have been defined from related forced binary questions, and act as both combination indicator and comparison statistic.

E: Effort Made to Protect Personal Data
Scale defined from questions: 30 a,b,c,d,e,f,g,h,i,j,k,l,m. Forced Binary 0-1 The Mean result from this scale is as follows. (M = .331, SD = .223)

F: Willingness to Engage Third Parties
Scale defined from questions: 24, 25 Forced Binary 0-1 The Mean result from this scale is as follows. (M = 2.44, SD = .966)

5.3.6 Likert Scales - Pearson's Correlation

Likert Scales have been analysed through a Pearson's Correlation statistical test to consider any relationships which may be present. The results of this test are illustrated in table 4, followed by a detailed outline of any test deemed to be significant. When considering the results, the following guidelines are used to evaluate any evidence of correlation.

- .00-.19 "very weak"
- .20-.39 "weak"
- .40-.59 "moderate"
- .60-.79 "strong"
- .80-1.0 "very strong"

(Evens, 1996)

As illustrated in table 4 and subsequent detailed description, only 2 tests, C and E, exhibit what can be considered a significant weak correlation.

		LS1_UNDERSTANDING_PD_VALUE	LS2_COMFORT_LEVEL_NETWORK_ENGAGEMENT	LS3_IMPORTANCE_OF_PD	EFFORT_MADE_TO_PROTECT	WILLINGNESS_ENGAGE_AI_3RD_PARTIES
LS1_UNDERSTANDING_PD_VALUE	Pearson Correlation	1	.150**	.056	.390**	.122*
	Sig. (2-tailed)		.010	.340	.000	.036
	N	293	293	293	293	293
LS2_COMFORT_LEVEL_NETWORK_ENGAGEMENT	Pearson Correlation	.150**	1	-.262**	-.092	.373**
	Sig. (2-tailed)	.010		.000	.117	.000
	N	293	293	293	293	293
LS3_IMPORTANCE_OF_PD	Pearson Correlation	.056	-.262**	1	.177**	-.149*
	Sig. (2-tailed)	.340	.000		.002	.010
	N	293	293	293	293	293
EFFORT_MADE_TO_PROTECT	Pearson Correlation	.390**	-.092	.177**	1	.059
	Sig. (2-tailed)	.000	.117	.002		.311
	N	293	293	293	293	293
WILLINGNESS_ENGAGE_AI_3RD_PARTIES	Pearson Correlation	.122*	.373**	-.149*	.059	1
	Sig. (2-tailed)	.036	.000	.010	.311	
	N	293	293	293	293	293

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 4: Likert Scales - Pearson's Correlation Test

A: Understanding of Personal Data and its Value / Comfort Level with Network Engagement

There is a significant positive relationship between the Understanding Of The Value Of Personal Data / Comfort Level With Network Engagement, r(293) =.150, p=.010

B: Understanding of Personal Data and its Value / Willingness to Engage 3RD Parties

There is a significant positive relationship between the Understanding Of The Value Of Personal Data / Willingness To Engage 3RD Parties, r(293) =.122, p=.036

C: Understanding of Personal Data and its Value / Effort Made to Protect Personal Data

There is a significant positive relationship between the Understanding Of The Value Of Personal Data / Effort Made To Protect Personal Data, r(293) =.390, p=.000

D: Comfort Level with Network Engagement / Perception of Importance of Personal Data

There is a significant negative relationship between the Comfort Level With Network Engagement / Perception Of Importance of Personal Data, r(293) =-.262, p=.000

E: Comfort Level with Network Engagement / Willingness to Engage 3RD Parties

There is a significant positive relationship between the Comfort Level With Network Engagement / Willingness To Engage 3RD Parties, r(293) =.373, p=.000

F: Perception of Importance of Personal Data / Effort Made to Protect Personal Data
There is a significant positive relationship between the Perception Of Importance of Personal Data / Effort Made To Protect Personal Data r(293) =.177, p=.002

G: Perception of Importance of Personal Data/Willingness to Engage 3RD Parties
There is a significant negative relationship between the Perception Of Importance of Personal Data / Effort Made To Protect Personal Data r(293) =-.149, p=.010

5.3.7 Ordinal Dependent - Independent Variable Comparison

In this section, the results of a variable comparison are reported. The defined Likert Scales have been compared with the independent variables: Gender, Suffered A Data Breach, Considered Controlling Personal Data, and Motivated to Manage Personal Data. An Independent Sample T Test is used to consider the probability that any difference between two groups is of significance, or due to random chance. This result is illustrated in the following results as a P Value. Anything under p=0.5 is considered significant, and anything over is considered random chance. This test also considers a Hedges'G value. This reports on the effect size of any variance between groups and can be evaluated through the following guide.

0.2 => small effect
0.5 => medium effect
0.8 => large effect.

The results also include a percentage value, which provides a clear indication of the size of any difference between the considered groups. Any result which is considered significant is highlighted within the following list of results:

A1: Gender / Understanding of Personal Data and Its Value
F: n102 (M =3.39, SD=.815) M: n183 (M =3.75, SD=.758) Male = +7.2%
Conditions; t(283)=-3.79, p=.000 Hedges' g = 0.46
A2: Gender / Comfort Level with Network Engagement
F: n102 (M = 2.11, SD=.662) M: n183 (M =2.36, SD=.798) Male = +5%
Conditions; t(283)=-2.62, p=.009 Hedges' g = 0.33
A3: Gender / Perception of Importance of Personal Data
F: n102 (M =3.98, SD=.828) M: n183 (M =3.91, SD=.873) Female = + 1.4%
Conditions; t(283)=.648, p=.519 Hedges' g = 0.08
A4: Gender / Effort Made to Protect Personal Data
F: n102 (M =.252, SD=.179) M: n183 (M =.359, SD=.221) Male = +10.7%
Conditions; t(283)=-4.17, p=.000 Hedges' g = 0.51
A5: Gender / Willingness to Engage 3RD Parties
F: n102 (M =2.33, SD=.867) M: n183 (M =2.51, SD=1.00) Male = +3.6%
Conditions; t(283)=-1.52, p=.128 Hedges' g = 0.18

B1: Suffered A Data Breach / Understanding of Personal Data and Its Value
N: n208 (M =3.66, SD=.760) Y: n85 (M =3.57, SD=.893) N = +2.5% No = +1.8%
Conditions; t(291)=.952, p=.342 Hedges' g = 0.11
B2: Suffered A Data Breach / Comfort Level with Network Engagement
N: n208 (M =2.33, SD=.749) Y: n85 (M =2.15, SD=.826) N = +8.3% No = +3.6%
Conditions; t(291)=1.82, p=.069 Hedges' g = 0.23
B3: Suffered A Data Breach / Perception of Importance of Personal Data
N: n208 (M =3.88, SD=.850) Y: n85 (M =4.11, SD=.850) Y= +5.9% Yes = +4.6%
Conditions; t(291)=-2.15, p=.032 Hedges' g = 0.27
B4: Suffered A Data Breach / Effort Made to Protect Personal Data
N: n208 (M =.326, SD=.215) Y: n85 (M =.343, SD=.245) Y= +5.2% Yes = +1.7%
Conditions; t(291)=-5.80, p=.562 Hedges' g = 0.07
B5: Suffered A Data Breach / Willingness to Engage 3RD Parties
N: n208 (M =2.45, SD=.980) Y: n85 (M =2.42, SD=.933) No = +0.6%
Conditions; t(291)=.286, p=.771 Hedges' g = 0.03

C1: Has the Participant Considered Controlling Personal Data / Understanding of Personal Data and its Value?
N: n108 (M =3.21, SD=.755) Y: n185 (M =3.88, SD=.720) Yes = +13.4%
Conditions; t(291)=-7.56, p=.000 Hedges' g = 0.91

C2: Has the Participant Considered Controlling Personal Data / Comfort Level with Network Engagement?
N: n108 (M =2.27, SD=.681) Y: n185 (M =2.28, SD=.826) Yes = +0.2%
Conditions; t(291)=-.073, p=.942 Hedges' g = 0.01

C3: Has the Participant Considered Controlling Personal Data / Perception of Importance of Personal Data?
N: n108 (M =3.79, SD=.838) Y: n185 (M =4.04, SD=.855) Yes = +5%
Conditions; t(291)=-2.34, p=.020 Hedges' g = 0.29

C4: Has the Participant Considered Controlling Personal Data / Effort Made to Protect Personal Data?
N: n108 (M =.236, SD=.162) Y: n185 (M =.386, SD=.236) Yes = +15%
Conditions; t(291)=-5.80, p=.000 Hedges' g = 0.74

C5: Has the Participant Considered Controlling Personal Data / Willingness to Engage 3RD Parties?
N: n108 (M =2.33, SD=.86) Y: n185 (M =2.51, SD=1.01) Yes = +3.6%
Conditions; t(291)=-1.50, p=.134 Hedges' g = 0.18

D1: Motivated to Manage Personal Data / Understanding of Personal Data and Its Value
N: n91 (M =3.46, SD=.754) Y: n202 (M =3.72, SD=.809) Yes = +5.2%
Conditions; t(291)=-2.60, p=.010 Hedges' g = 0.32

D2: Motivated to Manage Personal Data / Comfort Level with Network Engagement
N: n91 (M =2.61, SD=.636) Y: n202 (M =2.13, SD=.787) Yes = -9.6%
Conditions; t(291)=5.11, p=.000 Hedges' g = 0.64

D3: Motivated to Manage Personal Data / Perception of Importance of Personal Data
N: n91 (M =3.63, SD=.905) Y: n202 (M =4.09, SD=.794) Yes = +9.2%
Conditions; t(291)=-4.34, p=.000 Hedges' g = 0.55

D4: Motivated to Manage Personal Data / Effort Made to Protect Personal Data
N: n91 (M =.243, SD=.188) Y: n202 (M =.370, SD=.227) Yes = +12.7%
Conditions; t(291)=-4.64, p=.000 Hedges' g = 0.58

D5: Motivated to Manage Personal Data / Willingness to Engage 3RD Parties
N: n91 (M =2.52, SD=.849) Y: n202 (M =2.41, SD=1.01) No = +2.2%
Conditions; t(291)=-869, p=.385 Hedges' g = 0.11

5.3.8 Emergence of Longitudinal Data

The primary method used to gather responses to this survey involved approaching individuals in person and asking them to complete the survey using a handheld digital device. Tables and banners were erected at various locations across the University of Salford campus. Several digital devices where available, and individuals were approached in passing, and asked to participate. The target number was set at between 300 and 400 responses. It soon became very clear that reaching that number would be a challenge and require a considerable amount of time.

The first sessions where held over four days in December 2017, at the University of Salford's New Adelphi campus and its main library. During the first session around 100 responses where gathered. Due to work commitments, a second session of data gathering was not scheduled until April 2018. In March 2018 the Cambridge Analytica incident became public and received a considerable amount of media attention. As a result, the decision was made to delay the second data gathering session until September 2018. The rationale for this was to reduce the probability of the ongoing media exposure influencing the general survey responses.

Although a period of time has been placed between the initial data gathering activity and the final sessions, it is still considered acceptable to combine the results and analysis as originally planned. In addition, Likert scales from the first session of data gathering, n=100, and the second, n=200, have been compared to attempt to ascertain if the Cambridge Analytica event has changed or influenced public opinion across several variables. The following results are reported in the same way as the Ordinal Dependent - Independent Variable Comparison. Through an Independent Sample T Test, Hedges'G value and percentage difference. Results which are considered significant are highlighted.

Longitudinal Comparison Results

Comparison of the two groups via descriptive statistics
Group A: n=87 Before Cambridge Analytica
Group B: n=206 After Cambridge Analytica

E1: Comfort Level with Network Engagement
A: n=87 (M=2.11, SD=.665) B: n=206 (M=2.35, SD=.809) B = + 4.8%
Conditions; t(291)=-2.37, p=.018 Hedges' g = 0.31
E2: Effort Made to Protect Personal Data
A: n=87 (M=.258, SD=.180) B: n=206 (M=.361, SD=.233) B = + 20.6%
Conditions; t(291)=-3.70, p=.000 Hedges' g = 0.47
E3: Has the Participant Considered Controlling Personal Data
A: n=87 (M=.59, SD=.495) B: n=206 (M=.65, SD=.478) B = + 6%
Conditions; t(291)=--1.04, p=.299 Hedges' g = 0.12
E4: Motivated to Manage Personal Data
A: n=87 (M=.68, SD=.470) B: n=206 (M=.69, SD=.462) B = +1%
Conditions; t(291)=-2.70, p=.788 Hedges' g = 0.02
E5: In All Honesty How Concerned Are You About the Personal Data You Disclose
A: n=87 (M=2.76, SD=1.04). B: n=206 (M=2.77, SD=1.04) B = +0.2%
Conditions; t(290)=-1.23, p=.902 Hedges' g = 0.02

This concludes the results section for the Public Survey. The following section presents the results of the 3 stages of semi-structured interviews and focus group with experts, and the subsequent thematic analysis.

5.4 Expert Interviews Results

This section presents the results of the expert interviews conducted within this research. The interviews where undertaken in three stages. The first focused on seven individuals with in-depth knowledge of the decentralised domain. The second stage informed by the first, engaged a further nine individuals with specialist knowledge in the realm of *Digital Identity*. A final stage engaged ten focus group participants, each with extensive knowledge in the field of usability and interface design. The first phase explored the broad decentralised space with the objective of understanding the trend and direction of travel, the technological usability barriers, and opinions and views around sustainable adoption. A second phase of interviews focused on individuals from the decentralised space, who had active interest in self-sovereign technologies. These interviews were narrower is scope and focused specifically on user interaction and adoption. A third phase related to data gathered from the focus group conducted as part of the Practice Led Component of this research. As this data has value in the context of these expert interviews, it has been subjected to, and included in the same analytical process.

This section begins by presenting a key number of themes, which emerged from a basic thematic analysis of the first seven interviews with experts from the decentralised field. The section then progresses to present the process of in-depth thematic analysis conducted on all interviews combined. The section then illustrates a map of the themes generated during analysis, before moving to describe and detail each theme in turn. Finally, the section draws a summary of the knowledge compiled through the process.

5.4.1 Phase 1 Semi-Structured Interviews

Phase 1 involved conducting semi-structured interviews with prominent members of the decentralised community. The interviewees were a representative cross-section, able to comment on the broad ideology of decentralisation, project development and technology innovation. Each covered a different facet of the decentralised field, and there is representation

from both the UK and the US. Each interview was between 20-35 minutes in length. The aim of these interviews was to gain a clearer understanding of the decentralised space, while informing the research trajectory and practice-based component in line with the research design (see: 4.1).

The participants included:

- **Philip Sheldrake** - Author of The Business of Influence (Sheldrake, 2011), Partner, Euler Consultants Ltd, Founder of The Human Interface Project 'Hi-Project', PhD investigating Network Agency under Wendy Hall at Southampton University
- **Adrian Gropper MD** - CTO at Patient Privacy Rights, Key contributor to the Internet Identity Workshop, MIT, Harvard Medical School
- **Jeremy Rushton** – Founder Tiddy Wiki – Ex Head of Open Source Innovation, British Telecommunications
- **David Irvine** – Co Founder of the Maidsafe Network
- **David Alexander** - Co Founder of Mydex
- **John Laprise PhD** – Principle Consultant - Internet governance, public policy, technology strategy and technology forecasting
- **Michael Linton** – Originator of the term 'Local Exchange Trading System', Prominent figure in the world of community currencies

The 5 themes for investigation were as follows:

- Participants background and work within the decentralised field.
- The evolving trajectory of decentralised technologies and its associated application.
- The perceived barriers to adoption and the challenges ahead.
- User Journeys within a decentralised personal data framework.
- User interaction and interface issues.

A full transcript of phase one questions can be found in the appendix of this thesis (see: C3).

Once conducted, the interviews where transcribed and analysed through basic thematic analysis. Codes where applied and memoing was conducted, allowing patterns, themes and general consensus to be drawn.

The codes used during this first cycle of analysis were as follows:

- Issues around the adoption of decentralised technologies
- Developing concepts and innovations
- Public awareness
- Technology
- User interaction
- Value proposition
- Legislation and the Law

The key themes generated from stage one is presented below. Please note that this does not represent all of the information derived from the first phase of interviews, but those that are considered to be the dominant themes, which influenced the direction of this research at the time. A complete analysis of the phase one interviews, has been conducted and presented through the second cycle of thematic analysis, which combined all three sets of interviews.

5.4.1.1 Themes from the First Cycle of Thematic Analysis

Five dominant themes emerged, namely:

- The Decentralised Internet cannot be Marketed.
- Decentralisation must supersede to find Adoption.
- Technology is not the Issue.
- Getting to the Interface Layer.
- Identity and Blockchain technology unlock the Model.

Each of these themes will now be discussed in detail.

5.4.1.2 The Decentralised Internet Cannot be Marketed

Throughout the expert interviews, this powerful statement emerged on many occasions. The argument posits that decentralised technologies, services or applications cannot be marketed solely on the fact that the technology is decentralised. In a question regarding which technology would be first to emerge, Philip Sheldrake responded:

> *None of them, none of them can, because decentralisation cannot be marketed. Why would anyone? So, for example let's take Solid, the project at MIT under Tim Burness-Lee, were they are trying to take the best of linked data and apply it to a social networked environment. Why would anybody leave Facebook to join a distributed social network? Why would they rip themselves away from the convenience, the comfort, the connection of Facebook to go and be by themselves over here on a distributed version?* (P Sheldrake 2016, Personal Communication, 19th October).

5.4.1.3 Decentralisation Must Supersede to Find Adoption

Following on from the theme that the *Decentralised Internet Cannot be Marketed*, the general consensus was that in order to achieve adoption, decentralised technology needs to do more to offer a value proposition. There is also consensus that the decentralised model can offer real innovation. When asked what the decentralised community needed to do to drive adoption, David Irvine responded: *'I think it is about supplying better, faster, cheaper: you can*

actually do that with decentralised technology' (D Irvine 2016, Personal Communication, 9th Nov).

Philip Sheldrake commented on the need to move away from providing decentralised versions of existing technologies. He argued: *'You cannot effect it, and that's the same for almost all decentralised versions of existing services. It's only when you come to the market with a decentralised architecture which has no precedence that you have the opportunity for adoption'* (P Sheldrake 2016, Personal Communication, 19th October).

5.4.1.4 Technology Is Not the Issue

A powerful theme that emerged is that technology, in terms of computer science, is now no longer a barrier to decentralisation. Through encryption, distributed ledger technology and existing standards, the solutions and relevant technology layers are now in place to deliver a broad range of decentralised applications. Jeremy Rushton commented: *'There are multiple obstacles, but technology is not one of them. We have loads of plausible ways that people like me can design you on the back of an envelope, to arbitrary levels of detail, decentralised systems'* (J Rushton 2016, Personal Communication, 27th Oct).

5.4.1.5 Getting to the Interface Layer

As the dominant forces control and drive to channel Internet traffic through their own interface layers, the challenge of reaching these layers becomes difficult. Whether this is hardware, application, search, commerce or entertainment portal, accessing the user through a portal, service or mechanism that is not either directly or indirectly controlled by a dominant force is a significate issue. Philip Sheldrake comments:

> *There would seem to be a great deal of work to be done here. Once you have got them at the interface layer, once you can aggregate through network effects at an interface layer, that has a concentrating effect that runs down through the layers of the tech stack far more powerfully then any centralising dynamic might ripple up* (P Sheldrake 2016, Personal Communication, 19th October).

5.4.1.6 Identity and Blockchain Technology Unlock the Model

The game changing nature of verifiable identity through blockchain technology is a significant step forward in making decentralisation a reality. The concepts of Self-Sovereign Identity and the supporting standards offer a concrete mechanism through which to achieve many of the principles of a Decentralised Internet. When asked what he considered to be the most significant developments in the decentralised space, David Alexander responded:

> *Self-Sovereign Identity. The idea that you have identity that you control, and everyone else is a relying party to it, and those relying parties choose what evidence they want from you to prove you are you is a complete paradigm shift from the way identity is being done at the moment, but I think it is almost at the point of being able to get to scale* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

5.4.2 Phase 2/3 Semi-Structured Interviews

Phase 2 of the Expert Interviews drew participants from the decentralised community with specific knowledge of Self-Sovereign Identity. The existing Semi-Structured Interview questions developed for phase one, have were used again in a simplified form, and reframed around the narrowed theme of Self-Sovereign Identity. A full transcript of questions can be found in the appendix of this thesis (see: C2).

The three narrowed investigative themes were as follows:

- The concept of Self-Sovereign Identity.
- The complexity at the interface layer of decentralised technologies.
- The adoption and value proposition for decentralised technologies.

Within Phase Two the participants included:

- **Colin Wallis -** Executive Director - Kantara *Initiative Inc*
- **Darrell O'Donnell -** CULedger / Technology & Strategy Advisor - Continuum Loop
- **Gregor Zavcer -** Protocol & Foundation - Co-Founder Datafund
- **Ian Forester -** Senior Producer at BBC R&D / Founder DataPortability.org
- **Jonny Howle -** Product Designer at uPort (ConsenSys) / UX-UI Designer IBM
- **Kaliya Young -** Identity Woman / Co-Founder Internet Identity Workshop / Leader in the field of Self-Sovereign Identity or Decentralised Identity Workshop
- **Micheal Becker -** Founder, managing Partner - Identity Praxis, Inc.
- **Phil Windley -** Chair at Sovrin Foundation / Founder and Organiser - Internet Identity Workshop / Author of Digital Identity: Unmasking Identity Management Architecture (IMA)
- **Sunil Malhotra -** Founder CEO - Idea-Farms - New Delhi - India

Within this selection there is representation from the UK, US, Canada, India and Germany. Each interview lasted between 20-45 minutes.

Phase 3 of the expert interview data was derived from the focus group undertaken as part of the Practice Led Component of this research. The second part of the focus group posed the same questions as those asked of the individual interviewees previously listed. This collection of individuals added depth to the data as they came from a predominantly design based usability prospective as opposed to a position focused on the Decentralised Internet.

Within Phase 3 the participants included:

- **Paul Wilshaw** - Head of Innovation / Barclays UK
- **Ian Forester** - Senior Producer at BBC R&D
- **Suzanne Clark** - Senior UX Designer BBC R&D
- **Tom Cheesewright** - Applied Futurist - Broadcaster
- **Jasmine Cox** - Product Designer BBC R&D
- **Le T Hoai** - MA Digital Experience Design at Hyper Island
- **Marc Tobia** -MA Digital Experience Design at Hyper Island
- **Tim Panton** - Co-founder and CTO at Pipe
- **Lisa Ortega** -UX Expert and Digital Psychologist at Keep It Usable
- **Nathen Broadbent** - Head of Creative Technology at The Trunk Agency

Within this collection of participants there was representation from the UK, US and Taiwan. The second phase of the focus group lasted for 45 minutes.

5.4.3 Thematic Content Analysis

Expert interviews from all three stages of data gathering have been transcribed and time coded. All transcriptions were loaded into the qualitative analysis application NVivo. The transcriptions were coded through a number of cycles of generation and combination. While undertaking the analysis a total of 48 codes were generated. A code book has been exported from NVivo, and this can be found in the appendix of this thesis (see: C5).

Once coded, a process of memoing was undertaken. Collections of quotes associated with codes were printed, and the process of memoing was undertaken physically. Through this process a significant number of themes were identified. Themes were then categorised into 3 core areas, Adoption, Interface and General Concept. In total 64 themes have been defined.

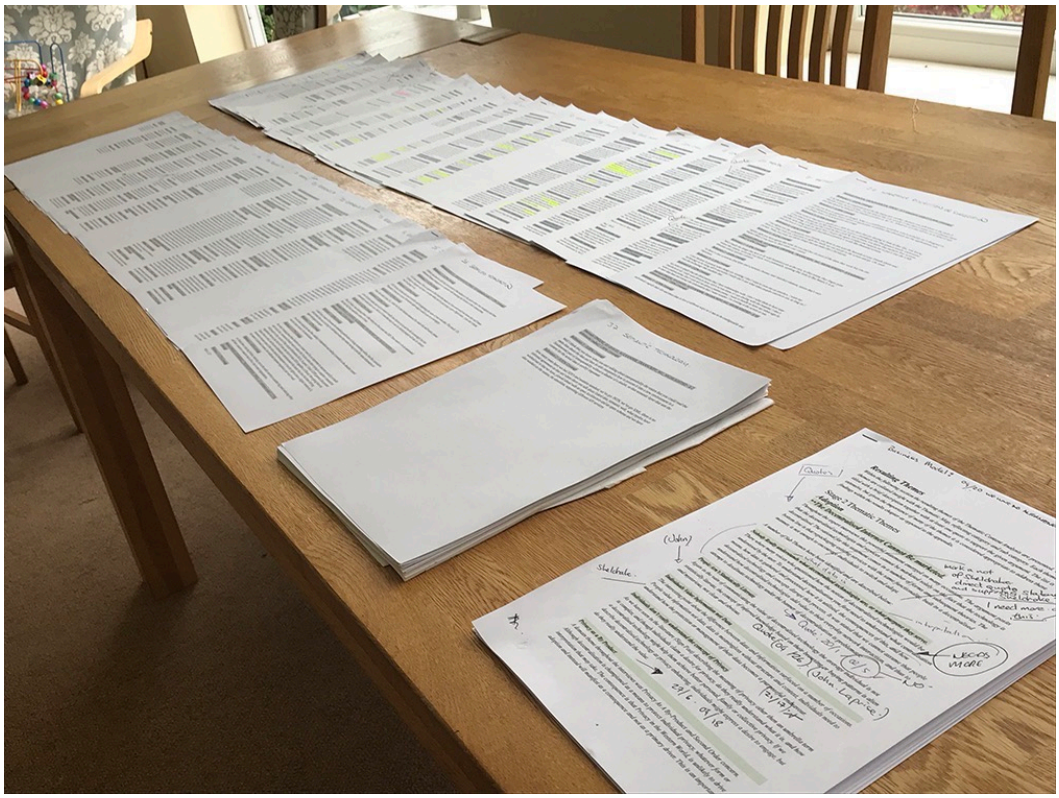


Figure 22: Documenting the Memoing Process

Figure 23 on the following page illustrates a mapping of the themes generated through the memoing process.

5.4.4 Thematic Map

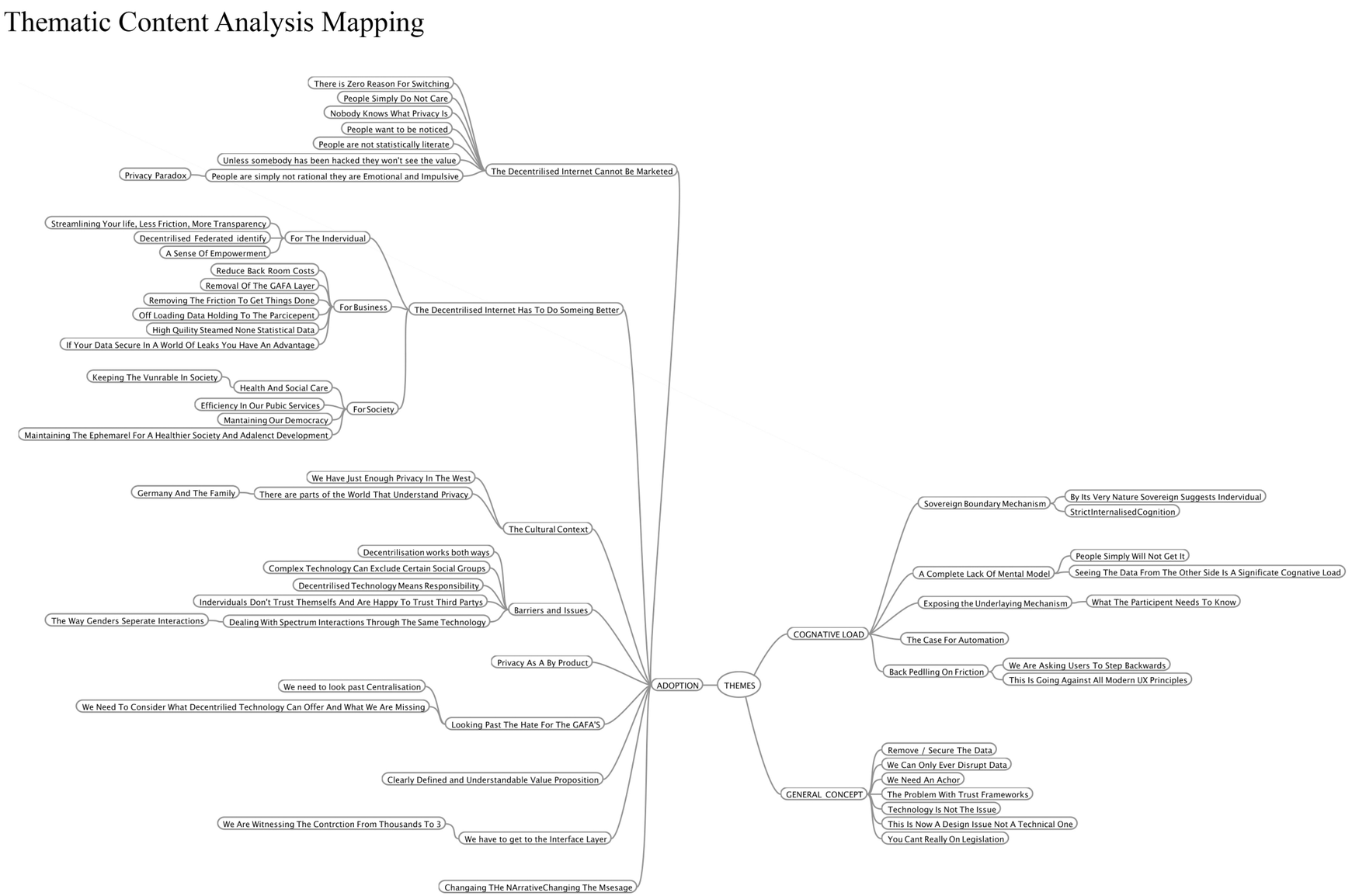


Figure 23: Thematic Content Analysis Mapping

5.4.5 Resulting Themes

Within this section the resulting themes of the content analysis are presented. The themes were collated in line with the thematic map, reflecting category and subcategory. Each theme is titled with a brief description. The themes are of both semantic and latent, and where available direct quotes are included to support a given argument. The list is extensive but given the importance of many of these themes, it is considered appropriate to address the findings in full within the results chapter as opposed to the appendix.

5.4.6 Adoption

5.4.6.1 The Decentralised Internet Cannot Be Marketed.

Throughout the expert interviews this powerful statement emerged on many occasions. The argument posits that decentralised technologies and services cannot be marketed solely on the fact that the technology is decentralised. The rationale for this statement comes in a number of forms and subsequent theories. The consensus is that repackaging existing services with the same functionality, built in a decentralised manner, is not enough to encourage individuals to switch and adopt. Sheldrake responded when asked which technology might be first to market with the statement: *'None of them, none of them can, because decentralisation cannot be marketed'* (P Sheldrake 2016, Personal Communication, 19th October).

A number of sub themes regarding the marketing of decentralised technologies have been identified and are described below:

- **Nobody Really Understands Data, What Decentralised Technologies Are, or What Purpose They Serve.** There were many voices who questioned how the notion of decentralised technologies would be understood by the participant. To understand decentralisation, you first need to understand what centralisation means, how data is gathered and processed, how it is utilised, the consequences of this, and how decentralised technology and concepts disrupt this process. It was stated that we cannot assume that people understand how decentralised technologies add value to their everyday network interaction, and thus to attempt to disseminate technology under the banner of decentralisation is problematic. Paul Wilshaw commented: *'You know, you tell them you're wearing a watch and it gives off millions of bits of data every hour, and they go seriously, they don't understand that complexity, what that data is, or what that data is doing, or what that data is for'* (P Wilshaw 2019, Focus Group, 15th March).

- **People Aren't Statistically Literate.** Again, in the vain of communicating the value of decentralised technology the average individual is not statistically literate. The concept of inferred knowledge based on their browsing or buying patterns is alien and mysterious, and not clear or obvious, John Laprise commented: *'Most people are not statistically literate, so explaining to them how you can take a large data set and figure out you're actually pregnant is just as much magic as I can put these two minerals together and come up with gun powder'* (J Laprise 2019, Personal Communication, 2nd Dec 2016).

- **Individuals Value Information Not Data.** The interesting point about the difference between data and information surfaced on a number of occasions. People value information not data. Data is meaningless without structure and context. Individuals need to understand how the protection or management of their data becomes a purposeful endeavour. Philip Scheldrake commented: *'The vast majority of people connected to the Internet couldn't care less about their personal data, people don't deal with data. People deal with information'* (P Sheldrake 2016, Personal Communication, 19th Oct).

- **Individuals Don't Really Understand the Concept of Privacy.** The concept of privacy understanding emerged as a latent theme. As seen in the literature (see: 2.2.1) describing the meaning of privacy other than an umbrella term is complex, and though individuals claim a desire for privacy, do they really understand what it is, and how a decentralised technology might help them achieve better personal, family or collective privacy. If we describe decentralised technology as privacy enhancing, individuals might express a desire to engage, but do they really understand the value. Collin Willis agreed: '*I think that's right. though that doesn't necessarily mean to say that the conversation should stop'* (C Willis 2018, Personal Communication, 11th Oct).

- **Privacy as a By Product.** An important theme throughout the interviews was privacy as a by-product and second order concern. Although decentralisation is championed as a means to protect individual privacy, whatever form or interpretation that may take, the consequence is that the prospect of privacy in the western world is unlikely to drive adoption, and instead will manifest as a consequence and not as a primary driver. This is an important point as it potentially acts as a major consideration in the development of strategy for the decentralised community. Darrell O'Donnell commented: *'What we learned on our project, is that by handing you trusted information, the Paradigm shifted from, how we make sure this is privacy respecting, to privacy becomes a second order issue'* (D O'Donnell 2019, Personal Communication, 9th Feb).

- **Unless They're A Die-Hard Activist, Is A User Really Going to Jump to An Inferior Decentralised Alternative.** A topic raised on a number of occasions, and one that forms a latent theme, questioned the logic of anybody leaving established high-quality centralised services to move to a lesser quality service just because it is decentralised. The consensus seemed to be that in reality this is simply not going to happen. Sheldrake commented: '*why would anybody leave Facebook to join a distributed social network, why would they rip themselves away from the convenience, the comfort, the connection of Facebook to go and be by themselves over here on a distributed version'* (P Sheldrake 2016, Personal Communication, 19th October).

- **Individuals Don't Want to Hide.** This latent theme argues that in a world where our participatory digital culture is becoming increasingly broadcast focused, where social standing, interaction and omni presence are socially inclusive norms, is decentralisation really that attractive and does it really warrant social reclusion or off grid mentality? Do people really want to hide?

- **People Simply Don't Care.** The statement of apathy was made on a number of occasions. People simply don't care and have better things to worry about and lives to lead. Unless they have suffered a hack or data breach people are unlikely to see any value in decentralisation or engage in decentralised technologies. Jeremy Rushden commented*: 'Everything I observe about people in the field is that they don't care. They don't care and give their passwords up for chocolate'* (J Rushton 2016, Personal Communication, 27th Oct).

- **People Are Not Rational, They Are Impulsive and Emotional Creatures**. On a number of occasions, the point was made that human beings are not always rational, and that many of the actions that we undertake can be based on emotion and impulse. We see this in the theory of the Privacy Paradox and Instant Gratification Bias, (see: 2.2.8) but in the conceptualisation and design of decentralised tools this is certainly something that should be considered. Michael Becker commented: *'Don't make me think. It's not a rational decision. We're not rational. That's the thing when you try to approach this from a technology prospective, you try to approach it with rational decision making and Neuroscience is proving that were not rational'* (M Becker 2018, Personal Communication, 28th Sept).

5.4.7 The Decentralised Internet Must do Something More

Continuing from the theme that the Decentralised Internet cannot be marketed, the general consensus is that in order to achieve adoption, it has to do more to add value and advantage. Decentralised technology has to develop innovations to supersede a centralised model. This process has identified a number of opportunities for innovation. These have been divided by context into three sub-categories: The Individual, Business, and Society.

5.4.7.1 The Individual

A number of themes concerning the topic of value for the individual have been identified, and are described below:

- **Streamlining Your Life, Less Friction, Instant Gratification, More Transparency**. This theme embodies a core value of decentralisation as it argues that the collection and control of personal data, the capability to combine, to reuse, to utilise and understand personal credential and data streams, offers an opportunity to simply make our lives easier. It streamlines our everyday activities and reduces the friction in getting tasks accomplished. David Alexander commented: *'You can have everything you need, just to get that done seamless frictionless experience, job done booked sorted, no threat to your personal safety security, no risk of marketing'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

- **Decentralised Federated Identity.** The federation of established identity across interactions and services is a powerful concept, and one which is utilised by the centralised Internet, as participants develop rich identities that they themselves will never truly own. To control an identity, disseminate and reuse it in a way the participant chooses, and to invest and build persona and digital presence in a Sovereign Boundary Mechanism is a powerful concept as it liberates users from the dominant digital oligarchs. Johnny Howle commented: *'Right? so we can get traction there, and the thing that Self-Sovereign Identity really has to solve is centralised federated identity'* (J Howle 2018, Personal Communication, 4th Feb).

- **A Sense of Empowerment, Transparency and Agency.** Decentralised technology offers the participant a sense of empowerment, affording the same degree of control in the digital space that they enjoy in the real world. Individuals can understand their network relationships, their past transactions, their patterns of interaction, combine data streams to infer and inform, and ultimately decide who or what has access to the data and subsequent information they control. David Alexander commented: '*we never got any further than some transaction. They never got to a point of empowerment or awareness about their life. We were still sitting in a world where there where filing cabinets full of paper'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

- **Avoiding the Cost of Surveillance Capitalism.** A powerful theme and narrative is the cost of surveillance capitalism in the sense that a centralised model locks in and prevents users from engaging in alternative business models, and also from a societal and democratic prospective. Control and cost are powerful narratives, and if coupled with tailored value propositions could prove compelling arguments for adoption. Adrian Gropper commented: *'Individuals simply don't understand just how much surveillance capitalism is costing them'* (A Gropper 2016, Personal Communication, 23rd Nov 2016).

- **Security in The Ephemeral**. This latent theme derived from conversation regarded public benefit and privacy, is that of the ephemeral. The prospect of decentralised technology restoring the ephemeral back into exchange of opinions and views and general conversation is a powerful narrative if expressed and understood correctly.

5.4.7.2 For Business

A number of themes concerning innovation opportunities for business have been identified and are described below:

- **Removal of the GAFA Stranglehold.** The reality of the GAFA portal 'Google, Apple, Facebook and Amazon' was raised a number of times during interviews. We have seen a significant contraction of entry points to the network reduce significantly. Hardware, storage, supporting technology, marketing and federated identities are now reliant in some way on the network's dominant forces. The prospect of decentralisation offers liberation from this monopoly and is seen by many as an energising innovation. Philip Sheldrake commented, *'So this if you like is the final conclusion to the incredible collapsing down of the number of companies that feature in our digital lives', from hundreds ten years ago, to which of the two you happen to pick, plus Facebook'* (Sheldrake, P. 2016, Personal Communication, 19th Oct).

- **Removing the Friction to Get Things Done.** Friction is a significant obstacle to transactional completion and overall customer experience, for which reducing or removing through decentralisation is considered to offer great efficiency. David Alexander commented: *'Decentralised models are considered to offer a significant means of reducing friction, driving efficiency and ultimately competitive advantage'* (Alexander, D. 2016, Personal Communication, 21th Oct).

- **Off Loading the Responsibility, and Cost of Holding Data.** Recent strengthening of data protection law has seen greater accountability placed on data holders and processors. The costs involved in securing data are increasing for existing companies and present a significant overhead and risk for start-up businesses. If a decentralised model can allow companies to access verified customer data in a remote location, on demand and only when needed for the purposes of the transaction in hand, the commercial benefits are clear. Tim Panton commented, *'I think what this would be good for as soon as smaller businesses and especially if you could guarantee your customer database security through something like this that didn't cost millions and require experts'* (T Panton, Focus Group, 15th March 2019).

- **If Your Data is Secure In a World of Leaks You Have a Competitive Advantage.** In a world where news of data breaches continues to be a regular occurrence, if a business's processes are clearly more secure than their competitors, then the secure data model offers a clear competitive advantage. David Irvine commented, *'if you've got a decentralised technology that's incredibly secure and not losing data, then you've got a company that will be much more reliable, more secure and not getting sued and able to operate faster'* (D Irvine 2016, Personal Communication, 9th Nov).

- **Reducing Back Office Costs.** Backroom costs featured as a prominent point in light of participants being able to provide verified credentials within a decentralised model. It is noted that although companies have off loaded labour costs for data entry, any data they received as part of a transaction still had to be checked, whether at a fundamental level of units or address, all the way through to licenses, certificates, ratings and applicant background. Any model that could significantly reduce these processing costs would give a competitive advantage. David Alexander commented, *'There is massive back office costs for organisations, there is an army of people behind the scenes'* (Alexander, D. 2016, Personal Communication, 21th Oct).

- **High Quality Streamed, Realtime, Non-Statistical Data.** The concept of clean, streamed, up to date, non-statistical data is a latent theme that emerged through the interviews. At present, commerce is reliant on statistics that support market intelligence and advertising. Customer data can be inaccurate and can quickly become dated. The majority of current data is static, approximated and by its very nature outdated. Decentralised models through vendor relationships offer the prospect of live streamed, real time data, which is potentially far richer and relative. Customer loyalty, relationship, service, business planning and efficiencies would vastly improve through dynamic data streams.
- **New Forms of Business Based on Vendor Relationship Management.** The reversal of relationships between participant and vendor, the availability of real-time ethically sourced data empowering the participant to cast intentions, emancipation from the GAFA's and the levelling of the playing field will disrupt the current centralised model and will generate new forms of business opportunity. A great deal of R&D has been undertaken through past VRM projects, and they act as a valuable resource for future

developers. When asked if she thought there was opportunity in historic VRM projects, Kaliya Young commented '*oh totally yeah*' (K Young 2018, Personal Communication, 21st Nov).

- **Customer Relationships, Trust, KYC**. This latent theme centres around the building of customer relationships through a *Web of Trust* which will provide a means for business to satisfy their responsibility to *Know Their Customer*. This will reduce friction and costs in this area and contribute to seamless transactions.

5.4.7.3 For Society

A number of themes concerning the benefits to society have been identified and are described below:

- **Maintaining the Ephemeral for a Healthier Society and Adolescent Development**. During interviews and within the literature the concept of perceived privacy, surveillance capitalism, profiling and the impact this has on our social norms, especially as an adolescent population, has become important. The premise that the ephemeral is important in our daily social interactions, and the development and shaping of identity and relationships in younger persons, is a very compelling argument. In essence, individuals need to explore the boundaries of society and develop as human beings without the fear of every interaction, conversation and choice being a matter of indefinite record in the ether of the network.
- **A Stronger More Cohesive Society.** The benefits of data sharing and utilisation across a social context in terms of education, health care, social care, crime prevention, planning, efficiency, the smart city, and the environment would benefit significantly from a decentralised network. The means to share personal data in a secure and trusted manner would mean powerful real-time data streams feeding into all facets of a data driven society. Philip Sheldrake raised the issue of structuration: *'should you wish to make a difference, the structuration and the structures of society that either enable or disable you for expressing or welding your personal agency are important'* (Sheldrake, P. 2016, Personal Communication, 19th Oct).

- **Maintaining Our Democracy.** Conversations reinforced the risk to our democracy of centralised data gathering, processing and inference poses. Although the eradication of targeted influencing through individual profiling is unlikely, a Decentralised Internet may at least offer a means of disruption, and a return to a more private and internalised semblance of political position and voting intention.

- **Efficiency in Our Public Services.** Efficiency in public services can be improved and ultimately costs reduced significantly through the availability of real time data. Public health, planning, transportation, energy consumption and social care are but a fraction of societies services that could be improved. David Alexander commented, *'so we got to come up with a better way, reducing friction, reducing the cost to the public services, reducing the effort and the massive back office cost in most organisations'* (Alexander, D. 2016, Personal Communication, 21th Oct).

5.4.8 The Cultural Context

A number of themes concerning cultural considerations have been identified and are described below:

- **The West Has Just Enough Trust.** Just enough trust is a theme that considers the culture attitude, and the geo economic and political landscape. The argument suggests that the western world currently enjoys just enough trust not to concern itself with decentralised tools and services. There are data breaches, and the topic of surveillance and privacy in the digital realm is prominent in the media, but there is still a sense of security. We aren't completely exposed, and very few individuals have experienced a catastrophic data breach that has resulted in personal reputational or financial harm: *'You know in the Western World my sense is that it'll take a long time before there's actually any value proposition in many developed countries, because there's enough trust that there's no motivation to change'* (C Willis 2018, Personal Communication, 11th Oct).

- **Parts of the World and Cultures that Value Privacy.** A strong theme throughout has been that of cultural difference, attitudes and understanding. Differing parts of the world consider privacy differently, and value it for a number of varied reasons. A western liberal democracy understands value proposition within privacy enabling technology in a different way to a repressed culture. Cultures who live within or have a recent history of state surveillance may have a differing world view to others. Considering cultural differences is a key consideration when developing for adoption. Tim Panton commented: *'I mean our experience in Germany is completely different from this country and our experiencing in the Far East just different again'* (T Panton, Focus Group, 15th March 2019).

- **One Size Does Not Fit All.** The notion of one size not fitting all relates to the above themes, in that designing one decentralised solution or engineering a value proposition to suit all cultures and conditions may require reconsideration. The prospect of multiple angles and the requirements of different user groups offers design opportunity. If the objective is to roll out a standardised Sovereign Boundary Mechanism, differing problem spaces can be identified, and applications and messages tailored to the

circumstance, raising the probability of adoption. David Alexander commented: *'individuals cognify their life and their data in lots of different ways, so it has to be configurable and ultimately very personalised'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

5.4.9 Routes to Adoption

A number of themes focusing on routes to adoption have been identified and are described below:

- **High Value High Friction.** Concepts have been expressed regarding routes to market, adoption and traction. A significant concept is the steering away from the mass consumer market and instead targeting big business and identifying transactions which experience high friction and regulatory burden. It is suggested that there is real value proposition in this area. Kaliya Young commented: *'Focusing on high value, high friction business first, the problem of focusing on the mass consumer market and not specific high value. High friction Identity transaction points, high regulatory burdens that are costing institutions serious money, those things. That's where we go after first'* (K Young 2018, Personal Communication, 21st Nov).

- **Targeting Cultural Context as a Break-Through Mechanism**. The theme around targeting of cultural context relates to seeking cultural, social, political and geographic groups that might see a significant value proposition within their current reality. The argument suggests that in order to drive innovation and momentum, pockets of respective user groups should be pinpointed and supported: Germany was raised as a culture that understands privacy from a family perspective. The Unbanked within failing economies. Individuals who are undocumented within immigrant or refugee communities. And repressed populations living without many of the liberties enjoyed in the western world. Targeting cultural context would allow efforts to foster adoption and develop innovation to be strategically planned, while supporting those who would arguably benefit from decentralised technology the most. Colin Wallis commented: *'It's going to be rather a play thing for a while, at least for the majority of down to a socio-economic level of banked, when you get to under-banked or unbanked I think then you*

know, potentially you can, you can bring in self sovereign identity at this point' (C Willis 2018, Personal Communication, 11th Oct).

- **On-boarding and Companies Bringing their Existing Customers with them**. The on-boarding concept through existing customers comes from the argument that, given the complexity of the technology, many users are likely to be introduced to, and lead through initial engagement instead of adopting independently. This leads to the idea that if a large organisation already has an existing registered user base, and a decentralised innovation can be proven to support or streamline that organisation, existing customers can be transposed to a decentralised system en masse, without even becoming aware they have moved to a new paradigm of network engagement. From there, new innovations, ideas and applications can slowly be introduced. This model can be seen in the experimentation being undertaken in Canada through a number of credit unions. Darrell O'Donnell commented: *'I look at this as the winner, here is the trojan horse. People have no clue that their wallet is now running verifiable credentials that they're doing zero knowledge proofs'* (D O'Donnell 2019, Personal Communication, 9th Feb).

5.4.10 Barriers and Issues

A number of themes concerning barriers to adoption and general issues have been identified and are described below:

- **Getting to The Interface Layer.** Targeting the interface layer is an issue that's raised as both an objective and, in this context, also a significant challenge. As the dominant Internet forces drive to channel traffic through their own interface layer, the challenge of reaching these layers becomes difficult. Whether this is hardware, application, search, commerce or entertainment portal, accessing an interface layer that is not either directly or indirectly controlled by a dominant force is a significant issue. Sheldrake comments: '*It's that layer that will make or break the ability to allow a de-concentrating effect to ripple down through the rest of the technology layers, giving them room to breathe and manifest themselves'* (P Sheldrake 2016, Personal Communication, 19th October).

- **Decentralisation Works Both Ways.** An interesting latent theme identifies the issue of decentralisation working both ways. Once verified personal credentials and information can be easily established and disseminated, demands for such data could raise significantly and we may find ourselves in a situation where companies and organisations demand ever increasing access to personal data in exchange for goods and services. Information and data currently outside of the boundaries of availability, may quickly prove more problematic than helpful. To counter, the surrounding mechanisms of decentralisation need to be in place as well as the data itself, if we are to avoid these potential negative unforeseen consequences.

- **Complex Technology Can Exclude Certain Social Groups.** The notion of exclusion is a broad and complex consideration. For those who are time rich, well-educated and technology literate, the advent of decentralised tools may prove valuable, but for those in society who lack the fluidity to engage, the advent of such technologies may prove to exclude and alienate from valuable opportunities. The situation may be likened to an extended version of what is currently observed with computing technology and the older demographic in our society. Lisa Ortega commented *'is this going to just overwhelm people and they're not going to end up making the right decisions'* (Lisa Ortega, Focus Group, 15th March 2019).

- **Decentralised Technology Means Responsibility.** A powerful theme relates to responsibility. Decentralised technology by its nature realigns the management of personal data back to the individual. A paradigm shift away from centralisation creates a significant responsibility which may be embraced or rejected across many differing contexts. The argument is that decentralisation places a significant responsibility on the shoulders of the participant, and the ongoing friction associated with the management of personal data may prove problematic. Jonny Howle commented: *'It's just a lot to task and that is, to have people take responsibility and control over their identity and their sovereignty and the stakes are high. If you fail to do that correctly, you can lose a lot of value'* (J Howle 2018, Personal Communication, 4th Feb).

- **Individuals Don't Trust Themselves and Are Happy to Trust Third Parties.** Continuing from the theme of responsibility, an individual's confidence and competence to manage their own personal data needs to be considered. Once an individual is faced with the potential complexity and workload of managing their own personal data, it may transpire that many simply have no interest in doing so and may indeed be happy for a third party to do that for them. That may be the status quo, a family member or friend, or it may be a public service organisation, or a designated organisation specialising in this area of personal service. The argument would appear to be, that we should not assume that individuals will want, or have confidence in themselves to control their own data. Lisa Ortega commented: 'You're like, I don't know if I want to deal with that and so then maybe it's easier for me to just go to a third party' (Lisa Ortega, Focus Group, 15th March 2019).

- **Dealing with A Spectrum of Interactions Through the Same Technology.** An interesting theme came through a conversation about different social and gender groups, that males tend to feel comfortable conducting their affairs through one network domain, whereas females tend to prefer to separate their interests across a number of unrelated domains. Also, gay men might see the benefits in a self-sovereign decentralised system that helps their community manage sexual health, but that they might not be comfortable dealing with other personal issues through the same domain. The central argument is that a decentralised digital wallet for managing personal data, will need clear boundary domains, if people are to feel comfortable and confident using it, at least at the beginning. So that in early email adoption terms, the classic error of mistakenly sending an email to all contacts can be avoided. Kaliya Young commented: *'Women pro-actively managed persona separation in their lives more than men'* (K Young 2018, Personal Communication, 21st Nov).

- **Non-Profit Does Not Make A Good Business Model.** How technology and resource is funded without a capital driven business model is hard to understand. The centralised internet has a clear model of data collection and exploitation to drive market intelligence and advertising. The consensus is that everybody would benefit from a Decentralised Internet, including the traditional capital driven models, once established. But getting there will require significant investment, and it is difficult to

see how the ecosystem gets to a sustainable position and critical mass. Gregor Zavcer commented: *'It should be public good which directly conflicts, you know, with the profit-driven activities a company should do, like returning value for to its shareholders or investors'* (G Zavcer 2018, Personal Communication, 4th Feb).

5.4.11 Interface

A number of themes concerning human interfacing with decentralised technologies have been identified and are described below:

- **Sovereign Boundary Mechanism**. This concept is a latent theme which describes a realm in which a participant has a surrounding boundary of control within which they exercise agency over the terms by which personal data can be transacted through a number of core mechanisms. The boundary is sovereign, and traffic moving from the network into the participant's domain and back out again, stops at the boundary where management, access and dissemination of data is controlled, either on a case by case basis, or automatically based on a number of predetermined parameters.

- **By Its Very Nature Sovereign and User Centric Suggests the Individual**. Across interviews a topic related to a Human-Centred Data Ecosystem and what it means to be sovereign was both engaging and prominent. The notion relates to independence of the participant. In order to be sovereign on the network a user must control and take responsibility for their own data and information. The user requires transparency, understanding and the agency to decide how and where personal data is distributed and utilised.

- **Strict Internalised Cognition.** If a participant is to take control of their own personal data in a decentralised context meaning there is no third party in the equation, and if the participant is to have the confidence in a system, approach and interaction, a developed internalised understanding will be required, in much the same way as an individual understands the process of withdrawing money with an ATM, i.e. the proving of identity, withdrawal of wealth, the updating of a ledger and so on. A user of a decentralised system will require a rather complex and in-depth understanding of what

are arguably demanding abstract concepts if they are to confidently, efficiently, and sustainably engage with such a system.

- **The Technology Has to Be Open Source.** For any decentralised system to be truly considered decentralised it needs to be open source. The software has to be open to everybody, not owned by any third party, transparent and freely available for inspection. Adrian Gropper commented: *'it probably almost certainly has to be open source, otherwise somebody else owns it'* (A Gropper 2016, Personal Communication, 23rd Nov 2016)

5.4.12 The Missing Mental Model

Themes concerning this topic have been identified and are described below:

- **The Participant Simply Won't Get It.** The general expert position when confronted with the complexity of an analogue decentralised engagement is that the participant will simply not understand it. It is suggested that there is no precedent or real metaphor to describe the underlaying concepts and interactions and as such any participant will struggle to make sense of the environment. Many of the concepts and interactions are considered to be abstract and gaining an understanding and confidence to engage will potentially require a considerable amount of assistance or experimentation.

- **Changing the Narrative, Message, Language and Metaphor.** The need to re-evaluate and reconsider the narrative, message, language and metaphor from a strategic perspective is considered very important. In order to canvas value and a better understanding of what decentralised technologies can offer the participant, the communication needs to be considered, meaningful full and consistent: David Irvine commented: *'Stop talking about privacy because folk just don't get it, when you lose your privacy then the government or whoever is able to control you, and when you talk to people about control it becomes a very different thing',* D Irvine 2016, Personal Communication, 9th Nov'.

- **Individuals Would Have to Live and Breathe This to Understand.** An interesting point raised was concern that for an individual to engage with a full analogue decentralised system, a participant would have to be using such a system on a continuous basis, in order to develop and retain the required understanding.

- **Seeing the Data from The Other Side Is A Significant Cognitive Load.** An interesting view held in relation to a participant's mental model is that of an understanding of how data manifests itself on the other side of any transaction. This is considered important as participants make judgements regarding how they manage their own personal data. It is suggested that grasping the view of one's personal data from the opposing side of a data exchange poses a significant cognitive task.

5.4.13 Exposure of The Underlying Mechanism

Several themes concerning this topic have been identified and are described below:

- **What Participants Need to Understand, See and Have Access To**. A theme discussed and raised on several occasions was that of the exposure of the underlying mechanism. For a participant to engage with a system, learn how a system functions, and attain the required agency specified within the HDI, SSI principles, what level of exposure is required? Is there a balance between possible control and interaction and the development of a truly Decentralised Internet? In discussing the level of understanding required, John Laprise commented: '*To some degree yes, but not too deep'* (J Laprise 2019, Personal Communication, 2nd Dec 2016).

- **Exposure of the Mechanism and the Value Proposition**. An important theme is that of the exposure of the mechanism and value proposition. This is considered a balance in terms of perceived value. If the system is fully automated and controlled by AI the participant has no understanding of what the underlying system or mechanism is doing, and as such, unless the participant has complete trust in the system, they potentially see little of the value. By contrast, if the system reveals and provides access to all of its mechanism, the participant is potentially overwhelmed, the cognitive load is too great, and adoption is unlikely. There is a potential paradox, or at the very least variable, for balance in any design discussion process.

5.4.14 Back Pedalling on Friction

A number of themes concerning this topic have been identified and are described below:

- **We Are Asking Users to Step Backwards.** In the light of the emergence of Self-Sovereign Identity as the leading light of decentralised development, the issue of usability struggles to maintain the contemporary practice surrounding HCI and usability design. Any scalable analogue SSI interface is considerably complex and pushes against the design philosophies of simplicity and low friction. To engage with a UI enabling the full-scale SSI interaction poses a challenge to designers and stretches the limits of the participant's attention. It is felt, in the short term at least, engagement will require an expectance of complexity and friction. Johny Howle commented: *'In the short term, it's a downgrade and the value propositions have to be strong to convince somebody to do that'* (J Howle 2018, Personal Communication, 4th Feb).

- **This Is Going Against Modern UX Principles**. In terms of designing a user experience model from contemporary and more specifically, Self-Sovereign Identity systems, we are going against modern UX principles by introducing complexity and arguably internalised cognitive load. These systems are complex and in order to move forward we arguably need to step away from the concept of lowering friction.

5.4.15 The Case for Automation

- **An Agent or Personal Jeeves That Acts in the Best Interests of its Master.** The prospect of an agent or personal assistant is raised on a number of occasions, with the argument being that it would be unlikely that any participant would have the inclination to manage each and every exchange of data or credential through a decentralised system. As such, any system would require an agent driven by some means of AI, working on behalf of the participants, to monitor and manage the mundane, and flag any important issues to its master. John Laprise comments: '*A gentleman's man servant, who does everything in the interests of his client and no one else, even if the client doesn't realise it'* (J Laprise 2019, Personal Communication, 2nd Dec 2016).

- **Scalability.** The issue of scalability of systems was discussed as an issue on a number of occasions. Given the volume of data any Human-Centred Data Ecosystem would be tasked with managing, particularly in the form of SSI, it would soon run into scalability issues. Managing data and understanding transactions within a small scale, lab-based demonstration, is straightforward; in the real world it is quite something else. Phil Windley comments: *'so I think it's an interesting question about how this works its scale. I'm not sure I have any answers'* (P Windley 2018, Personal Communication, 12th Nov).

- **Setting Broad-Brush Stroke Policy.** In any agent or AI supported decentralised system, the concept of establishing a broad-brush policy which drives decision making is a recurrent theme. The participant either defines the broad policy or the system itself learns the participant's preferences. Decisions are then automatically made based on these policies. David Alexander commented: '*I think when you look at policy level work you set some broad-brush strokes'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

- **A Trust Network to Drive Agent Decisions.** The notion of a trust framework or affiliated group driving decision making is a powerful theme. Such a system, for example, could make automated decisions on the disclosure of personal information based on a collective intelligence or political affiliation.

5.4.16 Third Party Offloading

- **Power of Attorney for The Young, Old and Infirm.** A barrier to adoption for SSI is the cognitive loads and emotional impact of taking full responsibility for one's data and its management. At present, participants may not have control over their data but they have the comfort of knowing that it is being held by a third party who, at the very least, prevents all out data loss, and on the whole provides easy access to data in an understandable context. An interesting area around SSI is the notion of passing control over identity and elements of data to others. This may be to a family member or a loved one. The concept may benefit the very young, for example, a child's medical records managed by a parent. It might also help older people who are less technically literate, or the ill or infirm. This concept of giving trusted parties control over personal data may also find value for those who simply do not want the responsibility. Ian Forester commented: '*It could be the niece or nephew, or it could be the cousin that makes certain decisions rather than it always having to be on you'* (I, Forester, 2018, Focus Group, 15th Mar).

- **To A Group or Affiliation.** The notion of offloading part or all of the responsibility for one's personal data to a third party aligns with the concept of group or affiliation. This is very much in the vein of a trust network in which an individual's hand over responsibility for decision making through broad brush policy alignments. A very interesting evolution of this is the concept of political affiliation, a democracy in which the off-loading of the control of data becomes almost a dynamic way of continually casting a vote.

- **To a Public Service Operator.** Suggestions have been made during debate that one type of custodian of personal data and decision making might be a public service operator. In this case the organisation cited was the British Broadcasting Corporation (BBC). How exactly this would function is unclear, but it would presumably be under

some extended charter yet to be considered. In the context of offloading responsibility, it is a concept worth exploring. Ian Forester commented, *'And so that could be an established organisation like the BBC'* (I, Forester, 2018, Focus Group, 15th March).

5.4.17 Broader Themes

- **Remove / Secure the Data**. A powerful theme and subsequent argument comes from the notion that, wherever possible, data is removed from the equation altogether. The argument is that the primary issue is the underlying technical model, the centralisation of data on servers. A distributed addressable model of data storage removes the target and thus, in many cases, the vulnerability. IPSS and Maidsafe have already proved the concept at scale and embracing this method of data storage may supersede many of the decentralised ideas around access control and contracts. David Irvine commented, *'so, it's not about securing servers just get rid of them, the only way to secure a server is not to have it'* (D Irvine 2016, Personal Communication, 9th Nov).

- **We Can Only Ever Disrupt Data Access.** The general consensus is that we will never be able to control all of our personal data generated individually, in connection with others or inferred by organisations. It is also unlikely that we could control access to data without relying on some type of external authority or legal framework. Instead, we need to better understand what we can and cannot control, what we can't, and what we can disrupt. A clear understanding of this might assist in strategic planning, the design of technology, and the development of value proposition. David Irvine commented: *'as soon as you have given somebody the data, they have got it, and they can do whatever they want with it, and there's no way to stop that'* (D Irvine 2016, Personal Communication, 9th Nov).

- **The Problem with Trust Frame Works.** There has to be some trust and solid ground in the transactional and identity verification equation. The principles of a Decentralised Internet give the individual agency, placing the participant at the centre of data streams, enabling independent choices about who and what to trust. At present, we have certification to assure us that online entities are who and what they claim to be. With the advent of decentralisation, this is potentially going to change, requiring new forms of trust anchors or frameworks across contexts. Trust could be seeded and maintained by community, state or public service organisation, or could be developed through new forms of dynamic reputation mechanisms. The concept of a trust framework is a complex issue which to date, it is argued, has been considered through a very narrow lens. Kaliya Young commented, *'It's kind of from a very privileged, male, North American, white centred world view about what trust is, marginalised, like marginalised people want accountability between the system and the people in it'* (K Young 2018, Personal Communication, 21st Nov).

- **This Is Now A Design Problem**. A powerful theme is that the problem of achieving decentralisation has moved from the development of the technology layers, through to one of user experience, interaction and product design. Through encryption, distributed ledger and existing standards, the pure computer science solutions are now in place to deliver decentralised technology. As Jeremy Rushton commented: *'there are multiple obstacles, but technology is not one of them'* (J Rushton 2016, Personal Communication, 27th Oct). In parallel, individuals have commented on the need for these issues now to be considered through design thinking. As Sunil Malhotra comments**,** *'I totally agree with you, It's now a design thinking problem. Yeah, absolutely'* (S Malhotra 2018, Personal Communication, 5th Oct).

- **Demonization is Energy Poorly Spent.** The argument is that many in the decentralised community are wasting their energy by focusing on an Orwellian two minutes of hate towards the capitalist GAFA's. This of course is a political position, one of many within the decentralised community. The argument is made that the narrative is being hijacked by a regressive activist position, when the objective should be to look past the dominance of the current technology giants and instead look forward to how decentralised technologies can supersede, innovate and drive the next stage of the Internet's development. David Alexander comments, *'I think this has been corrupted to some extent by the language of decentralisation, to try and force people down the activist campaigning view where the man is not involved'* (D Alexander 2019, Personal Communication, 2nd Dec 2016).

This concludes the results section for the Expert Interviews. The following section presents the results of the Practice Led Component of this research. The Conceptual Modelling, Development of User Interface, Critical Reflection, and Focus Groups.

5.5 Practice Led Component Results

The following section presents the results of the practice led component of this research. The component endeavoured to explore the practical realisation of both Human Data Interaction and Self-Sovereign Identity at the interface layer. The objective is not to develop a final user experience and user interface, rather to extend current concepts and prototypes to support the analysis of a full-scale analogue interaction model based on the current trajectory of the domain. Self-Sovereign Identity in the context of a Sovereign Boundary Mechanism has been selected for this practice led component. The rationale for the selection of these concepts is communicated in detail within the introduction of the SSI chapter of this thesis (see: 3). In brief, SSI arguably satisfies the principles of Human Data Interaction, providing a model that offers the user full agency, control, and, ultimately, responsibility for their personal data.

This section begins with information regarding preliminary investigation. The section then continues to present the outcomes of the three distinct phases of practice. The first phase engages a Conceptual Modelling exercise (Johnson & Henderson, 2002), which requires clear steps to be undertaken. The first, is to define the High-Level Functionality enabled by the proposed system which establishes the core system capabilities. The next stage is to define the Major Concepts and Vocabulary. In the context of this research the concepts are linked to SSI technology and developing standards. The method then requires the consideration of a User Profile and Task Analysis. The user in the context of this domain is broad and can be defined as an individual with a reasonable level of computer literacy, with a desire to control their personal data. The task analysis is divided into two sections: a general list of user scenarios and a task analysis of specific operations. The next step, and arguably the most important, is the Objects and Operations Analysis. This stage involves the identification of the required visible objects within the system and their specific operations. Finally, the method requires a mapping of the identified objects and their relationships.

The second phase of the Practice Led Component involved the development of a preliminary user interface based on the outcome of the Conceptual Modelling process. The word preliminary is used as this is a first interpretation of a user interface. A full annotated wireframe has been constructed to aid visualisation and further analysis.

Finally, the practice has been analysed, through a critical analysis conducted by the researcher. This element communicates general observations regarding usability, mental model and cognitive load, while commenting on elements of the system which require further consideration or technical evolution. In order to achieve a high level of validation, the practice has been presented to a selected group of user interface and usability professionals through a focus group, providing a collective critical analysis from a neutral expert position. The outcomes of the focus group make up the final part of this section.

5.5.1 Preliminary Work

Prior to commencing work on this Practice Led Component, the literature was examined, a technology Artefact Catalogue was established, and the majority of expert interviews were conducted. As described in the methodology, this investigation acted as the initial stages of an explanatory sequential mixed methods approach (see: 4.1).

In the initial stages of this practice, a series of whiteboard exercises were conducted to help develop knowledge of the concepts, components, and interactions required for a decentralised eco-system. Figure 24 is an example of an early diagram exploring the component parts of a general decentralised model, a data repository, overarching system state, and related functions. In this diagram the focus, highlighted in red, is on negotiated access control to data.

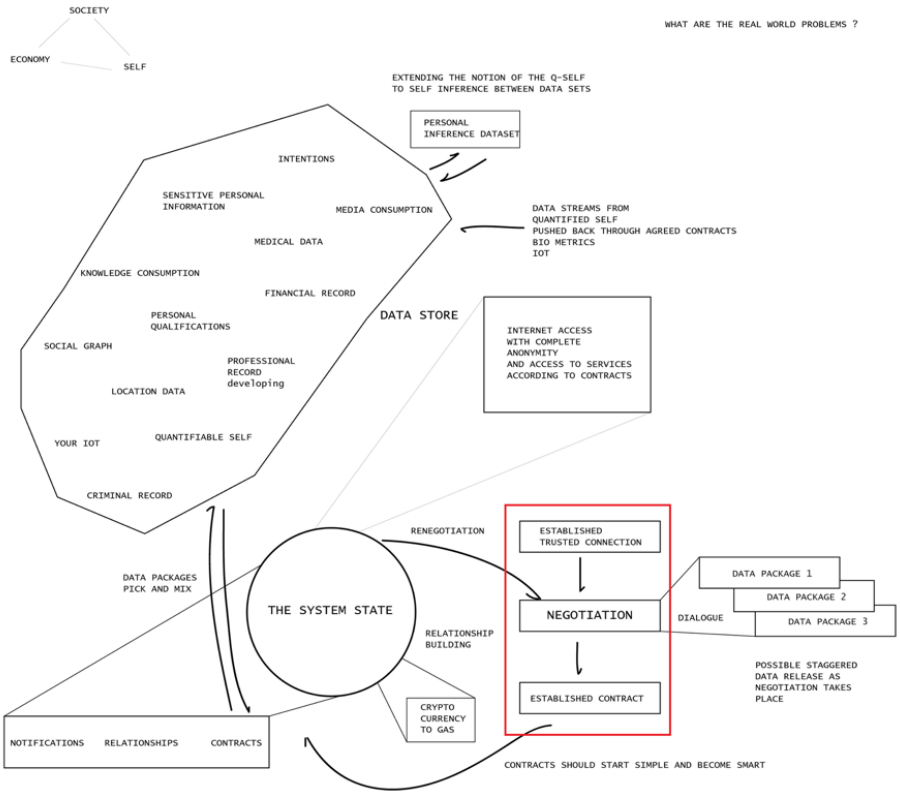


Figure 24: Early Visual Modelling of a Decentralised System

Figure 25 is a later developed example, communicating the component parts of an SSI system. The semi-circle to the right signifies the notion of a boundary shell within which the participant controls mechanisms and functions to engage the wider network transacting data, metered through identity and access control.

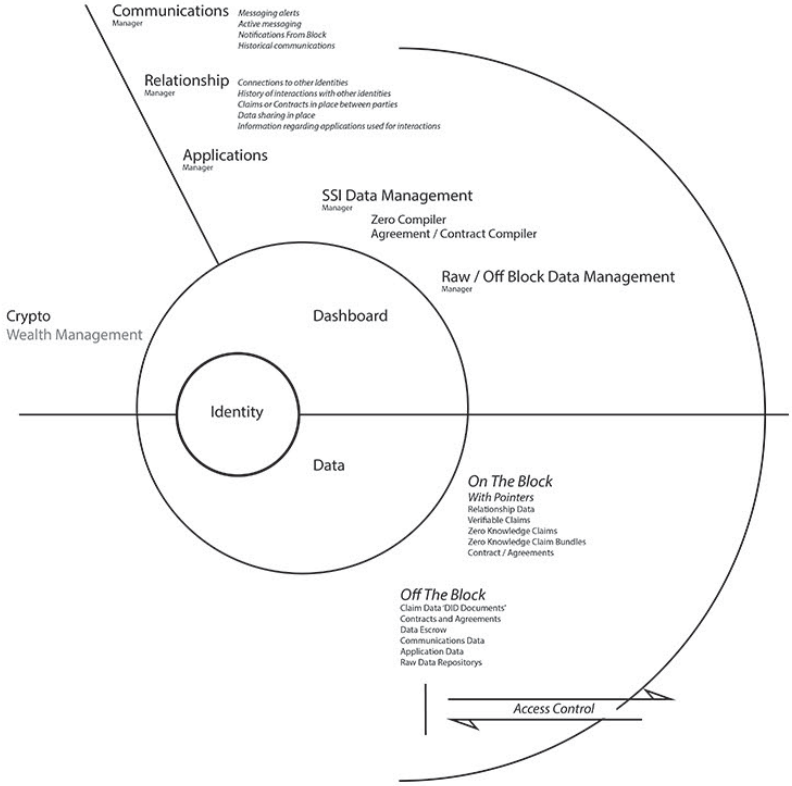


Figure 25: Visual Modelling of a Self-Sovereign Identity System

As the investigation and preliminary planning for the later conceptual model and subsequent user interface unfolded, the Evernym *Connect Me* (2018) SSI wallet application was dissected and laid out (Fig. 26) before an analysis was undertaken to understand how their approach to interaction and interface had been designed. This was done in co-operation with Evernym, through dialogue with their lead developer and product owner. The application was selected as the time, October 2018, it was the only available example of an SSI client-side wallet application.

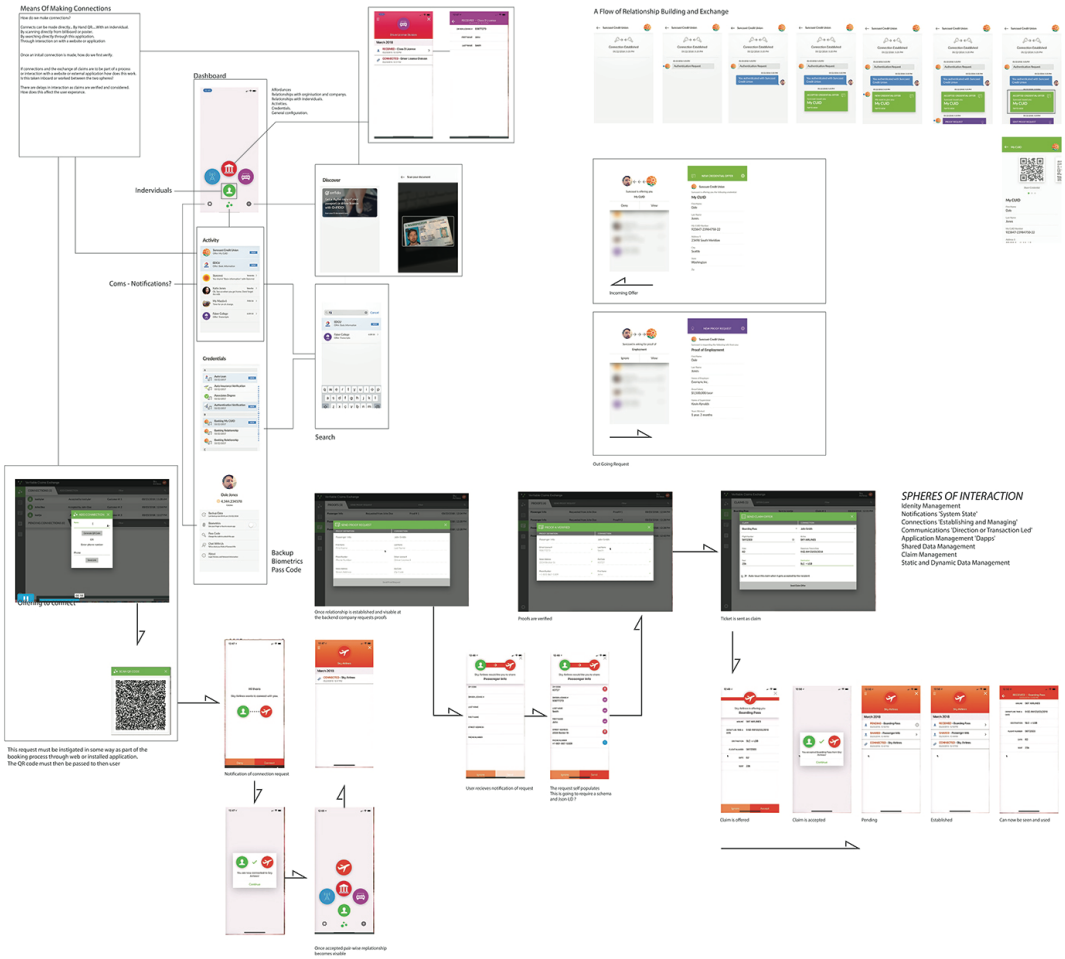


Figure 26: Dissection & analysis of Evernym's *Connect Me* SSI wallet prototype. An enlarged version of this graphic can be found in the appendix, See: G1.

To reiterate, the objective of this practice was not to develop a final user experience and user interface, rather to extend current concepts and prototypes to support the analysis of a full-scale analogue interaction model. Evernym's *Connect Me* represents one of these prototypes. It is limited in its functionality and is a proof of concept in a limited lab-based scenario. In the case of figure 26, it illustrates both the front and backend of a system in which the user presents Verifiable Credentials to streamline the purchase of an airline ticket. When considering this prototype, there are some interesting questions regarding how this looks at scale, and how this manifests across a full-scale collection of interactions to satisfy the full gamut of HDI and SSI principles. This was the question posed to Phil Windley, Chair of the Sovrin Foundation, by

the researcher, during a demonstration of *Connect Me* at MyData 2018. The response was, *'we have yet to see,'* and that *'there is still work to do'*. The Practice Led Component of this research, supported by preliminary investigations aims to do some of that work, and move closer to an answer.

5.5.2 Engaging A Conceptual Modelling Method

Johnson and Henderson (2002) define a clear number of steps which should be followed when developing a conceptual model. The following sections present the outcomes of this process.

5.5.2.1 High-level Functionality

The first step of the process is to define the high-level functionality enabled by the system. The following definition has been considered in the context of a general Decentralised Internet, and the defined principles of Human Data Interaction (see: 2.4) and Self-Sovereign Identity (see: 3.2), alongside current and evolving technology development and standards.

The High-Level Functions required by the system are to:

- Manage Digital Identities
- Manage Connections and Relationships
- Establish A Boundary Control
- Facilitate Transparency of Data Usage
- Facilitate Secure Encrypted Channels
- Facilitate Credential Exchange and Management
- Transact Data with Minimal Disclosure
- Support Trust Networks
- Facilitate Portability
- Enable Data Repository and Binding

5.5.2.2 Major Concepts and Vocabulary

The Second step is to establish an annotated list of the major concepts embodied within the application. Some of these are clearly defined, functional and available, and some are still in development. The vocabulary is important, not only to aid in communication for designers and developers, but also as we begin to establish understanding of new concepts from the participants perspective through metaphor.

The core tangible components and concepts required by the system are identified as:

- **DID -** A Persistent Sovereign Decentralised Identifier, Paired Uniquely for every relationship to disrupt correlation.
- **DID Auth –** Cryptographic method of proving control of DID private key and method of authentication.
- **Verifiable Claim –** An independently verifiable claim about an object, organisation or individual, signed by all relevant parties.
- **Zero-Knowledge-Proof –** A proof of single or multiple claims, compiled with limited exposure of information.
- **Machine Readable T&C's and Contracting** – Method of establishing terms of claim, proof and data collection, distribution and sharing.
- **Distributed Ledger –** An immutable record of transactions and truth.
- **Identity Data Binding –** Means of binding a decentralised data repository to an established identity.

5.5.2.3 User Profile and Task Analysis

The method requires that a user profile should be established before moving to define the system's core tasks. In the context of this system the user base is broad, and essentially a sovereign individual with a competent level of computer literacy, and a desire to control their personal data. The task analysis is divided into two sections: a general list of user scenarios, and a task analysis of specific operations within the system.

In order to consider the broad range of applications, user scenarios have been generated that engage the central mechanics of SSI: defining relationships, building identity, authenticating transactions, managing data, compiling claims, and the redistribution of credentials. These activities are considered within practical applications, as for example, in authenticating a credit card transaction, engaging an Intent Casting application, or gathering data from IOT devices to share with third parties. The compiled list broadly fits into four categories. Connection, authentication, and the gathering and sharing of data. This list of scenarios is by no means exhaustive, but careful consideration has been given in establishing a list of required interactions, for a broad range of engagements.

Table 4 provides a simplified list of the considered scenarios; a full illustration can be found in the appendix of this thesis (see: D1).

SSI CASE STUDYS	
	CONNECTION
	Establishing A Pairwise Identifier with an Individual in the field, or generate and send.
	Establishing A Pairwise Identity through a Website
	Establish A Pairwise Relationship from a public DID
	AUTHENTICATION
	Sign Into A Website With a Pairwise Relationship
	Authenticate a Credit Card Transaction Online
	Authenticate a Credit Card Transaction Within A Retail Environment
	Authenticate Identity With A Physical Space
	Authenticate In The Field to Access a door
	Authenticating A Ticket At A Gate
	SHARING DATA
	Establish relationship and terms of data use when visiting website
	Prove your eligebility to hire a vechical, age, licence capbility to pay, additional insurances
	Supply a Number Of Proofs For An Employment Licence
	Apply for Credit Providing proof of employment, address and income
	Share realtime Data for an IOT health data device with medic
	Share Purchase History and Finiacal Position With An Intent Casting Application
	Provide A Claim For A Personal Reference
	Provide a Claim To Allow Child On School Trip
	DATA GATHERING
	Request A Claim Of Education Record
	Establish a Repository For IOT Data
	Download and Redistribute Social Network Data

Table 4: Simplified version of Conceptual Model User Scenario table

Developed in parallel with the user scenarios, is a list of system tasks. Table 5 provides a simplified list of the required internal tasks for the listed scenarios; a full illustration of this table can be found in the appendix of this thesis (see: D2).

SSI System Task Analysis

	Task Category	Task Title
	CONNECTIONS	
	Connections Establishing DID's - Generation and Send All of these scenarios involve the generation and transmitting of DID Addresses. The did is then picked up by the opposite party via either direct address input or scanning QR. Once DID information is picked up by Agent, a reciprocating DID can be sent in response	**Generate new DID and Send**
	Respond to incoming Request / or in the field	**DID Pairing**
	Needs Further Consideration	**Publically visible DID**
	AUTHENTICATION	
		Pushed Authentication
	CONNECTIONS ***'Historys and Management'***	
	Accessing information about existing pairwise connections. Prerequisite for the following tasks.	**Accessing Pairwise Information and Interaction History**
	Subset Action once Pairwise connection is accessed.	**Accessing Pirewise connection details and subsequent actions**
	Subset Action once Pair-Wise connection is accessed.	**Sending direct Message to Pairwise DID**
	Subset Action once Pair-Wise connection is accessed.	**Requesting a Claim**
	Subset Action once Pair-Wise connection is accessed.	**Requesting a Proof**
	Subset Action once Pair-Wise connection is accessed.	**Data Binding to DID**
	Subset Action once Pair-Wise connection is accessed.	**Claim Request**
	Subset Action once Pair-Wise connection is accessed.	**Proof Request**
	Subset Action once Pair-Wise connection is accessed.	**Claim Send**
	Subset Action once Pair-Wise connection is accessed.	**Direct Message**

Table 5: Simplified version of Conceptual Model System Task Analysis table

5.5.2.4 Objects and Operations Analysis

The Objects and Operations Analysis required the identification of the conceptual objects that the user can see and manipulate. Once the objects are identified, their attributes are specified and listed. Operations within a given object are considered, together with their shared relationships. As the conceptual model develops, outstanding and resolved issues are listed. Within the simplified table that follows, the identified, required, 29 objects are listed. A full illustration of all elements can be found in the appendix of this thesis (see: D3).

SSI OBJECTS AND OPERATIONS ANALYSIS

	IDENTITY MANAGEMENT TOOLS *'Constituting a Wallet'*
	System State 'Notifications / Activity / Data Flow'
	ADMIN AND AUTH
	Administration
	Push Authentication
	CONNECTION
	DID Creation
	DID Distribution
	CONNECTION *'Incoming and Pairing'*
	Inputing DID 'By String'
	Review DID Doc and Create Pairing
	CONNECTION *'Established Connections Management'*
	Listing Of Established Connections
	History of Interactions
	Detailed Pairwise DID Information
	History Log Detail
	CONNECTION SECTION 'Functions'
	Direct Message
	Claim Request
	Proof Request
	Claim Send
	PROOF BUILDING
	Auto Population and Confirmation
	Selecting Alternative Zeros
	DATA MANAGEMENT
	Sent Proof Listing
	Sent Proof Dig Down
	Received Proofs
	Received Proofs Dig Down
	Received Claims
	Received Claims Dig Down
	Sent Claims
	Sent Claim Dig Down
	DYNAMIC DATA
	Data Binding
	Existing Data Binding Listings
	Data Binding Dig Down

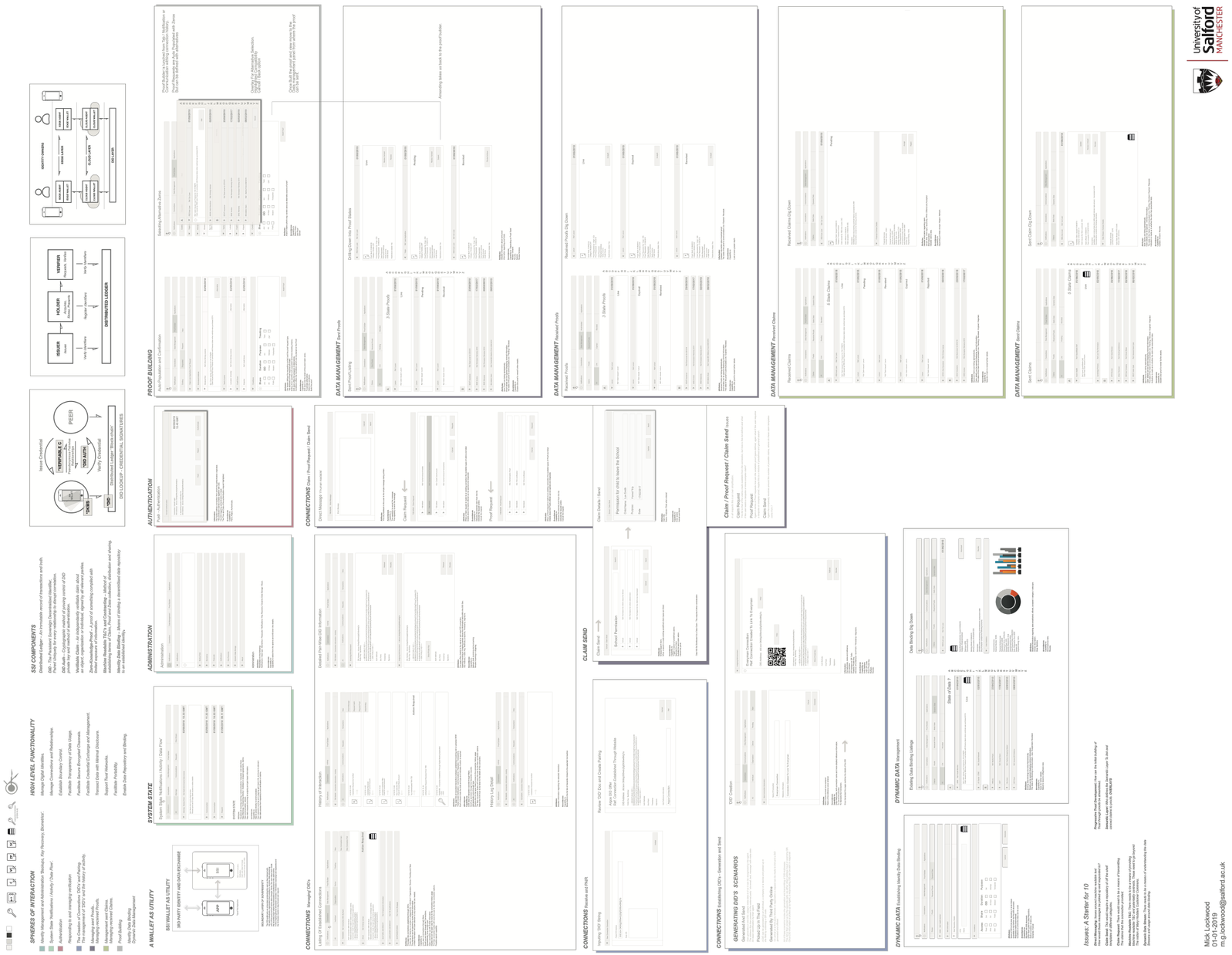
Table 6: Simplified version of Conceptual Model Objects and Operations Analysis table

5.5.3 Preliminary UI

With an Objects and Operations analysis in place, an initial user interface has been developed. This can be found within a folded section on the following page. At this point this is a basic wireframe using a simple table layout. A panel for each of the identified objects has been created, supplemented by a description of the required attributes and functions.

To aid in the reading of the wireframe, a colour key has been provided highlighting the identified Spheres of Interaction, these relate directly to the Objects and Operations analysis. A link to a PDF copy of this diagram can be found in the appendix of this document (see: D8).

Figure 27: Preliminary User Interface



5.5.4 Critical Reflection

The Conceptual Model suggests that a minimum of 29 objects are required for a functional Sovereign Boundary Mechanism. When observing the resulting UI, it is evident that this is complex, and in terms of a cognitive load it can be argued that it is significant. The user experience requires a clear interpretation and understanding of the system state, within which the interactions are reliant on complex sequences, there are a number of sub-domains within the system and a general understanding of the majority of objects and their relationships would be necessary in order to enable confident engagement. In subsequent phases of development, efforts can be made to reduce complexity, and many of the system processes might be reordered, automated or streamlined following usability testing. It remains however, that a User-Centred Data Ecosystem through Self-Sovereign Identity, offering full agency through a Sovereign Boundary Mechanism presents a collection of original concepts and interactions, which may prove challenging in the context of a participant's mental model.

It can be argued, based on the evidence derived from the Conceptual Model and subsequent UI, together with the defined principles of HDI and SSI, coupled with the value of personal data, that a Sovereign Boundary Mechanism requires a considerable amount of internalised understanding before meaningful interaction can be achieved. The introduction of such a system introduces considerable friction and is a backwards step in the context of contemporary user interaction design. The notion of sovereignty and independence place a great weight of responsibility on the participant, and potentially results in what is considered later in this thesis as the Paradox of The Sovereign Boundary Mechanism (see: 6.4.2).

It is important to separate what can be argued to be a high friction demanding user interface experience, and the internalised knowledge and understanding of system and concepts that will be required in order to engage. Critical reflection suggests that the dominant issue in any future development of an analogue Self-Sovereign Identity system, is not the physical interface

design. Many office management tasks, media editing tools and social networks, require engagement with complex UI structures and interactions. It can be argued that, in this context, the dominant issue is that of the understanding of concepts and mechanisms. An interesting comparison has been voiced by Paul Wilshaw, Head of Innovation, Barclays UK, who shared his experiences of introducing *Ping-It*, Barclays' application which allowed for the transfer of funds, via a mobile phone number. Paul recalls that communicating this single, simple, concept proved to be incredibly challenging (P Wilshaw, Focus Group, 15th March 2019). Within a Sovereign Boundary Model of interaction there are potentially multiple novel concepts that lack precedent, existing mental model or metaphor to allow a participant to build a sufficient internalised understanding.

Below is a none exhaustive list of potential original concepts which may prove alien to a new participant:

- Identity Sovereignty
- Multiple Identifiers
- Individual Unique Relationships with Peers
- Peer to Peer Relationships and Equality
- Establishing Individual Relationships to Engage Services
- A Metaphorical Boundary between the Participant and the Wider Network
- The understanding of and the managing of Static and Dynamic Data
- Data Binding
- The Blockchain as a Source of Truth
- Verifiable Credential and the Cascade of Validity
- The Issuing of Credential
- The concept of Proving Truths through Partial Data Disclosure
- Finding Faith in Distributed Storage
- Finding Faith in Self and Responsibility of Managing Presence
- Understanding the Relationships between the Digital and Physical Manifestations

Initial critical analysis of the model also raises some interaction challenges and potential incomplete mechanisms, which need to be considered. These issues manifest around the edges of the core interaction and include:

- **Direct Messaging:** Issues concerning machine readable text. How would these messages be recognised, processed and responded to at scale?
- **Claim / Credential Sending:** This would require a repository of off the shelf templates across different categories.
- **Claim / Credential Request:** There would need to be a means of transmitting the claims that the connected entity provided.
- **Machine Readable T&C:** There needs to be a means of providing Machine Readable T&C and potentially Contracts. These need to go beyond the notion of Mary Hodders *Customer Commons* (Hodder, 2019).
- **Dynamic Data Streams:** There needs to be a means of understanding the Data Streams and usage around Data Binding.
- **Progressive Trust Development:** How can the initial building of Trust through Proofs be streamlined?
- **Semantic Layer:** Who defines the Semantic Layer to link claim elements to Zero Knowledge Proofs?

5.5.5 Focus Group

As part of the verification process, expert opinion was sought via a focus group. The participant list, agenda and process can be found in the methodology section of this thesis (see: 4.5.4). Specific questions where posed relative to a participant interaction with an SSI system, through the previously described developed conceptual model and user interface. The following section documents the responses, consensus, and emergent themes.

5.5.5.1 Reaction Rather than Response

During the focus group there were two instances where reactions rather than responses revealed what can be described as an emotional return. The first occurred when the UI was revealed as

a slide during the presentation. The pre-distributed website included a page dedicated to the user interface, but it transpired that the team from the BBC hadn't seen it. When revealed, this resulted in an audible gasp, followed by the comment, *'Now that's a work of Art'*. The response and phrase are interpreted to convey the shock at the scale of complexity and the required components in the model. The next instance came when the second question was posed, *'What are your views on the complexity of these interactions?*' In this instance the end of the question was met with an awkward silence, followed by laughter. The focus group was well spirited, comprising a group of highly creative people, with many years of industry experience. Although individuals where highly critical, they were also professional problem solvers, seeking constructive solutions to the issues encountered. There now follows a summary of themes and responses relating to the predefined questions.

5.5.5.2 How Understandable are the Concepts?

When this question was posed, a consensus was that the concepts will be problematic for individuals and pre-empted the later questions regarding metaphor. The basic understanding of the notion of proving something through a verifiable digital credential would be understood, but that the broader peripheral mechanisms are not relatable. Jo Cox commented: *'I don't even talk about having a relationship with the corner shop when I buy a paper, or the airline when I buy a ticket, or Facebook who hosts my Social Network'*. She went on to discuss the need to consider the language around any developed model:

I don't have relationships and establish those at all, and don't take personal information and plant access to get access to the system. So yeah, I think this vocabulary needs changing, the underlying concept of having a verified identity from somebody you trust and being able to prove your identity using it is completely sound, but that's as far as it goes (J Cox, Focus Group, 15th March 2019).

Paul Wilshaw made some very interesting comments about his work for Barclays Bank, and the lack of understanding of what was arguably a very simple concept. He recalled:

I worked on Barclays Pingit, back in 2012, and that was hard to market, because of the concept of a digital wallet. People didn't get it you could transfer money via mobile phone number. It was an alien concept, and probably took the best part of three years for the concept to get out (P Wilshaw, Focus Group, 15th March 2019).

Tim Panton made interesting comments about his work in Germany, and the need to present concepts in context, in a form that individuals understand. His points centred around German families and the protection of data relating to children, baby monitors and remote access cameras. He suggested that this was a niche market, that people understood and were motivated to protect, he explained *'you have to get it at that real concrete level that people care about, otherwise you can talk about abstract concepts all day and people say "yes". And they don't do anything about it'* (T Panton, Focus Group, 15th March 2019)

The general consensus to this question in the context of the system presented, was that its concepts would be alien to most people, and that they would struggle with the abstraction of the components. This would be beyond all but the most technical and motivated of individuals. The system needed simplification, it needed to be put in context, and niches of value needed to be found in order to drive interest and engagement.

5.5.5.3 Views on the Complexity of these Interactions?

Following initial laughter from the group, the posing of this question raised conversation about the acceptable level of complexity. This was followed by comments by Paul Wilshaw arguing that this should not be complex, and where possible, the interface layer should be removed. He argued that, *'it shouldn't be complex at all, the more you talk about it the more it shouldn't even be on a screen'* (P Wilshaw, Focus Group, 15th March 2019). Points were made regarding the participant needing to see something in order to understand the value and comprehend what the system was doing. Tom suggested a separation of layers to support simplicity as he argued:

I think you have to separate out, what is a middleware workflow which probably isn't designed in a user interaction fashion, from a series of application workflows, and these are absolutely the applications you can design from a UX perspective (T Cheeseman, Focus Group, 15th March 2019).

Comments were made regarding reducing the friction found in the proposed system, with gesture and biometrics. Tom commented on the system presented as an attempt to visualise the whole environment, and that in any redesign, a great amount of the initial interface might be hidden from the user. He commented:

I think this is an attempt to lay this out, to give a consumer understanding, a unified concept of this Sovereign Identity as a single whole. And I think even if that exists it would be like the classic Iceberg problem; I only need to see 10% of it (T Cheeseman, Focus Group, 15th March 2019).

Following Tom's observation regarding the hiding of the mechanisms, Lisa Ortega then went on to suggest that this was a general problem of any complex system, hiding what the user doesn't need to see to reduce friction, but then having to face the problem of what to do if at any time hidden elements are needed. *'That's that problem isn't it, if you do need to access any of that 90% how do you do it?'* (L Ortega, Focus Group, 15th March 2019).

The general consensus to this question, in the context of the system presented, was that the current manifestation is over complex, that many of these issues might be addressed, but that a considerable reduction in friction needs to be achieved. Within the conversation, it was recognised that a balance between exposed and hidden interaction needs to be considered if the user is to fully understand the system's functionality.

5.5.5.4 How do we build something that's Accessible?

Once posed this question immediately lead to a debate about what the user does and does not need to see or understand within the system. Jasmine commented on value and accessibility: *'I*

think when you talk about the value, that value is not that it's secure the value is something else, that makes it easier. You know, it's a click away' (J Cox, Focus Group, 15th March 2019). This is a very pertinent point, as it relates to the idea that the Decentralised Internet cannot be marketed. The debate continued with a discussion about the Canadian Government's move to begin to issue verifiable credentials, and *Evernym's* initiative to work with Canadian credit unions and banks to develop digital wallets, first populated by financial credentials, before being introduced to additional applications once a Self-Sovereign Identity has been created. Tom highlighted the fact that many attempts to develop verifiable credential technology have failed, and that this mechanism might act as a solution. Tom commented on some of the recent attempts to establish digital verification systems, *'Verification is absolutely broke, GDS seems to be falling apart, Verifies a total failure'* (T Cheeseman, Focus Group, 15th March 2019). The conversation progressed, and the group began to discuss who and what people can trust. Ian Forester raised the issue of solid ground, and the fact that at some point people need to be able to trust in something commenting: *'the example we gave about using a phone, they were saying, well do you trust your phone, and it's like, okay, do you trust a fingerprint hasn't been tampered with, we got to trust something'* (I Forester, Focus Group, 15th March, 2019). Tom began to discuss the notion of trust frameworks with the statement: *'My understanding of why or how this works, is that part of the reason for having a single identity, based on these relationships, is that they come together to reinforce that identity overtime, to reinforce trust'* (T Cheeseman, Focus Group, 15th March 2019). The conversation around trust evolved, and Marc Tobia made some very interesting points about potential exclusion, with regard to the issue of individuals being unable to trust themselves in managing what might be high value personal information and data. He commented:

I think that it may be creating an anxiety in people, if they don't know how to manage their data, and so that creates another layer of words like, so if my data gets out there, then it's my fault, and so that creates anxiety with him. You're like, I don't know if I want to deal with

that, and so then maybe it's easier for me to just go to a third party, and be like, you deal with that (M Tobia, Focus Group, 15th March 2019).

The consensus to this question, while only being answered in part, was to reduce friction in the interaction, and to build a system that people could trust, while developing confidence in their own capabilities.

5.5.5.5 Mental Model and Metaphor

The discussion around this question followed on from previous debate with the researcher asking, *'so there's an acknowledgement that in this as it stands, there would be a lack of understanding in terms of the mental model and metaphor?'* Paul Welshman's view was telling, with a one-word answer. *'Totally'* (P Wilshaw, Focus Group, 15th March 2019). Tim continued to communicate his company's experience of cultural differences and individuals' understanding of privacy:

I think our experience, that the way that people respond to these concepts are very much dependent on the culture they're in. I mean our experience in Germany is completely different from this country, and our experience in the Far East just different again, where privacy is a concept but it's around the family, it's not individual privacy. The Germans really understand individual privacy, and the Americans don't understand privacy at all (T Panton, Focus Group, 15th March 2019).

Paul referred to the different types of documents we have, and the different requirements of companies, alluding to the confusion in the forms of identity, with an argument for standardisation. He argued that: *'the concept of identity is different for every single company, what they need, and what they want, like we've got driving licenses and passports, and like should they be the same thing'* (P Wilshaw, Focus Group, 15th March 2019).

Lisa Ortega voiced concerns that the mental model for this kind of system could be tinged with fear and nervousness about being locked out of services in the real world, if you don't subscribe to or understand the ecosystem. She commented:

I think it's interesting that we're talking about a kind of fear and nervousness, and if I use this or didn't use it, what service do I get locked out of, online services and real-world services? What can't I use if I can't connect, if I don't have any power on my device, or if I've forgotten some key information? It's lots of stuff again that's reliant on you (L Ortega, Focus Group, 15th March 2019).

Marc made interesting points, aligning the development of a mental model for this kind of engagement to his own experience with crypto currencies. He commented:

I tried adopting cryptocurrency, like a year or two ago and it was getting the wallet and everything. It was just like this is too much for me. And then it was also the thing of because other people weren't using it. I wasn't having an understanding of the real framework of how this can be applied. So, it just got to a point where I was just like, you know, I just don't even want to deal with this. (M Tobia, Focus Group, 15th March 2019).

He continued to argue that this sort of system might require an introduction and instructions as opposed to a system that might be picked up independently: *'I feel like they would still probably pick it up through a third-party first, before someone's going to go out and just engage with it on their own'* (M Tobia, Focus Group, 15th March 2019). Both Paul and Tom expressed concern that in its current form, individuals simply wouldn't be able to grasp these collective concepts: *'I think some of these Concepts without living and breathing them would be difficult to grasp'* (P Wilshaw, Focus Group, 15th March 2019). *'I don't honestly think they will understand'* (T Cheeseman, Focus Group, 15th March 2019). Finally, Paul commented again on the need for standardisation: *'I think the real solution to this is the standardisation of what information we need to provide, and that that's where it falls down'* (P Wilshaw, Focus Group, 15th March 2019).

The general consensus around the question of mental model and metaphor was that in its current state, the system would be difficult to comprehend and understand, and that standardisation might play a role in simplification while supporting understanding. There was also agreement that the lack of understanding of the system might exclude certain types of individual, and that on the whole, participants would need assistance in understanding and adopting the concepts.

5.5.5.6 Can any of this be Automated?

When this question was posed to the group the instant collective response was 'Yes'. The conversation turned quickly to AI, and the issue of an individual's comfort level with allowing a third party to manage their affairs, and the argument that once you relinquish control to a third party, you are no longer sovereign. Tim made the point that AI can be sovereign, that it *'can be yours'* and that it can operate under your control.

Paul discussed: *'Robotic process automation'* and contrasted the configuration of that, to potential configuration of a personal AI or personal Jeeves, '*you can control the workflow yourself*' and that *'it was really easy to do'*. Paul went on to discuss the trajectory of technology, and that AI was quickly becoming part of everyday life and would soon become the norm. Paul drew the conversation towards authentication by commenting that even with AI doing much of the mundane tasks for you, there is still the issue of authenticating the peer on the other side of the transaction. (P Wilshaw, Focus Group, 15th March 2019). Ian discussed having agents working for you that do so automatically and can be inspected by you if you choose to do so. He commented: *'Agents that work for you, but you can inspect them if you want to, they're under your control'* (I Forester, Focus Group, 15th March 2019). The concept of inspection is considered important, the way we can trust AI and how the participant interacts with and inspects the AI system.

Finally, interesting points were made about the different trends and direction of travel of both Apple and Google. Apple processing data and running AI locally, and Google dealing with the computation in the cloud. It was argued that this kind of decentralised model may sit more comfortably in an Apple Universe.

The general consensus that emerged from the conversation regarding Automation and AI is that it could solve many of the complexities and frictions identified, that individuals are becoming accustomed to AI, that AI can work for the individual in a sovereign way, and that

the means and methods of inspection were an important consideration. Concerns were voiced that even with AI, authentication of actors is still an issue, and finally, that differing philosophies of where data is processed is an important consideration.

This concludes the results section for the Practice Led Component. There now follows a summary section followed by a discussion chapter, where the results are considered alongside the literature and research questions.

5.6 Results Summary

This chapter presented the results generated from 4 components of primary data gathering. The initial results presented for the Artefact Catalogue included the number of elements considered and the chronological span. The section communicated the 12 measures used across the evaluation matrix and showcased a visual colour coded section of the overall outcome. The section proceeded to define the 16 categories of artefacts, with a detailed explanation of each supported by exemplifier. A visual mapping of artefacts is presented chronologically, which supports the communication of the domain trajectory towards Self-Sovereign Identity. The section proceeded to present the pertinent findings drawn. These included the impact of the identity layer, the driving trends and the development of business models. Finally, commentary was made on the value of artefacts and the observations made regarding usability and specifically mental models.

The chapter progressed to consider the Public Survey. Initially, the participation and demographic and the raw descriptive statistics were reported across all questions, with significant results including, public concern, trusted sectors, and societal benefits. The section then proceeded to consider the Likert and Forced Binary scales, and their reliability and significance. The section continued to present the weak correlation and comparison statistics. Finally, the section presented the unforeseen longitudinal component, with an interesting observation made regarding measures taken to protect personal data that emerged following the Cambridge Analytica revelations.

The chapter then proceeded to convey the outcomes of the Expert Interviews, the stages and the participation. The initial stage of basic thematic analysis and its influence over the second phase of interviews is reported. The inclusion of a third phase of data generated through the focus groups was communicated, before the section moves to present the substantive thematic analysis and subsequent results. A visual mapping of the 64 generated themes was illustrated,

followed by a detailed description of each within the 3 main categories of adoption, cognitive load, and general concepts.

The final section of this chapter presented the Practice Led Component, initially this consisted of the text-based outcomes of a conceptual modelling process, culminating in the generated tables of user scenarios, user interactions, and objects and operations analysis. The section then moved to illustrate the development of preliminary user interface and subsequent critical analysis, highlighting original concepts which may lack existing user mental model. Finally, the section identifies components of the wider User-Centred Data Ecosystem which still need further consideration and development.

In the following chapter, the results of this research are discussed alongside the research questions and relevant literature, exploring the generation of potential new knowledge, and the defining of findings and conclusions.

Chapter 6.0 Discussion

The following discussion chapter examines the results of this research. Each component is considered in turn, and the pertinent topics are examined relative to the research questions and existing theory. The Artefact Catalogue is debated, identifying the trajectory of the decentralised domain, the impact of an identity layer, the complexity of a Sovereign Boundary Mechanism, and the values found in contemporary and past artefacts and projects. The Public Survey is discussed, a comparison with existing surveys is made, alongside a commentary of the descriptive statistics of interest. Scales of significance are communicated, followed by a discussion concerning the correlation and comparison statistics and the advent and relevance of the unintentional longitudinal element. The expert interviews are investigated across the 3 stages of data gathering, focusing on the core derived themes of noted significance. Finally, the Practice Led Component is discussed, through the communication of core findings: The concept and paradox of a Sovereign Boundary Mechanism, the introduction of friction, the potential for missing mental models, the issue of internalised cognition, the relevance of a spectrum of HCI theory, and finally, the balancing of cognitive load against value proposition.

6.1 Artefact Catalogue Discussion

The rationale for developing an Artefact Catalogue is to investigate the scope and scale of work that has already been undertaken, and its alignment with the decentralised objective. The catalogue is not a definitive list, rather a broad slice of related artefacts. The cataloguing process aims to understand the concepts, the driving technologies and the value propositions of existing endeavours. It aims to gain an understanding of the level of friction and complexity illustrated in the concepts. Focus has been drawn, not only to contemporary ideas, but also to now defunct concepts and projects which may still have relevance and value given recent technological developments. Importantly this process aimed to understand the trajectory and trends within the Decentralised domain. There now follows a number of themed areas of discussion, derived from the cataloguing process, that are considered to be most relevant.

6.1.1 The Impact and Relevance of an Identity Layer

The progression of the Federated Identity layer towards Self-Sovereign Identity, as illustrated in the work of Christopher Allen (see: 3.1), can clearly be seen within the Artefact Catalogue as we observe the trends across Decentralised Trusted Identity, The Centralised Hub, and ultimately SSI (see: 5.3.4.2). When we consider the primary mechanism of the centralised model, the control of identity, whether that be through loaned or federated identity, identity derived from registration with online accounts and services, or identification of a user through tracking methods. Unlocking the participant and providing a genuine means of identity ownership and control, would seem to be the final obstacle in the progression towards genuine decentralisation, and the disruption of the *Domains of Identity* (Young, 2018).

The advent of SSI focused Blockchain technology (hyperledger.org, 2019; Veres.one, 2019; Uport.me, 2019, Sovrin, 2019) and its capability to facilitate an identity layer is a significant step forward. Providing the capability for participants to generate sovereign verifiable identifiers at scale, enables powerful SSI mechanisms and interactions.

- A participant can generate multiple persistent identifiers that do not rely on any centralised authority.
- A participant can verify ownership and control over any given identity through encryption methods.
- By pairing identifiers participants can establish an encrypted communications channel.
- The participant can generate a new identifier for every transaction undertaken and choose to link and verify one identifier against another.
- Participants can sign documents, accept and issue verifiable credentials and combine and redistribute part or full elements of a credential.
- Identity's and credentials can be verified cryptographically by any peer within the ecosystem.

The concept of multiple identifiers controlled by an individual, organisation or object, disrupts the tracking and profiling mechanic of the centralised model. The extended functionality made possible through an identity layer; addresses the asymmetric power imbalances we currently observe across the network. Observations drawn from the Artefact Catalogue, would suggest that the claims made by Jeremy Rushden that technology is no longer a barrier to decentralisation, would seem to be correct (see: 5.4.1.4). And given the capability a functional identity layer enables, the claims by David Alexander with regard the game changing impact of SSI, would also seem to be justified (see: 5.4.1.6). The capability to generate and control practically thousands of identifiers on the network, and engaging mechanisms that allow complete agency over personal data is a powerful prospect. Managing this at scale from the prospective of the individual raises several obstacles, and it remains to be seen if Tobin's claim that the rise of SSI is inevitable can surmount the rather significant challenges of participant engagement and usability (Tobin, 2016).

6.1.2 The Pursuit Towards a Sovereign Boundary Mechanism

At the time of writing, December 2019, the drive towards SSI is spearheading the decentralised communities' efforts. In the pursuit of a sustainable business model, artefacts have chosen to limit sovereignty and control components of the decentralised system, to arguably retain some degree of leverage around which to develop a business model. Staking a claim to semi decentralised models at the early stages of evolution is one approach, but investigation suggests another trajectory that is driving for a fully sovereign decentralised system. This direction of travel looks past the realisation of a fully functioning Human-Centred Data Ecosystem, developing applications and supporting services that can be employed once a genuine ecosystem is achieved. An observed causality dilemma is the need for a number of components to be in place before the model becomes self-sustained. These components were communicated by Kaliya Young, during an expert interview. Kaliya argued that we would begin to see a mature ecosystem, once developers can focus solely on one of the following: Wallet and agent code, issuer and verifier code, and code driving the ledger, without having to develop every supporting element independently (K Young 2018, Personal Communication, 21st Nov). This is arguably unfolding through Evernym, (2018) a company closely associated with the Sovrin Foundation (Sovrin, 2017), who are beginning to develop forward looking wallet solutions and enterprise tools.

In the pursuit of a truly sovereign system in terms of participant interaction, the trajectory would seem to be progressing towards what has been termed in this research as a Sovereign Boundary Mechanism. An independent realm of interaction which enables the functionality required for the principles of decentralisation. This research suggests that an analogue manifestation of an SBM is complex, and high in friction and would require considerable internalised cognition and knowledge in order to engage. The model potentially places the participant in an isolated position, one which generates practical usability and

adoption issues (see: 5.5.4). These issues are considered and discussed at length in the subsequent sections of this chapter.

6.1.3 The Value of Decentralisation

A critical variable in the adoption equation for decentralised technology will be the value proposition they present. A central aim of the Artefact Catalogue has been to explore the current decentralised landscape for existing and emergent value.

When considering value across existing artefacts, the argument that the *Decentralised Internet Cannot Be Marketed* is an important prospective (see: 5.4.1.2). It has been argued that in order to gain adoption, decentralised technologies need to do more. There has to be innovation, and this needs to go beyond the vague common understanding of privacy.

This research focuses on a decentralised trajectory that potentially enables complete agency over personal data. It is important to consider that many decentralised technologies also enable transparency, partial agency, disruptive capability or form part of a broader Human-Centred Data Ecosystem. Technologies such as Bit-torrent and Crypto Currencies offer peer to peer transaction and provide significant disruption within specialised domains. Distributed storage technologies offer a dismantling capability and are arguably a core component of a wider SSI interaction. Technology's which offer privacy protection, such as Ad Blockers, VPNs and Browsers offer clear value proposition, but disrupt the centralised model rather than innovate. Many of the Collation Personal Data Stores which go some way to enable a deeper understanding and management of centralised data streams, are in reality pseudo decentralised, with limited portability and redistribution capability, it can be argued that these artefacts would be better aligned with the concept of the quantifiable self, (see: 2.4.12) rather than genuine decentralised agency. Much of the drive towards SSI through Limited Identity Mechanisms enabling DAPPS, still require a degree of centralised acceptance by the participant, who to a greater or lesser degree is locked into the system.

It can be claimed that without the full agency functionality enabled through an SBM, artefacts are restricted to offering semi decentralised versions of their centralised counter parts, and as such fail to offer the full value decentralisation can potentially provide.

When considering the landscape, many of the artefacts which fall under the category of VRM 'Vendor Relationship Management' warrant further investigation. These technology's endeavour to address the asymmetric power imbalances enabling genuine peer to peer transaction, disrupting Contracts of Adhesion, facilitating negotiated relationships through limited data disclosure. Many of the concepts of intent casting developed to date, demonstrate significant potential value, (see: 5.2.2.5) and could potentially be rejuvenated through the introduction of a functional identity layer. Further investigation of what in many cases are now defunct projects are of value to anybody considering the development of decentralised artefacts. The redevelopment of these concepts with contemporary technology may come with the caveat of increased interactive friction and participant responsibility, but this research suggests that in order to establish the required levels of value to drive adoption, these identified issues need to be recognised and mitigated.

6.2 Public Survey Discussion

The following section discusses the results of the Public Survey, specifically the descriptive statistics of interest, significant data emerging from Likert Scales, data relating to correlations and comparisons, and finally the significance of the serendipitous longitudinal element.

6.2.1 Descriptive Statistics of Significance

Q28 *What concerns you most about sharing your personal data?*

The results of this individual question are significant, with 68.5% of participants citing concerns that they don't have control over how their personal data is shared. The concept of

Control, as a means of communicating privacy harms and the risks associated with the sharing of personal data, has been highlighted elsewhere within this research (see: 5.4.12). The notion of Control is powerful, and this result supports the argument that the narrative of *Being Controlled*, should form part of a communication strategy to drive adoption of decentralised technology.

Q37 *Which sector do you trust the most with your personal data?*

The results of this individual question are significant, with 38% of participants voicing Financial and 34.1% Public Sector. This result is similar to that found within the Catapult, *Digital Trust in Personal Data Survey* (Catapult, 2016) which resulted in Public Sector 43% and Financial Services 28%. It can be claimed that both areas are favourable focal points for initial product development and adoption strategy.

Q38 *Which one of the following would most convince you to share your personal data?*

The results of this individual question are significant, with 58.2% of participants citing *Improving Society* as a motivational driver. This result is similar to that found within the *Catapult Digital Trust in Personal Data Survey* (Catapult, 2016) which resulted in 42% opting for societal gain. Arguments for the affordance of privacy rights and the benefits of data sharing for society are a central argument for decentralisation (Solove, 2008; Pentland, 2012; Van Kleek, 2014; O'Neil, 2016; Monbiot, 2017; Schneier, 2015). The academic arguments aligning with the position of the general public, presents a primary direction for product development, and strong narrative for adoption strategy.

Q41 *Do you read the terms and conditions when agreeing to a service online?*

This is a significant result as it confirms that people don't read terms and conditions, with only 9% answering 'Yes'. The eventual impact of GDPR on the problem of meaningless consent, participant understanding of the outdated Contracts of Adhesion, engagement with terms and access control, and ultimately the potential of Smart Contracts, are predicated on

participant awareness and engagement. There is a considerable amount of work to be done in raising awareness and battling complacency, if the status quo is to be influenced.

Q43 *Have you ever been a victim of what you would consider a fraud, breach or an abuse of personal data?*

With a result of 71% of participants answering 'No', a central justification for the adoption of decentralised technology might be absent. David Irvine made the argument that unless a serious data breach has ever been experienced, participants are unlikely to be interested in decentralised technologies (D Irvine 2016, Personal Communication, 9th Nov). This is compounded further when we consider that the consequences of the majority of data breaches are financial, for which there is a common understanding that insurances are in place to rectify. Adding to this is the general confused picture held by participants with regards risks and harms, which for many will never become a reality (Jarvis, 2011), (see: 2.2.2). This supports arguments around the communication of the positive advantages of decentralisation, (D Alexander 2019, Personal Communication, 2nd Dec 2016) rather than the negative consequences that the majority may never experience. There is though, the hidden exploitation of personal and collective data, individuals are not aware of, and the gathering, inference and secondary use (Van Kleek, 2014). The communication of this type of unconscious self-inflicted data disclosure, running alongside the positive advantages of decentralisation, potentially provides a compelling argument for adoption.

6.2.2 Scales of Significance

Scale: B: *Understanding of the Value of Personal Data,* resulted in M=3.64, from a maximum potential of 5. This suggests a general population with a high perceived understanding of the value of personal data. This result has been derived through a number of questions that explore the process of data collection and the value of data not only to the individual, but also as a broader commodity. The results suggest that the population

understands that data is bought, sold, processed and ultimately exploited by capital, and that there is a general awareness of *Surveillance Capitalism,* (Zuboff, 2015).

Scale: C: *Comfort Level with Network Engagement,* resulted in M=2.28, from a maximum potential of 5. The results across the elements of this scale are consistent. Participants expressed views regarding the fairness of personal data exchange for services provided, the amount of control the participant felt, the trust that data would be kept secure, the perception of inferred data, and over all opinion of the practice of data collection. The results would suggest a tolerant population who are marginally disaffected with the current centralised system.

Scale: D: *Perception of the Importance of Personal Data*, resulted in M=3.95, from a maximum potential of 5. This suggests a population that is highly conscious of the importance of different data types shared across the network. The consistency of results across elements is split, with perception being high in data disclosure which might be obvious. For example, email, file download, location information and online chat. However, a lesser perception was recorded within engagement which might be argued to be more inferred, browsing patterns, search terms, downloaded applications and times of day online. These results suggest a population who perceive their personal data as important at a surface level, but potentially lack an appreciation of the deeper methods of data analysis. This result is interesting when considered against the arguments made by John Laprise regarding statistical literacy (see: 5.4.6.1).

Scale: E: *Effort Made to Protect Privacy* resulted in M=.331, from a maximum potential of 1. This is considered to illustrate a low level of engagement by participants to protect their personal data. Other than clearing cookies and browser history, and deleting or modifying Internet posts, little effort would appear to be made. It could be argued that participants are unaware of the spectrum of more obscure methods available but equally, it

could be argued, contrasted with the *Understanding and Importance of Personal Data* results, that this is evidence of the Privacy Paradox (Norberg, 2007). This is further supported by the results of **Q31** and **Q32,** which both signify that individuals have a strong interest in controlling personal data and an interest in engaging with emergent decentralised technology. However, when asked within **Q33** if current concerns about data privacy would sufficiently motivate participants to actively manage part, or all of their personal data, the answer is contradictory, with 68.9% of participants answering 'No'. Further support is found in the results from, **Q42** When asked: *In all honesty, how concerned about the disclosure of personal data are you?* Participants concern level seemed to be moderate at M=2.77, from a maximum potential of 5.

Scale: F: Willingness to Engage Third Parties, combined **Q34** and **Q35** to define a result which indicates the participants willingness to allow either third party or AI management of personal data. The results indicated a low comfort level with this prospect at M=2.44 from a maximum potential of 5. This is an important statistic as the efficient management of personal data within a User-Centred Data Ecosystem may ultimately require a degree of automation. (see: 5.4.11)

6.2.3 Pearson's Correlation

As part of this investigation, given that the Likert scales are ordinal and exhibit a standard distribution. Each scale has been compared through a Pearson's Correlation, in order to ascertain any significant relationship or trend. The results of these correlations can be found in the results section of this thesis (see: 5.3.6). Of the seven scales compared, only two resulted in a weak correlation:

- A: Understanding of Personal Data and its Value / Effort Made to Protect Personal Data resulted in a correlation of r (291) =.390, p=.000
- B: Comfort Level with Network Engagement / Willingness to Engage 3rd Parties resulted in a correlation of r (291) =.373, p=.000

A stronger correlation within these categories may make a contribution to the way decentralised tools are conceived and disseminated. It is possible that through the better design of Likert elements, and the removal of Forced Binary questions, stronger evidence of correlation may be possible (see: 7.5.3).

6.2.4 Ordinal Dependent - Independent Variable Comparison

A series of independent binary questions have been developed to define differing sample groups, in order to allow comparison with Likert and Forced Binary Scales. The binary groups included: Gender, Suffering A Data Breach, Considered Controlling Personal Data and Sufficiently Motivated to Control Personal Data. The results of these comparisons can be found in full within the results section of this thesis (see: 5.3.7). In the following section only comparisons resulting in a > +/- 5% difference and a P Value of <0.5 are discussed and documented as significant.

In the collection of comparisons relating to Gender, there are 3 notable results. The first, A1 relates to the participants perception of their own understanding of *Personal Data and Its Value.* Within this comparison, Males resulted in +7.2%, p=.000, Hedges' g = 0.46. A2 relates to the participants *Comfort Level with Network Engagement***.** Within this comparison, Males resulted in +5%, p=.009, Hedges' g = 0.33. Finally, A4 relates to the participants *Effort Made to Protect Personal Data.* Within this comparison, Males resulted in +10.7%, p=.000, Hedges' g = 0.51.

These results would suggest that males feel more confident with their understanding of personal data, that they are more comfortable with network engagement, and that they go to greater lengths to protect personal data. These results might be considered relevant when developing tools and considering value proposition for differing groups, but the direct causal effect of these findings is outside of the scope of this research. These results raise the issue of Gender and the way in which different groups perceive and utilise technology. This was a topic

discussed with Kaliya Young during expert interviews, when the issue of single and multiple identities was discussed (K Young 2018, Personal Communication, 21st Nov). The argument being that males are comfortable for all formal and personal communications to go through a single identity, whereas females tend to prefer separation, through multiple identifiers.

In the collection of comparisons relating to *Considered Controlling Personal Data*, there are 2 notable results: C1, relates to the participants *Understanding of Personal Data*. Within this comparison, participants who had considered controlling their personal data resulted +13.4%, p=.000 Hedges' g = 0.91. Suggesting that participants who had considered controlling their personal data, had a significantly better understanding of personal data and its value. C2 relates to the *Effort Made to Protect Personal Data.* Within this comparison, participants who had considered controlling their personal data resulted in a +15%, p=.000 Hedges' g = 0.74. Meaning participants who had consciously considered protecting their personal data, demonstrated a significant increased effort to then go on and make an effort to protect it.

In the collection of comparisons relating to *Motivated to Manage Personal Data,* there are 3 notable results: D2 relates to the participants *Comfort Level with Network Engagement* resulted in a negative -9.6%, p=.000 Hedges' g = 0.64, Meaning participants who are less comfortable with their network engagement where significantly less likely to be motivated to manage their personal data. D3 relates to a participants *Perception of Importance of Personal Data, resulting in* +9.2%, p=.000 Hedges' g = 0.55, meaning participants who have a stronger perception of the importance of personal data are significantly more likely to be motivated to manage it. Lastly D4 relates to the *Effort Made to Protect Personal Data* with a result of +12.7%, p=.000 Hedges' g = 0.58, which it could be argued contradicts arguments surrounding the Privacy Paradox (Norberg, 2007), as this result indicates those who are motivated, do actually go on to make efforts to protect their personal data.

6.2.5 The Insights from a Longitudinal Study

As described in detail within the results chapter (see: 5.3.8) the Public Survey inadvertently yielded a longitudinal element, when a pause in data gathering was interluded by the Cambridge Analytica scandal. Following this revelation, a period of time was allowed for coverage in the media to dissipate, and for the raw public awareness of the event to pass. Within the analysis of the survey data, three Likert scales, and two binary choice scales where compared to ascertain whether the impact of such a seismic personal data controversy could affect the opinion and views of individuals. This element was not planned, the comparison group sizes are different, and the validity of this element could be debated. The observation though is of interest and is worthy of consideration.

Out of the 5 scales considered, pre and post Cambridge Analytica, 2 scales resulted in moderate gains of between +0.2% and 4.8%. E3: *Has the Participant Considered Controlling Personal Data,* saw a significant gain of + 6%, p=.000, Hedges' g = 0.47. The notable variance is found within E2: *Effort Made to Protect Personal Data,* which saw a change of +20.6%, p=.000, Hedges' g = 0.47. This would seem to be very significant. The survey was not designed to capture this data, and other related questions demonstrated little notable change. But it would seem that E3 and to a greater degree E2, potentially recorded a significant change in attitudes that may warrant further investigation.

Another possible explanation for the significant rise in activity to protect personal information is the introduction of GDPR 'General Data Protection Regulation'. This EU directive came into force during the interlude in data gathering and although there is no direct evidence, the prospect of regulation changes influencing the population may warrant further investigation as it might illustrate the impact of government policy on public awareness and individual behavior.

6.3 Expert Interviews and Thematic Analysis Discussion

The semi-structured interviews and subsequent thematic analysis, generated findings aligning broadly into 3 categories, arguments surrounding adoption, interaction and general concepts. The following section endeavours to distil these themes down to the pertinent arguments, discussing topics most relevant to the questions posed within this research.

6.3.1 Marketing Privacy Is Not Enough

A dominant theme throughout the expert interviews, and indeed a seminal pillar of this research, is the value of decentralised technology and how this is embedded within artefacts and communicated to participants. The communication and understanding of value is critical to the preliminary stages of the Diffusion of Innovation (Rogers, 2003). The contemporary value proposition for the adoption of decentralised technology is privacy. Philip Sheldrake et al (see: 5.4.1.2) have clearly stated a position that the Decentralised Internet cannot be marketed solely on the fact that it is decentralised. It can then be argued that individuals don't perceive the value or context of privacy, and subsequently don't see the advantages of switching to technology that offers little more, or indeed less, functionality than their centralised counter parts. The literature describes privacy as a complex and misunderstood concept. It is clearly difficult for individuals or indeed academics to define and contextualise privacy as an overarching concept, and this is repeatedly argued in the literature (Thomas, 1975; Post, 2001; Solove, 2008). Jarvis (2011) describes concerns regarding privacy on the Internet, as a *'confused web of worries, ill-conceived, and unjustified'* (p. 9). Danial Solove argues that privacy is an umbrella term for intrusions in a myriad of contexts across a spectrum of cultures and social norms (Solove, 2008). Solove suggests a bottom up approach based on a taxonomy of privacy harms, through the notion of family resemblance, in order to clearly define and understand privacy concerns within the digital domain (see: 2.2.10). It would appear that this theory offers a starting position from which to consider the specific domain of network data

privacy, through which one might identify privacy infringements, emergent advantages, and the potential benefits and innovations of a decentralised model. There are a number of other factors compounding the participants perception of privacy harms in the context of a Decentralised Internet. As Colin Willis comments: *'In the West, we have just enough privacy'* (C Willis 2018, Personal Communication, 11th Oct), meaning direct individual privacy infringement is measured, and hasn't yet reached a point of comprehendible harm. There are arguments concerning changing social norms. Campbell and Carlson (2010) suggest an acceptance and apathy towards privacy issues, and Cohen (2012) has argued that the concept of privacy is becoming old fashioned. Zuboff (2015) argues that an acceptance of Surveillance Capitalism is now seen as necessary in order to achieve an effective life. Ian Brown (2017) argues that Immediate Gratification Bias and the Privacy Paradox, are demonstrations of individual actions and cognitive biases that led to *'non-optimal privacy decisions by individuals'* (p. 13). The evolving landscape is arguably perpetuated and indeed orchestrated by those holding power, O'Hara's (2014) rebuttal of *Zuckerbollocks* shines light on the power of influence, as arguments are made for the justification and disruption of social norms relating to privacy. This research suggests that privacy is a vague concept, and that a systematic investigation following the theories defined by Danial Solove (2008) should be engaged to better understand the domain, the potential value of a decentralised model and its relationship to privacy, to ultimately develop a clear, comprehendible understanding of problem and solution for both participants and developer.

6.3.2 Privacy, A Primary or Secondary Concern?

Throughout the expert interviews, there is a sense that the dominant concept of privacy, as a justification for engaging with decentralised technologies, may be masking other potential value propositions and positive narratives. And that the argument of protecting one's privacy, might become a secondary concern or positive consequence of decentralisation. If Danial

Solove's (2008) position is to be considered and privacy is seen as an umbrella term instead of a definitive catchall definition, arguments might be built through the taxonomy of privacy (see: 2.2.10), to communicate specific privacy problems and the solutions offered by decentralisation. At the same time recognising the benefits offered through decentralised innovations. It can be argued that this is not an issue of whether privacy is relevant or not, rather this is an issue of semantics in the communication of value proposition. In some situations, the narrative will be focused around privacy protection, but in others the narrative will be framed around innovation and advantages, even though as a by-product, engagement may also bolster privacy protection.

6.3.3 Building A Message

When considering the communication of value within the decentralised domain, it can be argued that this falls into two categories, arguments against privacy infringement, and arguments defining the advantages and potential innovations decentralisation supports.

Interviews suggested a need for a consistent narrative, to communicate the justification of decentralisation. A significate theme was that of *Control*. David Irvin, founder of Maidsafe, argues that people don't understand or indeed care little for the notion of privacy, but that when people realise they are being *controlled*, it is something very different. (Irvine 2016, Personal Communication, 9th Nov). The literature supports this idea. The concept of the Panopticon, (Bentham, 1791; Himmelfard, 1968), the concept of control being metered in the mind (Foucault, 1975), and the notion of Social and Panoptic Sort, (Lyon, 1993; Gandy, 1996). These arguments of *Control*, and indeed exploitation, are drawn into the digital realm, and to the depths of Marx's theory, through the *Prosumer Proletariat*, with notions of class, exploitation and surplus value. (Fuchs 2012). The notion of resisting being *Controlled* offers a clear means of expressing the rationale for adoption, which may potentially strike more resonance with the average participant then the notion of privacy.

An additional powerful message is that of failing to benefit from the innovations and opportunities decentralisation potentially offers. This is supported in the literature. Hall argues *'how vital the sharing of personal data is in technological, and specifically, digital innovation'* (Hall, 2016, p. 03). Van Kleek argues that we are jeopardising the realisation of Web 3.0 technologies (Van Kleek, 2014). Pentland highlights the potential, positive societal impacts, if we can move from data based on beliefs, to data based on behaviours (Pentland, 2012). It can be argued, that the decentralised community should be looking positively forward to the innovation's decentralisation offers, to identify the emergent value through which to build a positive narrative. Indeed, great frustration is voiced that the *'Decentralised Brigade'*, have to a degree highjacked the argument, focusing primarily on a vague battle for privacy with the objective of reversing the status-quo. (D, Alexander 2019, Personal Communication, 2nd Dec 2016). In summary this research suggests 2 core strands for a decentralised communication strategy, the notion of being *Controlled*, and the significant benefits and missed opportunities of decentralisation.

6.3.4 Finding Value in Decentralisation

Throughout the expert interviews, there has been significant debate, regarding what decentralised innovation may offer. The themes generated from these conversations are valuable, as they act as an inspirational catalyst for innovation, in addition they form compelling narratives through which value can be established to promote adoption. The themes are broadly divided into 3 areas: the individual, commerce and society.

6.3.4.1 For the Individual

It is argued that decentralised models, which provide agency through reusable and verifiable personal data, offer considerable advantages. A prominent theme is that of streamlining and acceleration of daily transactions, reducing friction, and making it easier to complete tasks. Gaining control over federated identity currently controlled by third parties, is another notable example. The Identity that you invest in, that is developed and refined over time has great value and should belong to its subject and not indefinitely held by a third party. The power of federation, or redistribution of personal information, on the user's terms, is a powerful mechanic of decentralisation.

The concept of empowerment is a compelling idea. The participant controlling their digital presence, using the validation of identity, verifiable credential and mechanisms of negotiation and contract, form a powerful message that a Decentralised Internet delivers the same agency in the digital realm, as that experienced in the real world. This empowerment manifests from the capability to communicate with anonymity, through to the means to avoid echo-chamber and political manipulation, the concept of a Sovereign Boundary Mechanism, and the metaphorical ring of steel between the participant and the network. Collectively these ideas can be woven into persuasive metaphors and value statements.

A significant digestible example of empowerment is Vendor Relationship Management (ProjectVRM, 2019). The principles of VRM are predicated on the rebalancing of the current asymmetric relationships between participant and vendor, freeing the participant from contracts of adhesion across a spectrum of transactions (see: 2.4.8). This is a powerful narrative, re-decentralising through a peer-to-peer model goes beyond privacy protection, and arguably presents an array of opportunities for individuals to transact independently within a rebalanced landscape.

The cost savings for a free agent on the network is another notion that might build a persuasive message. During an interview with Adrian Gropper MD, CTO at Patient Privacy Rights, the comment was made: *'individuals simply don't understand just how much surveillance capitalism is costing them'* (A Gropper 2016, Personal Communication, 23rd Nov 2016). If this could be quantified, in real terms, it would constitute an immediate understandable value proposition.

To surmise, the notion of streamlining, the ownership of identity and the power of federation, the prospect of empowerment and the rebalancing of relationships with vendors, offer a collection of themes around which to build individually focused value proposition. If this is wrapped in the narrative of emancipation from a controlling and manipulative dominate force, it provides a powerful argument, more so than the vague prospect of privacy protection alone.

6.3.4.2 Societal Gain

Societal gain, as an understandable justification for adoption, is a central narrative that was discussed at great length during expert interviews and during focus groups. The importance of privacy for the well-being of society is well documented in the literature (US-Gov. 1973; Gavison, 1984; Solove, 2008; Ohara, 2016). Our ability to protect the vulnerable, improve health and social care, education and the efficiency of public services are all components of a functional society, that will benefit from open sharing of personal data. Silverman expresses concerns about our trajectory of travel and our lack of understanding regarding the social benefits of privacy (Silverman, 2017). At a macro level, the argument that we need to safeguard our democracy (Grassegger, 2016; O'Neil, 2016; Monbiot, 2017), and build a healthier society and support adolescent development by maintaining the ephemeral (Schneier, 2015), offer a further dimension for the benefit to society argument. Indeed, the concept of societal gains aligns with the theories of Danial Solove (see: 2.2.1): that any granting of privacy rights should

be afforded if it benefits society. The results of the Public Survey have illustrated the favoured motivation for the sharing of personal data as societal gain (see: 5.3.2). It can be argued that the rewards for a functional, open, decentralised mechanism are clear for society. And a narrative should be framed, in terms of these missed opportunities society faces, by being locked into a centralised model.

6.3.4.3 For Business

Positive sentiment was held across the majority of experts consulted with regards the potential benefits to commerce decentralisation offers. A functional Human-Centred Data Ecosystem is considered to offer significate opportunities for new business models and efficiencies. Haddadi (2015) argues that the locking in of network participants is *'preventing the formation of a truly competitive market'* (p. 1). Levine expresses a view that the Internet could provide an environment which resembles the vitality of an ancient bazar (Levine, et al, 1999). Searl's (2012) argues that the internet makes *'obsolete, the Industrial Revolution business models of mass marketing, and mass media'* (p. 159). In a relatively short period of time, the Internet has gone from an open marketplace of thousands of individual businesses, to businesses that are forced to engage with, and or go through one of four major players. There would seem to be a great appetite to break these monopolies, and release commerce from being forced to operate through controlled mechanisms. It is argued that this provides opportunities for established larger organisations, but more importantly, acts as a leveller for smaller operations and entrepreneurial endeavour. Indeed, many of the potential models for innovative business through decentralisation have previously been conceptualised and developed, to a degree through the principles of VRM (Vendor Relationship Management) (see: 5.3.4.5). With the advent of a functional identity layer, many of these concepts would now seem to be within grasp. During interviews, a number of specific ways decentralisation might offer value to

commerce were voiced, these include: the removal of back room costs, reduction in friction, off-loading the responsibility of data holding, the prospect of real-time high-quality data marketing intelligence, and the competitive advantage of direct trusted relationships with customers. As well as clear advantage for business, the related notion of emancipation from the current centralised model, and the cost savings, offers a valuable marketing message for both vendor and consumer.

6.3.5 The Cultural Context and Niche Pockets of Value

This discussion falls into two strands, the cultural context of decentralisation and the recognition of niche pockets of value. The cultural context is important, and in any effort to design, build and disseminate decentralised technology, the consideration of the cultural dimension and its relevance to any overarching strategy is critical. The notion of strategy in this context, relates to designing decentralised tools and services, that are aligned with the requirements and worldview of a recognised culture. It can be argued that identifying a cultural niche, may offer an opportunity to realise adoption. If the overall community objective is to achieve a critical mass for a global ecosystem, identifying genuine cultural need, with lower barriers to entry, and targeting these domains first, raises the probability of realising a sustainable ecosystem. This notion aligns itself with Moore's Technology Lifecycle Theory (see: 2.4.5) where in order to gain adoption, identification of niche markets is required.

During the expert interviews, the argument was made that in a western liberal democracy, we currently enjoy *just enough privacy*, and care little enough to see the value in decentralised services. (C Willis 2018, Personal Communication, 11th Oct), This is supported by the theories of the Privacy Paradox (Norberg, 2007), and Instant Gratification Bias (Acquisti, 2004). But equally other arguments are made, with German society identified as a group that values privacy highly in a family context. Points have been made regarding

community groups that sit outside the mainstream, countries that don't enjoy the same levels of democracy and freedoms, peoples who are without recognised identity and documentation, the unbanked, refugees and asylum seekers, or those that simply don't subscribe to the established social norms. It can be argued that there is a great deal of work to do in identifying cultural groups, that might benefit from a Decentralised Internet outside of the western vein. When considering the varied cultural contexts, a signal standardised ecosystem maybe suitable, but the developed services and applications, and the targeting for adoption is varied. Simply put, one size does not fit all.

6.3.6 Unforeseen Barriers of Decentralisation

Pertinent insights emerged through the theme of *Barriers to Adoption* and suggested a number of issues that could be argued to be unforeseen consequences of decentralisation. These issues centred around conceptual barriers, which may emerge once interaction with the network becomes enabled through a Sovereign Boundary Mechanism.

The issue was raised of *Decentralisation Working Both Ways,* meaning once access to extensive personal data becomes normalised, third parties may begin to demand more of it, in order to provide transaction and services. There is a sense that the concept of decentralisation could rebound, leaving individuals increasingly exposed. Debate did not reveal specifics, but this is an interesting angle which requires further study.

Differing user groups who do not understand the technological concepts or struggle with the mental models, may find themselves excluded from the benefits of decentralisation. This topic was heavily debated during the focus group and was a theme which required serious further consideration. In parallel debate, the concept of responsibility was raised. The issue that taking control over personal data through a Sovereign Boundary Mechanism, defining relationships, making judgments of trust, the monitoring of dynamic transactions, and being ultimately responsible for backup and fail safe, represent a significant on-going responsibility

and potential isolation. This was considered to pose considerable friction and potential anxiety. The risk that the participant may lack trust in their own capabilities and competence represents a potential adoption obstacle.

It is important to consider that outside of the primary focus around value proposition and functionality at the interface layer, there are many nuanced variables across differing user groups which need to be further investigated and fully understood.

6.3.7 The Trust Framework

A central component of a Human-Centred Data Ecosystem is a Trust Framework, indeed a driving organisation behind decentralisation is known as *Rebooting the Web of Trust* (RWOT, 2019). There has to be some solid ground so that peers can trust one another over the network. At present trust is facilitated across a string of usernames and passwords, issued through various degrees of verification, centralised organisations federating loaned identifiers, and a pyramid of certificate providers. These centralised mechanisms, combined with secure payment services offering a degree of insurance, establish an acceptable level of trust that allows interaction and transaction. If the Internet is to move to a decentralised model, the evolution and mechanisms of trust need to be considered carefully, to establish what is an acceptable and functional level of anchorage across differing kinds of transaction. The distributed ledger is one part of the equation, providing a means to prove control over encryption keys and identifiers: It is a way of verifying credentials through digital signatures and establishing agreements through smart contracts. But where is the anchor? How does one verify a credential, an identity or a reputation? One answer is to seed identity from state or corporate sources. A personal credential issued by a commonly known root identifier, for example the driving licence association or a passport issuer. Identity may be seeded by corporation or financial institution, such as a public service provider or bank. It may be that biometrics come into play, for example physical identity shops, an early exemplifier of which

is *Archive* highlighted within the Artefact Catalogue (Archive, 2016). How does a centralised anchor relate to a decentralised objective? Is this still a centralised model? If the central anchors on which the verification of an identity is built can be retracted without notice, this contradicts the principles of *Existence* and *Persistence* defined by Christopher Allen (Allen, 2016). An identifier can be persistently controlled by the participant, but the potential verification of that identity is ultimately reliant on a third party. Are there other methods of building trust? Perhaps in the same way as centralised identities are developed overtime, through content, ranking and reputation? Are there existing models for this elsewhere? And is trust even needed when smart contracts can lock in agreement through the notion of *'Code as Law'*? Many of these questions are yet to be resolved or explored, and there would seem to be a rich stream of research materialising within this domain.

6.3.8 Looking Past the Technology, Turning to Design

Throughout this research, supported by conversation during expert interviews, there is a sense that the objective of a Decentralised Internet has now moved out of the realm of the purely technical, towards the domain of design. Investigation has concluded that the majority of the technical stack layers are now available, and the mechanisms for interaction with a full personal data eco-system are evolving rapidly. It can be argued that the balance of development has now moved into the realm of design. The development of value propositions, digital services, interaction, and underlying narrative, are all elements that can be considered, and resolved through design thinking. The problem space can be considered systematically, and processes can be engaged to develop solutions. It is telling that at the time of writing, December 2019, if we consider the strands published for the MyData conference, there is a great deal of opportunity to hear speakers discuss technology, computer science, ethics, law, and commerce. But there is a clear lack of a dedicated design strand, exploring and identifying the fundamental questions that need to be resolved. Indeed, a contribution to knowledge within this research, is

a body of work that will help the design community to better understand the decentralised domain, the opportunities it presents, and the variables and constraints within which new products and services could be developed.

6.3.9 Getting to The Interface Layer

A powerful argument that warrants further discussion is that of *Getting to The Interface Layer*. Any attempts to decentralise the Internet face the issue of access to the literal screen space, that many of the dominant forces have monopolised to a greater or lesser degree. The barriers to overcome are significant. Apple devices and operating systems are closed and controlled, Android is in essence open source, but the influence of Google is significant. Most web portals are under the control of the dominant Internet forces, and the power or search and targeted marketing may favour centralised offerings. With the normalisation of network activity moving to smart handheld devices, accessing this interface layer in a sustainable way, needs to be considered in any strategic planning by decentralised advocates. Indeed, anecdotally, a detailed conversation was had during MyData 2019 with a senior designer at a globally recognised telecoms provider, who claimed: *'without access to the hardware and the interface layer, without a fundamental change to the interaction model within mobile devices, the prospect of Decentralisation is limited'.*

6.3.10 Community Agendas

The conducted interviews, together with conference attendance and the reading of the literature, reinforces the inevitable camps of political perspective, and motivation within the decentralised community. It is interesting to observe these differing, and potentially problematic positions, as attempts are made to define manifesto and realise collective cooperation.

For many, the resistance to the dominant Internet forces is almost militant in nature, arguably driven by a negative world view towards capitalism, or an anti-disestablishment and incredulous position towards the state and surveillance. This is contrasted by individuals and organisations, who see the commercial opportunities of decentralisation, and are focused on capitalising from models of limited sovereignty with a semi open ecosystem. There are other groups who see the missed opportunities of Big Data and the social advantages a data driven society has to offer. And there are those with a passion for technology, who are motivated through the building of new innovations, standards and infrastructures. The following examples illustrate a selection of these positions.

The *MyData* organisation defines its objective: *To empower individuals with their personal data, thus helping them and their communities develop knowledge, make informed decisions, and interact more consciously and efficiently with each other as well as with organisations*. (MyData, 2019). The MyData position is reasonably neutral, but might be argued to be more activist led, with a focus towards social responsibility. In contrast *BlockStack*, is a company that is clearly focused on a market share. It aims to be first to the table with a semi open ecosystem, offering Identity, Distributed Storage, and a DAPP 'Decentralised Application' marketplace (BlockStack.org, 2018). *Sovrin* and its associated company *Evernym*, would seem to be focusing on the bigger picture, publicly building infrastructure, while at the same time developing peripheral business models through commercial tools and agent and wallet software that participants will later require (Sovrin, 2017; Evernym, 2018). Finally, projects like *Veres One (2018)* and *Uport (*2018), would seem to be purely technology and developer focused, with little evidence yet of practical application.

It can be argued that the realisation of a sustainable Human-Centred Data Ecosystem, is unlikely to be achieved by one organisation or individual, and will instead require coordination, and collective effort. But this may prove challenging in a community of tribes

with conflicting agendas. This research does not take a position on this issue, nor does it offer a solution. This is an observation that one may need to be mindful of, when considering overall strategy, and offers an interesting landscape for further research.

6.3.11 The Need for A Cohesive Strategy

Following on from the discussion concerning community agendas, the need for a cohesive strategy would seem to be evident. There are a great many stakeholders who believe in the benefits of a Decentralised Internet. The first wave of concepts, applications and the technology infrastructure are beginning to materialise, many are driving to be first to market with solutions through semi decentralised architectures. Others are attempting to develop a full ecosystem, which once established, provides a foundation for commercial opportunities. In trying to develop something which is arguably a paradigm shift against a powerful monopoly, it could be argued that a cohesive decentralised community strategy is required. To rely on individual break through, or a serendipitous moment is not enough. A cohesive strategy, standardised methods, seeded trust frameworks, targeted opportunities and establishing consistent narrative, are all examples of how collective cognition will increase the probability of achieving a sustainable ecosystem.

6.4 Practice Led Component Discussion

The development of a conceptual model, primarily focused on the extension of existing efforts to develop SSI applications, considered the principles of SSI and HDI, and the developing technologies and standards. The objective of this component of research, has not been to develop a finalised artefact, instead the practice explores a preliminary practical implementation of a system at the interface layer. Developing the conceptual model and preliminary user interface and verifying and testing through critical analysis and expert focus group, has led to the following sections of discussion, highlighting arguments that are considered to be most relevant to the questions posed within this research.

6.4.1 The Sovereign Boundary Mechanism

The development of a conceptual model in-line with the principles of the decentralised domain, following the trajectory of technologies and standards, realises a system that allows participants to manage their data, information, communications and affairs independently through a digital wallet and agent. The concept of sovereignty in this context translates to individualism, and this in turn, given the complexity of the required interaction, poses several issues. Within this research, this independent domain of interaction, has been titled A Sovereign Boundary Mechanism. This means there is a clear boundary between the management of identity, relationships, credential data, the engagement with applications, and the rest of the network. This is a strict task-based interaction; one which incorporates new distinct ideas and concepts (see: 5.5.4). Given the gravitas and value of the personal data being transacted, this research suggests that a strong internalised understanding of the domain will initially be required to engage users. This requirement for internalised knowledge relates to the traditional notion of internalised cognition and mental model. (Craik, 1967; Norman, 1986; Payne, 2003).

Creating a situation where a participant is required to engage in significant internalised cognitive processes is counterintuitive to the evolution of HCI theory, and accepted design thinking, where externalised and distributed cognition and offloading are considered best practice (Scaife & Rogers, 1996, Hitchens, 2000, Perry, 2003).

The required degree of internalised cognition for individuals to initially engage and have confidence and understanding of the system for a Diffusion of Innovation (Rogers, 2012), is arguably a primary consideration for the decentralised community moving forward. This research has clearly highlighted the complexity of a Sovereign Boundary Mechanism, but further research and design practice needs to be undertaken to explore how complexity and friction can be significantly reduced.

6.4.2 The Paradox of a Sovereign Boundary Mechanism

The objective of decentralisation is to emancipate the participant from the centralised Internet. In doing so the negative consequences are mitigated, and the missed opportunities presented by a decentralised alternative can be realised. However, this research suggests, that the current trajectory may inadvertently replace one set of constraints with another. The complex landscape, the isolation, internalised cognitive load, the responsibility of managing one's own data, and the generation of friction not found in centralised counterparts, may replace one form of incarceration with another. It can be argued that these issues can be addressed and overcome, but the notion of the proverbial, *out of the frying pan and into the fire*, needs to be considered as future tools are conceived and developed.

The notion of a genuine Decentralised Internet is predicated on the principle that the owner of the data should have control over it. The individual should have command over multiple immutable persistent identifiers, and they should have agency to decide who to share a relationship with, and who on the network can observe their activities and transactions. The participant should have the capability to manage and redistribute their credentials or data to

whom they see fit, under their own terms and conditions (Mortier, 2013, Allen, 2016). This research has demonstrated what has been termed as a Sovereign Boundary Mechanism, representing a participant who sits within a metaphorical technological boundary, defining Identity, controlling relationships and managing data streams. This is achievable, through an analogue model that potentially liberates the participant from the centralised Internet. Paradoxically this model comes with a selection of caveats which pose significant problems. This research has raised the issues of complexity, internalised cognition, mental model, friction, risk, responsibility, trust, and exclusion. By investigating these issues, this research presents key barriers which need to be considered and addressed in order to achieve sustainable adoption.

6.4.3 Back Peddling on Friction

A topic highlighted within the expert interviews and exposed through the Practice Led component, is that of increased interactive friction. It can be argued that the SSI analogue model extended within this research, exhibits a higher level of friction across interactions than that found within centralised counterparts. This relates to the required cognitive engagement, the understanding of original concepts, vague mental model and metaphor, and the shouldering of more personal responsibility. This research doesn't offer a metric on this assumption, but through developing and analysing a Sovereign Boundary Mechanism, the array of conceptual components for interaction and the collective required objects and subsequent multi-layered user interface, it can be argued to be substantial (see: 5.5.3). If decentralised technologies are to find adoption, friction needs to be reduced not increased. Placing a number on this increased friction, potentially found in differing forms of decentralised interactions, is outside the scope of this research. However, this is a topic that certainly warrants further investigation, and should be prioritised within any continued endeavour.

6.4.4 A Missing Mental Model

When considering the underlaying assumptions driving this research, in the early chapters of this thesis (see: 1.3) the paper, *Why Johnny Can't Encrypt* is cited (Whitten & Tiger, 1999). The assumption is stated, that similar circumstances described within this work, may emerge when considering decentralised endeavours. Within the paper, it is argued that different methods of user evaluation are required when considering software. Where a user lacks the understanding of the underlaying concepts and mechanisms. In essence, the user lacks the mental model of the domain (Carroll & Olsson, 1988). The paper is concerned with sending emails with encryption software. The paper recognises, that even with a well-designed interface, users struggle to complete what is a relatively simple task. A Sovereign Boundary Mechanism represents a system which is arguably considerably more challenging. This research has demonstrated a user domain that is significantly complex. There are multiple concepts, processes and interactions, which when taken individually, potentially lack the mental model for meaningful engagement. However, when these elements are combined as a whole, it can be argued that without considerable guidance, the objective of sovereign agency and utilisation of personal data through a mechanism of this kind, is impractical. It is accepted that the developed user interface within the Conceptual Model is preliminary, and that further cycles of refinement will reduce complexity and potentially improve the mapping of interactions. However, the UI design can only go so far, and it has to be recognised that a full Sovereign Boundary Mechanism, in this guise, presents considerable barriers with respect to forming an operational mental model for the participant.

6.4.5 Internalised Cognition

Leading on from the discussion regarding the complexity and potential missing mental models, even with sufficient understanding of the system domain, the degree to which a participant may rely on internalised cognition in order to engage is of concern.

This research identifies that the required internalised understanding for both the interaction and many of the broader concepts is significant. It can be argued that the cognitive load for initial engagement poses an issue for adoption. The very concept of sovereignty, and the metaphor of a secure boundary, suggests a degree of user isolation and internalisation. The value and differing types of the personal data, the weight of being solely responsible, the complex processes through which data must traverse to engage in meaningful transactions, amounts to a sizeable load of internalised understanding and knowledge. If a mental model for this domain can be established, it can be argued that engagement will still require a considerable degree of internalised cognition.

Within the expert interviews, comparison was made with *Crypto Currencies*, and the difficulty individuals have in understanding and engaging with this ecosystem, when asset value and responsibility falls into the equation. It can be argued that a Sovereign Boundary Mechanism is considerably more complex. Within the product design process, the degree of internalised cognition needs to be accepted and measured, from there every effort needs to be made to reduce the internalised understanding and decision making required. As discussed in the *Back Pedalling on Friction* section (see: 4.4.14). The increased friction posed by decentralised systems is a primary issue, and it is the internalised processes that are arguably responsible.

6.4.6 A Spectrum of HCI Theory

It can be claimed, that a User-Centred Data Eco System through a Sovereign Boundary Mechanism requires the consideration of a spectrum of HCI theory and paradigms (Rogers, 2012), and that any further innovation needs to recognise this in its deliberation. The graphic below (Fig. 28) illustrates the applicable theory across a model of a User-Centred Data Ecosystem. Internalised Cognition and domain specific mental model (Payne, 2003) at the core of a Sovereign Boundary Mechanism. Externalised Cognition (Scaife and Rogers, 1996) at the interplay across representations between the core and the boundary. And the notion of Distributed Cognition (Hitchens, 2000), (Perry, 2003) and Situated Action, as transaction occurs outside of the user boundary across the wider network (Suchman, 1987). In addition, Interactions and decision making within the wider ecosystem will be reliant on community-based templates, and the development of trust networks, so societal and cooperative theories of HCI are also relevant (Schmidt, 2011). What is clear is that SSI falls into a number of HCI paradigms, and in identifying them, the application of Harrisons Phenomenological Matrix, where theories are considered and selected where deemed appropriate, would seem to be of relevance (Harrison, 2007).

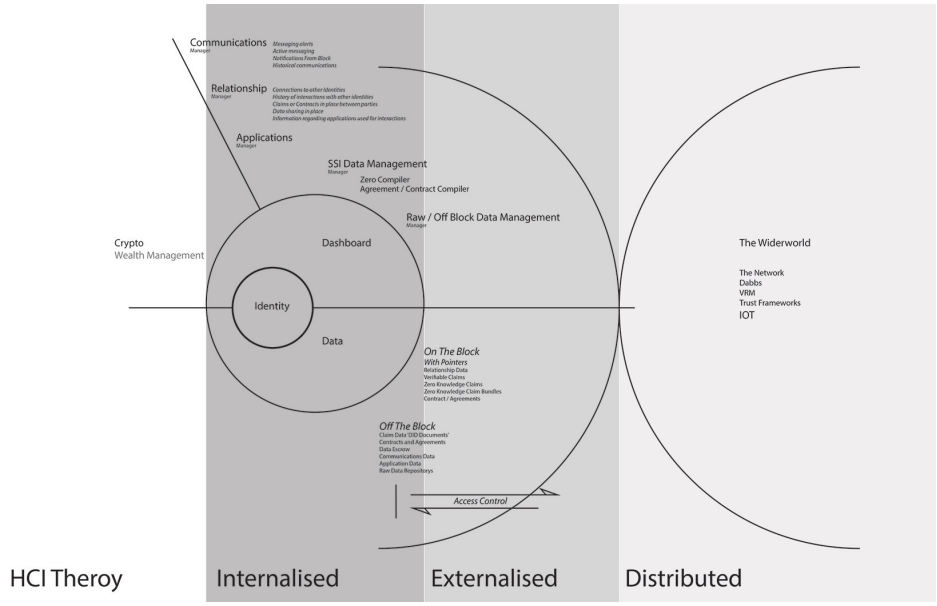


Figure 28: HCI Theory across a User-Centered Data Eco System

This research proposes that the decentralised community needs to understand the notion of cognition and systematically address both issues of direct usability and the interaction and internalised understanding of concepts. The community needs to comprehend the view held by Paul Wilshaw, Head of Digital Innovation, Barclays UK, when reflecting on the SSI conceptual model and UI developed within this research: *'They simply won't get it'* (P Wilshaw, 2019, Focus Group, 15th March). This research argues that Payne's theory (Payne, 2003) of specific mental models for domains should be followed to map the required user understanding in detail. Any direct interaction should consider how the interface layer can push as much cognition as possible into the externalised domain. Finally, interaction and transactions need to be fully understood outside the Sovereign Boundary Mechanism, so that Distributed cognitive relationships and Situated Actions can be defined. Above all, the consistency and cooperation across all stake holders is considered critical.

Types of activity recognised through this research, which may be reliant on Distributed Cognition and the consideration of Situated Action are:

- Validation / Reputation / Trust of Individual Actors
- Collective Decision Making
- Collective Production Activity
- Collective Data Sharing
- The Construction of Larger Cognitive Artefacts and Systems

A final notable HCI theory and potential movement, is that of Human Values, meaning any digital technology developed, should strive to consider human values within its development process and ultimate application (Harper et al, 2008). The theory introduces the notion of empathy into the design process, as a point of reflection to consider the human values within any technology development. This topic offers an interesting area for further research and poses some interesting questions: What are the tensions around the Decentralised Internet

being truly decentralised? How will this domain develop once established? How might the established forces react? Will we simply go full circle, in a *'Google, Do No Evil'* repetition of history?

6.4.7 Can AI or Collective Intelligence Reduce the Cognitive Load?

This area of discussion spans both the expert interviews and derives from observations within the conceptual model. Much conversation centres around the intelligent agent or personalised artificial intelligence, taking control of much of the mundane decision making and ongoing administration involved in a decentralised system. This might be based on broad sweep criteria defined by a participant or derived through machine learning based on the participant's history. This concept has the potential to drastically reduce the cognitive load required, drawing the participant's attention only to critical and important decisions. The concept of AI supporting decentralised engagement is attractive but can also be seen as a misguided panacea. There are issues of trust, understanding and value, which need to be considered as trade-offs against automation. If an objective of decentralisation is transparency, to hide critical decision making within a black box, may be contradictory. How can the participant trust the AI? Who defines the AI? And more importantly, how much of the inner workings of the system does the user need to see and understand, in order to both have faith, and see value in the engagement? This is a significant topic for further research and debate.

Another solution to cognitive load reduction and decision making, may be the establishing of group or tribe. A trust network that collectively makes decisions through shared values, for instance an environmental collective that marshals relationships and transactions through ethical reputations. This is part of the Trust Framework conversation, related to the theories of Distributed Cognition and Situated Action, and again offers a rich seem for future research.

An aside to the notion of collective decision making is that of democracy, political representation and vote casting. An interesting discussion might be found around the concept of decentralised systems acting as a voting mechanism. The logic being if knowledge is power, then sharing your data and subsequent inferred collective information, may offer a new and dynamic means of democratic process.

6.4.8 Balancing the Cognitive Load against the Value Proposition.

The practice-based component of this research has demonstrated that there is considerable cognitive load, complexity and participant responsibility that potentially manifests within an active personal data ecosystem. This friction might be mitigated through careful design considerations, but it can still be argued that even then, the required engagement demands more effort on behalf of the participant than that currently found within existing centralised services. To this end, in line with many of the arguments around the communication of privacy, control, missed opportunity and value proposition within the thematic outcomes, any development of interface layer needs to be balanced against the value that interaction serves to the user. Where this balance lies, and how it manifests, is a central conundrum in the delivery of a sustainable Decentralised Internet and stands as a source of considerable further research.

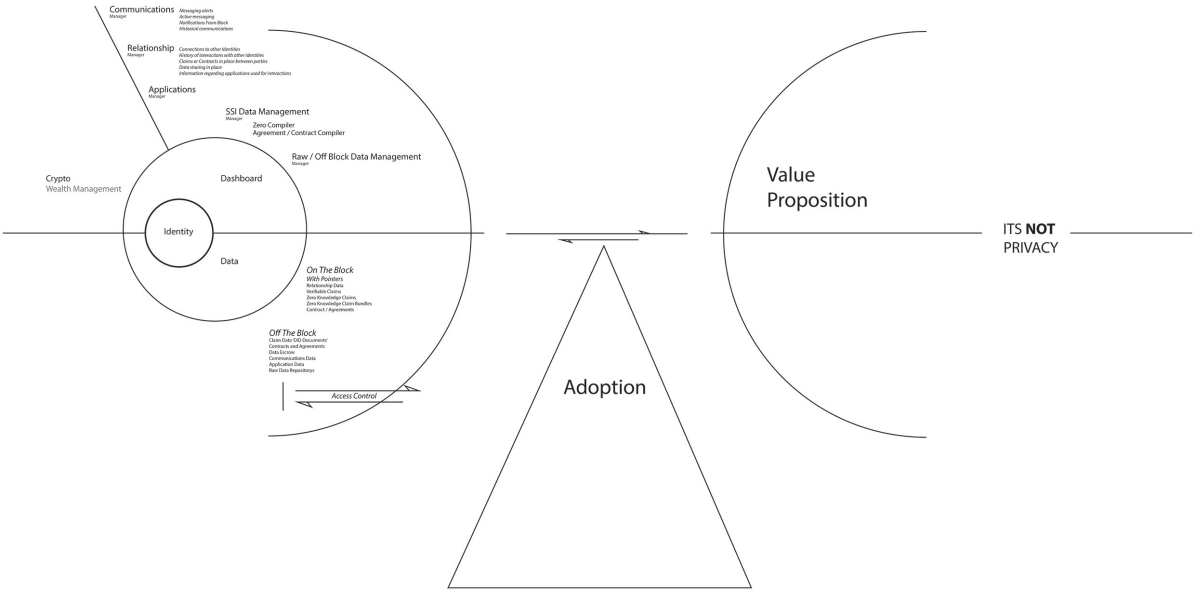


Figure 29: Balancing the Cognitive Load against the Value Proposition

6.5 Discussion Summary

Within this chapter, the results of this research have been interpreted and discussed in line with the relative theory and research questions.

The development of the Artefact Catalogue has been explored, considering the pertinent advent of a functional identity layer. This development of a sovereign means of immutable identity would seem to be a significant step forward in enabling full agency decentralised mechanisms and interactions. The trajectory of the decentralised domain has been discussed, highlighting the trend towards SSI and ultimately a Sovereign Boundary Mechanism. Finally, this section considered the identified and potential value in decentralisation, arguing that a full-agency engagement is required to realise the true value of decentralisation. Value in the artefacts catalogued was discussed, recognising the value and resource potentially found in now defunct VRM artefacts.

The chapter continued to consider the Public Survey, highlighting and discussing the descriptive statistics of significance, many of which offer interesting insights in terms of value

and targeting for adoption. The section considers the scales of significance which provide valuable information regarding perceptions, attitudes and understanding. Correlation analysis was discussed, communicating the weak evidence of relationships between scales. The comparison statistics are considered, with some significant results highlighted concerning *Gender* and the *Consideration* and *Motivation* for the *Control of Personal Data*. Finally, attention is drawn to the data developed through the unintentional Longitudinal study, which highlights significant changes in *Consideration* and *Efforts* made to protect personal data, following the Cambridge Analytica revelations.

The chapter continues to consider the core results of the Expert Interviews and Thematic Analysis, which centres around the 3 broad categories of Adoption, Interaction and General Concepts. The discussion considered the theme of privacy, citing literature to support the argument that it is complex and misunderstood, and the need to identify privacy harms from the bottom up. The section draws on opinions that privacy alone is insufficient as a justification for adoption. The notion that privacy is a secondary consequence of decentralisation is voiced, leading to the arguments for a consistent narrative for a decentralised marketing strategy. The section highlights 3 specific categories of value, including: personal, commerce, and society. The cultural context is considered before moving to explore some of the recognised unforeseen barriers of decentralisation. The concept of a trust framework is discussed as a necessary component of a Human-Centred Data Ecosystem. Arguments are then made that the decentralised community now need to look at the issues from a *Design Thinking* prospective, as opposed to pure technological and computer science. The problem of accessing the Interface Layer is discussed, before moving on to consider the varied community agendas and a need for collective cohesion.

In the final part of the chapter, the Practice Led Component is discussed initially reflecting on the extension of SSI through a conceptual modelling process and preliminary UI,

eventually arriving at a Sovereign Boundary Mechanism. The complexity of an SBM and general engagement with a data ecosystem is considered, and literature regarding HCI is discussed. The issue of internalised cognition is raised highlighting the problem that an SBM may result in a system which paradoxically replaces one form of incarceration with another. The section moves on to discuss the need to reduce friction in decentralised interactions and highlights the concerns raised through critical reflection and focus group of the potential for missing mental models. The section returns to the HCI arguments with the position that a spectrum of theory needs to be considered across Internal, External, Distributed and Situated Action when considering and designing for the whole participant journey. The section continues to discuss the recognised interactions within the distributed realm, before moving to the prospect of AI as a means to off load mundane interaction and administration, arguing that this should not be seen as a panacea. Finally, the section discusses that ultimately any engagement with an interface layer, must be balanced against a value proposition if it is to achieve participant adoption.

This thesis now progresses to the conclusion chapter. The research questions are addressed, the main findings are listed and the contributions to knowledge are discussed. The directions for further research identified within this work are communicated, followed by an examination of the limitations of the research methods. Finally, the chapter concludes with a personal reflection of the research journey.

7.0 Conclusion

Within this chapter, the results and discussion sections are brought together to communicate the main findings within this research. The research questions are outlined and re-examined, before being addressed directly. The chapter proceeds to communicate the knowledge contributions from both a practical and theoretical perspective. The efforts to disseminate the research are reported, followed by a critique of the limitations of the research methodologies. Recommendations for further research are outlined, and finally the chapter concludes with a personal reflection on the research journey.

7.1 Answering the Research Questions

In order to answer the research questions, they are first re-examined and deconstructed, before each component is addressed in turn.

7.1.1 Main Question:

Can a Sustainable Technology be Established to Allow for Individual Agency within a Decentralised Internet?

The main research question can be dissected and interpreted further. Firstly, a definition is required to understand what is meant by a Decentralised Internet, followed by a description of the principles and required functionality to achieve individual agency. Within this research, it is argued that a Decentralised Internet can be defined through the principles of both Human Data Interaction (see: 2.4) and Self-Sovereign Identity (see: 3). A Decentralised Internet can also be defined to mean any technology that allows degrees of transparency, partial agency or disruptive capability. Within this research, the decision has been taken to focus on the potential for, and drive towards, complete agency over individual data. The development of an Artefact Catalogue (see: 5.3.4.2) has suggested that this is indeed the trajectory, and current objective of the domain. Within this research we extend this development, arriving at what has been

termed as a Sovereign Boundary Mechanism (see: 6.1.2). Individual agency is interpreted to mean, placing the participant at the centre of data flows, with full understanding and control over their data landscape, through a sovereign identity layer. Though a participant is sovereign in this equation, their interactions are dependent on a wider ecosystem. The term which best describes this domain is a Human-Centred Data Ecosystem.

Once the definition for individual agency within a Decentralised Internet has been established, we can answer the question concerning the availability of the technology, from a purely computer science prospective. This can be phrased as; Are the required technical concepts and components in existence? This research demonstrates that many constituent parts are available (see: 3.3). The technology layers from the ground up are present, with existing and developing standards around encryption, key management, distributed ledger, decentralised identifiers, verifiable credential, distributed storage and access control. The combination of these elements in establishing what has been termed as the missing Identity Layer, together with broader decentralised mechanisms, significantly bolsters the probability of a truly Decentralised Internet becoming a reality.

With a position regarding pure technology drawn, focus can move to the element of the question which asked if a sustainable technology can then be established. The notion of '*sustainable'* and '*established'* within the main question, relates to the requirements for longevity, initial adoption, that the technology has value, that the technology is accessible, and can build a self-perpetuating critical mass of participants. The words '*sustainable'* and '*established'* within the primary question, required the posing of two additional questions. These questions and subsequent answers are discussed within the following paragraphs of this chapter.

7.1.2 Additional Questions

The substantive component of this research comes when we consider the adjectives, *'sustainable'* and *'established'*, in the context of individual agency found in the primary question. To establish a Human-Centred Data Ecosystem requires an interpretation of its objectives and principles, combined with a technical architecture to achieve a tangible interaction. For a technology to be established, it needs to be designed, built and tested. It needs to be adopted and the surrounding components of the wider ecosystem need to be established. If it is to be utilised by the individual participant, it needs to be accessible and usable. If a technology is to become sustainable, it needs to build gravitas. It needs to reach a critical mass and continue to attract participation long after its inception. In order to investigate and answer these fundamental elements of the main question, it became clear that two additional questions are required. The following sections first define, interpret and endeavour to answer these additional research questions.

7.1.2.1 Additional Question 1:

Can an Interface Layer for a Decentralised Internet be Designed to Allow for Accessible Interaction?

In the context of this question, the notion of an interface layer relates to the means by which the participant accesses the underlying decentralised ecosystem. This may relate to interactive screens on a handheld or static computer, a tangible interface in the real world, or passive interaction with IOT devices. The adjective *accessible*, relates to an evaluation of the user's interaction with an underlaying system. It also relates to the complexity, friction and required mental models, found in a manifestation of a tangible decentralised artefact. As a benchmark within this research, the investigation has considered a Sovereign Boundary Mechanism and its subsequent realisation as a Self-Sovereign Identity application. A broader

justification for this decision can be found within the SSI section of this thesis (see: 3). In essence, this research has projected current trends to understand and evaluate the anticipated interaction model.

Through projecting current trends, this research has established the term Sovereign Boundary Mechanism. The term describes a part of a network interaction where a participant steps back into their own metaphorical boundary. Which can be described as a ring of sovereignty, in which the participant manages peer to peer relationships, the generation or collecting of credentials, the management and storage of personal data, and the distribution of data or credential under sovereign terms, condition and contract.

In terms of an interaction model, this research demonstrates the minimum required analogue components and subsequent user interface required for this form of interaction (see 5.5.2.4). To answer the first part of this additional question: Can an interface layer for a Decentralised Internet be designed? The answer is “Yes” (see: 5.5.3). The combination of technology layers and decentralised concepts can be crafted through existing design patterns, into a series of user interactions which would allow a full spectrum decentralised engagement.

Following a critical reflection and expert consultation regarding the user interface developed within this research (see: 5.5.4 & 5.5.6), findings suggest that there are significant issues around the notion of accessibility. In terms of an everyday computer user, this research argues that there is considerable complexity, increased friction, cognitive load, and the obstacle of multiple missing mental models. Compounding this, is the argument that in order to initially engage with a decentralised model of this kind, a significant amount of internalised knowledge and cognition is required. It can then be argued that a full spectrum of decentralised engagement, through a Sovereign Boundary Mechanism, developed through existing design patterns is likely to be inaccessible without significant pre-knowledge, commitment, support and participant training.

This research has also argued, that there are a number of unforeseen consequences. Issues relating to complexity, trust, personal responsibility and exclusion that may result in *The Paradox of a Sovereign Boundary Mechanism*. A situation which releases a participant from one form of incarceration, only for it to be replaced with another. This research doesn't claim that the development of a meaningful interaction with an ecosystem of this kind is impossible, rather that there is a great deal of work to do at a design level, to mitigate and reduce the frictions observed in its raw form, while developing a strategy for gradual user introduction and acceptance. Ultimately the level of risk, complexity and friction a user is willing to persevere with and accept is related to the ecosystems value proposition. Given the complexity and frictions observed, in what is a projection of the current trends, the value proposition would need to be rather substantial to allow for meaningful adoption at any scale.

There is an argument, that many of the issues identified at the interface layer within this research can be reduced or eradicated through the incorporation of Artificial Intelligence. This research has argued that this solution should not be seen as a panacea and that users require a degree of understanding of the inner workings of the system in order to see its value. This research argues that any introduction of AI, should be balanced against the value proposition in this context.

In summary, an Interface layer can certainly be established, but critical reflection and expert consultation has revealed significant obstacles in terms of accessibility and usability, which need to be addressed. These obstacles are related to complexity, friction, mental models and internalised cognition. This research argues that these obstacles may well be reduced and overcome, but that this will require significant efforts, through a systematic design led approach.

7.1.2.2 Additional Question 2:

How Might a Decentralised Internet Emerge and be Adopted?

The final question concerns the sustainable and established component of the main question that relates to decentralised technology and the *Diffusion of Innovation*. How might this technology be established in the mainstream? How might it emerge and be adopted in a sustainable manner? The emergence of a truly Decentralised Internet, given the dominance of the centralised model, and the monopoly held by a limited number of dominant forces, marks a paradigm shift. A decentralised model is not only reliant on the technology and interface presented to the participant, but also a value proposition, a minimisation of risk, responsibility and friction, and a positive perception of the switching costs.

This research suggests that achieving a *Diffusion of Innovation* for a Human-Centred Data Ecosystem, is a significant challenge. Initial adoption, perceived value, a critical mass and longevity, pose a *Wicked* design problem. A problem which will require a broad application of design thinking, and considerable coordination and strategy. At this point, this research suggests that there needs to be a significate reduction in friction and a considerable increase in value proposition.

A compounding issue is the funding of sustained development and defining decentralised business models, a decentralised ecosystem in the guise of a Sovereign Boundary Mechanism, requires infrastructure, distributed ledger, data storage, and continuous software and potentially hardware development. It is argued that in order to be decentralised, these components need to be open source and freely available. The core of the system needs to operate independently, without reliance on any centralised organisation or authority.

The components of a wider ecosystem are unlikely to materialise without significant investment and in the absence of intellectual property rights, realisation is challenging. This research suggests an emerging model one which involves larger organisations contributing to

the development of key infrastructure, while acting as custodians. The trend appears to be to establish the foundations before building business models around public services, agent and wallet applications and Vendor Relationship Management principles. This does appear to be a causality dilemma, though much work has already been undertaken at a foundation level, and it remains to be seen how the domain evolves.

This research argues that the drivers for adoption, are unlikely to come from the desire for privacy. The privacy literature, the theories of the Privacy Paradox and Instant Gratification Bias (see: 2.2.8) and several themes generated through the expert interviews, provide evidence of this. Instead drivers for adoption are more likely to come from the innovations made possible through decentralisation and the capabilities of a Sovereign Boundary Mechanism. A number of these opportunities have been identified within this research, centring around the Individual, Business and Wider Society (see: 6.3.4). This research argues that the perception of privacy across a spectrum of cultures is likely to be as a secondary consequence of decentralisation, and not its primary motivator. In short, the benefits to the individual focus around the reintroduction of peer to peer relationships, private communications, streamlining transactions, and reduction of the hidden costs of centralisation, through VRM principles. For business decentralisation levels the playing field for small and medium sized businesses, allows the exploration of new business models, enables the retrieval of quality real-time data and supports the building of quality customer relationships. For society the opportunities are significant, governance, planning, health and social care are all areas that will benefit from decentralisation. There are more fundamental cultural benefits at a macro level, centred around democracy, the importance of the ephemeral as a function of privacy, and the well-being and mental health of the general populous.

In summary, finding emergence and adoption of a Decentralised Internet is a formidable challenge, given the dominance of established service providers. The problem

space is complex and multifaceted, and it will require a significant reduction in participant friction and a significate increase in value proposition. There is much work to do to establish collective understanding, narrative, mental model and language, and the primary driver for value is unlikely to be a desire for privacy. This research suggests areas where decentralised interactions, supersede current centralised models and it is these areas, that offer the strongest direction of investigation, for a greater probability of adoption.

7.2 Main Findings and Arguments

The following section lists a summary of the main findings for this research, although there is overlap, the findings have been listed in line with the additional research questions, focusing on Interaction and the Interface Layer and general Adoption.

7.2.1 Interaction

- This research concludes that the core technological infrastructure is now in existence to facilitate a genuine sovereign Decentralised Internet. One which satisfies the principles of both Human Data Interaction and Self-Sovereign Identity (see: 5.4.1.4).
- Investigations illustrate that a dominant trajectory for a user centred Decentralised Internet with the advent of a functional Identity Layer, is progressing towards a Sovereign Boundary Mechanism through the concept of Self-Sovereign Identity (see: 6.1.2).
- By projecting forward the current trend of a Sovereign Boundary Mechanism, through a Conceptual Modelling exercise, this research has demonstrated a potential interaction model that is complex, requiring significant internalised cognitive processes and knowledge (see: 6.4.5).
- Though the core technological infrastructure is in place. The development of a Conceptual Model and subsequent preliminary user interface suggests a number of mechanisms and interactions which still need to be developed, to allow for a full User Centred Data Ecosystem (see: 5.5.4).

- This research suggests that a Sovereign Boundary Mechanism developed through existing design patterns paradoxically replaces one form of 'Benthamesque' incarceration with another (see: 6.4.1).
- This research argues that the decentralised community should step back and reflect on the trajectory of the Sovereign Boundary Mechanism, address the issue of participant interaction and investigate the underlaying HCI theory. The rational being, too reframe the current thinking across a full range of HCI cognitive theory, or potentially investigate radical alternatives to existing models (see: 6.4.6).

7.2.2 Adoption

- This research concludes that the concept of privacy, in the context of a Decentralised Internet is poorly defined and miss-understood. That participants desire privacy, but struggle with it as a concept and fail to see its value across context and cultures (see: 6.3.1).
- Privacy as a justification for adoption should not be seen as the primary driver or message and that the privacy benefits of decentralisation are potentially a second order consequence (see: 6.3.2).
- This research concludes that privacy should be considered as an umbrella term, and that innovations should focus on the specific problems and frictions posed by the centralised model, possibly following the taxonomy of privacy defined by Danial Solove (see: 6.3.1).
- This research argues that the Decentralised Internet cannot be marketed simply on the fact that it is decentralised. And instead the Decentralised Internet needs to supersede the centralised model in terms of its offering in order to raise the probability of adoption (see: 6.3.4).
- This research concludes that value can be developed by looking progressively forward, exploring concepts that go beyond a centralised model, focusing on the advantages and innovations that will emerge through a functional identity layer and its peripheral mechanisms (see: 6.3.4).
- This research argues that a major barrier to the adoption of a Sovereign Boundary Mechanism is the proportion of internalised cognitive process and information needed for initial engagement, coupled with a number of additional unforeseen frictions (see: 6.4.5).

- This research suggests that a cohesive strategy is required by the decentralised community, in order to achieve widespread adoption of a Sovereign Boundary Mechanism. One which collectively identifies and develops offerings of value through design thinking, while defining a consistent narrative, message and language to deliver targeted solutions within cultural contexts (see: 6.3.11).
- Ultimately adoption will emerge through a balancing of cognitive load required to engage and the genuine value proposition found within decentralised participation. At the time of writing, this research suggests that the value proposition delivered by contemporary offerings is relatively weak and that the cognitive load of existing artefacts and those on the horizon is considerably large (see: 6.4.8).

7.3 Contributions to Knowledge

The following paragraphs endeavour to summarise the contributions to knowledge found within this research.

7.3.1 Theoretical Contributions

This research makes a knowledge contribution through additions to the Academic Discourse by delivering a number of ideas and themes. A central objective within this research is: *'to make a contribution to knowledge, and expand the academic field surrounding the Decentralised Internet'* (see: 1.4). To this end, a number of examples are now listed.

7.3.1.1 Linking SSI and HDI

The academic domain for the Decentralised Internet, specifically focusing on the interface layer, participant experience, practical implementation and its dissemination into the mainstream is currently limited. Any academic alignment could comfortably fit within the broader HCI branch of HDI 'Human Data Interaction' (Haddadi, 2015; Hornung, 2015; Mortier, 2015). Indeed, this research broadens the domain of HDI, with the supported claim that it is intrinsically aligned with the principles of Self-Sovereign Identity (Allen, 2016; Tobin, 2016). By aligning the empirical advances found within SSI, the field of HDI can be expanded, as developments can be categorised, investigated and publicised within its framework. Alignment with SSI adds a dynamic element to the field of HDI, through the Identity Layer, broadening the manifestations of HDI from the current inward looking prospective (Databox, 2018), outwards towards a Human-Centred Data Ecosystem.

7.3.1.2 The Identification of Required Components For SSI

This research generates new knowledge, through the identification of the components, and interactions required by a participant to engage in a Decentralised Internet, through what is known as a Human-Centred Data Ecosystem. By identifying the required interactions,

paralleled with an investigation of HCI theory, this research presents what has been described as a Sovereign Boundary Mechanism. A significant practical element within this research, extended the current investigated trajectory of the decentralised domain to realise a preliminary, complete, user interface and participant experience model for SSI through a conceptual modelling method. The resulting work is original and at the time of writing, December 2019, is the only publicly available interpretation, manifestation, and visualisation of a complete SSI interaction.

7.3.1.3 The SBM, The Ecosystem and HCI Theory

This research contributes to the academic discourse through the application of HCI theory. The analysis of a Sovereign Boundary Mechanism, considered in parallel with the required wider data ecosystem, illustrates a spectrum of HCI theory. Participation with a Human-Centred Data Ecosystem will require a balance of internalised, externalised and distributed cognition. This research argues that the current development of SSI is placing the participant too far into the realm of the internalised, and that firstly this needs to be acknowledged and subsequently addressed. As such the discourse around realising SSI, should focus on balancing the cognitive loads of the participant, across application, context and situation.

7.3.1.4 The Paradox of a Sovereign Boundary Mechanism

This research introduces the notion of '*The Paradox of a Sovereign Boundary Mechanism'* into the academic discourse. As discussed in the previous chapter, the emancipation of the participant from the centralised network made possible through a Sovereign Boundary Mechanism, may have unforeseen consequences which replace one form of incarceration with another. This research draws attention to this possibility, justifies the claim, and argues that these issues need to be addressed to increase the probability of adoption.

7.3.1.5 Further Areas of Investigation and Research

This research highlights important areas of investigation and development, which will need to be considered if an SSI driven Human Data Ecosystem is to be realised. These areas include: Design Patterns, Mapping of HCI cognitive theory, Automation, Trust Frameworks, The Measuring of Friction and Design Strategies and Agendas. A full description of topics can be found in the following section of this chapter.

7.3.2 Practical Contributions

7.3.2.1 Framework and Resource For further Innovation

Through the review of the literature, the artefact catalogue and the development of a conceptual model and subsequent User Interface, this research provides a valuable framework and resource for designers and developers. This research represents an original body of work that can be applied to the preliminary stages of research and development of artefacts, acting as a foundation of knowledge, and reference for design teams developing concepts, technologies and tools.

7.3.2.2 The Problem Space for a Human-Centred Data Ecosystem

Through this research, additional required components and mechanisms have been identified that are required for a complete Human-Centred Data Ecosystem. This research has raised the issue of complexity, internalised cognition, absent mental models, friction, risk, responsibility, trust, and exclusion. This research presents a defined, structured problem space, which can be utilised by the design community as a framework for future development.

7.3.2.3 Adoption Strategy and the Identification of Value Proposition

The synthesis of literature, Public Survey, the Artefact Catalogue, and thematic analysis of expert interviews provides an original resource for the development of adoption strategy and value proposition. The argument that privacy is not the primary conscious driver and that the narrative needs to change to reflect real world privacy issues, coupled with the suggested areas where decentralisation offers value and supersedes the current centralised model. Offers direction and catalyst for the development of adoption, communication strategy and value propositions.

7.4 Recommendations for Future Research

The following paragraphs communicate the identified areas for further investigation generated within this research.

7.4.1 More Problems than Solutions

At an early stage, it became clear that the realisation of a functional, low friction Interface Layer for a Human-Centred Data Ecosystem was out of scope for this research. Instead, from a design prospective, this research has endeavoured to define the preliminary stages of a design process, establishing the boundaries of the problem space, developing a framework of knowledge to support further development of full and partial agency models of decentralisation. During this process, specific areas of interest where identified which require further investigation, in order to progress further towards the shared overall goal of a functional decentralised offering. These areas are now presented and discussed.

7.4.2 Alternative Design Patterns for a Human-Centred Data Ecosystem

This research argues that the trend and trajectory of the Decentralised Internet is towards Self-Sovereign Identity through a Sovereign Boundary Mechanism. And that a system which allows a participant full agency is potentially a highly complex, isolated, internalised interaction. The research suggests that in order to achieve sustained adoption an extension and re-evaluation of existing design patterns is required. The argument being, that a more radical approach is needed as existing table, form and page patterns are insufficient. This is likely to require a restructuring of interactions at both the interface and hardware layers. The pursuit of this innovation presents an opportunity for significant, important further practice-based research.

7.4.3 Mapping HCI Cognitive Theory

This research argues that the model of a Sovereign Boundary Mechanism, presents a significant internalised cognitive model of interaction and that this is undesirable given the contemporary understanding of the friction reducing benefits, of externalised and distributed forms of interaction. This research argues that a full agency interaction model for a Human Centred Data Ecosystem, incorporates a full spectrum of HCI cognitive forms (see: 6.4.6). In addition, this research highlights arguments that there is still contemporary value in some older HCI theory. This focus on mental models, especially when considered within a specific domain (see: 2.3.6). There is further work to be done and opportunities for theoretical research, in understanding the forms of Internalised, Externalised and Distributed cognitive patterns required within decentralised models.

7.4.4 Automating Interactions

This research has concluded, that a decentralised interaction in the form of an analogue Self-Sovereign Identity system, is complex and cognitively demanding. There are multiple complex interactions, across varied components which arguably lack metaphor and mental model. Many tasks required for a decentralised interaction are mundane and demanding, generating friction, making such systems unattractive. A side from the practical interactions, there are issues concerning the everyday participant shouldering the risks and responsibility, for the management and re-distribution of what can be sensitive, valuable personal information.

Within this research, there is the sense that many of these problems can be solved through the use of a personal Artificial Intelligence, and that decision making and repetitive or mundane tasks can be streamlined and off loaded. This research argues, that AI should not be seen as a Panacea, and that in some respects it represents one black box solution, being replaced by another. There is also the concern that concealing too much of the underlaying mechanism may reduce understanding of the internal function and in turn reduce the amount of perceived value. However, a genuine argument for some degree of automation. Under who's control and through what ethical framework remains to be seen. The topic represents a valuable seam for additional research and needs to be investigated further.

7.4.5 The Notion of Decentralised Trust Frameworks

As highlighted within the discussion section (see: 5.4.17), the concept of a trust framework in the context of decentralisation, is an interesting and necessary focus for further research. A Human-Centred Data Ecosystem needs some degree of solid ground, which in a centralised model is provided through federated identity and certificates. The notion of state offered verifiable credential, a physical biometric identity shop, a reputational model developed overtime, community assurance, and smart contracts where the code is law, are all examples of how identities, credentials and claims might be verified to support trust amongst

participants. But how do these elements inter-relate across multiple identities over-time, across participating groups, within a spectrum of low to high value transactions? These are all questions which can be investigated and offer a rich seem of future research to support what can be referred to as the Web of Trust.

7.4.6 Measuring Friction

Within this research, a conclusion has been drawn that a Human-Centred Data Ecosystem in the form of a Sovereign Boundary Mechanism through existing design patterns, represents a higher degree of friction in comparison to existing centralised counterparts. Literature has been reviewed, that argues the need for new evaluation methods for complex technologies, in which users lack the mental models to engage with original concepts (see: 1.3). Within the limitations section of this chapter, attention is drawn to the short falls of investigation and measuring of friction, and it can be argued that there is a great deal of work to be done considering how friction in the context of this domain can be measured, and how comparison can be made between centralised and decentralised artefacts. This measuring and interpretation of friction within decentralised models, offers a valuable seam of further research.

7.4.7 Design Strategy and Agenda

Within this research, it has been argued that the realisation of a Decentralised Internet, has now moved from the technical domain, into the design space. The necessary technology is now in place to facilitate a spectrum of decentralised models and what is now required, is a coherent collective design strategy within the decentralised community, to overcome many of the complex value and user experience problems. Running alongside this is the need to formulate consistent language and mental models to convey narrative to potential participants. In addition, an agreed collective agenda is required to drive adoption, the community needs to

identify points of least resistance, and collectively focus on these opportunity areas to drive adoption. The consideration and realisation of a Design Strategy and Agenda for the decentralised community, offers a significate opportunity for further practice-based research.

7.5 Methodology Limitations

Within the Methodology section of this thesis, pre-empted limitations are identified for each of the components described. As the research progressed limitations emerged that where not foreseen and these are communicated in the sections below.

7.5.1 A Multidisciplinary Investigation, that Needs to Go Further

This research has been challenging due to the broad multi-disciplined nature of the field. Approaching this domain from the prospective of a designer focused on the issues relating to adoption and participant interaction, meant an investigation of a complex problem space from a theoretical, technical and empirical prospective. The research has exposed the landscape, established the focused areas of investigation, while revealing a number of areas that require further exploration. This research has made strong contributions to knowledge but is still defining the variables of the problem space, and significant further work needs to be undertaken, to evolve this research past preliminary investigation, towards valuable meaningful solutions.

7.5.2 An Artefact Catalogue with Limited Quantitative Outcomes

The rational for the introduction of an Evaluation Matrix within the Artefact Catalogue, has been to build a degree of quantitative data. The intention being to define metrics from which balanced judgements of an offering's capability across several criteria might be evaluated. During the process it became apparent that the validity of this data might be questionable. The consistency of evidence through which to make judgements is inconsistent and varied across artefacts. This approach may not have yielded the metrics envisaged, but the

method has acted as a mechanism through which to reflect upon artefacts, and aids in a broader understanding of the offering for the researcher. This element may have shifted to be more qualitative then quantitative, but the component still offers valuable categorisation, historical context, and a clear vision of the domain's trajectory.

7.5.3 A Broad Survey That Might Have Been Better Focused

The Public Survey was designed to capture a broad range of information, and provides valuable descriptive statistics, capturing participant perceptions and indicators for both product development and adoption strategies. Within the survey there are components which define correlation and comparative statistics. The value of these advanced statistical methods is limited, and on reflection these elements may have benefited from sharper focus and planning. Likert scales may have been better designed in terms of the collections, and Likert items might have been tested more thoroughly to enable stronger internal consistency. Comparative components might also have been better considered in line with overall objectives. The Public Survey is broad and captures valuable data, but any further study might focus on one of the many detailed points of investigation, for example issues relating to gender, cultural context, acceptable levels of friction, and detailed areas of value and opportunity.

It must also be acknowledged that in recruiting participants from a predominantly student population there is a degree of sampling bias. With the majority demographic consisting of younger academically accomplished participants, arguments might be made that the sample is potentially better informed and more topically aware when compared to the general population. The prospect of bias had been considered during the planning stages, and efforts have been made to stratify the overall sample to include non-university students, professional and non-professionals, and participants from a broader age range.

7.5.4 Expert Interviews with A Great Deal of Data

The Expert Interviews have guided this research and generated a valuable collection of themes, which support the corroboration of assumptions and development of new theory. It is though surprising just how much information is held in the data at both the semantic and latent levels. As this research continues to develop, and further design work is undertaken, a fourth cycle of investigation may be fruitful from a creative prospective, as returning to the original recordings may prove to be a great catalyst for idea generation.

7.5.5 A Practice Led Component with Depths Still Left to Plumb

The Practice Led Component of this research, through conceptual modelling, preliminary user interface, and critical analysis, has generated the first phase of a structured interaction model for SSI. As described in the discussion chapter, this work remains at a problem space stage. The results of the research identify issues of complexity, missing components and mechanisms, and a surface level mapping of cognition. This design work clearly has a number of iterations left to cycle, as it moves from problem space to solution development. Extended user journeys, measuring of friction, cognitive mapping, cognitive off-loading, development of metaphor, automation, and the alignment with defined value proposition, are all elements which can be factored into subsequent cycles. Real users need to be engaged from varied cultural context, and any further stages may see the development of actual tangible artefact. Though this first cycle is limited, the outcomes are strong, they are informative, and clearly define the roadmap for further investigations.

7.6 Dissemination of the Research

Following the successful completion of this thesis it is the objective of the researcher not only to continue with this work, but also generate quality academic outputs. To date this work has been communicated and disseminated in the following ways.

MyData 2018 - Session Host

Helsinki Finland Fri 30th Aug 2018 15:00 - 16:00

Title: DESIGN Adoption, Usability, Mental Models, Interface, Bridging the Gap.

Together with Ian Forester from BBC R&D, this research, and the general topic of the design issues for a decentralised internet and Self-Sovereign Identity where discussed and debated.

A transcript of the session can be found in the appendix of this document (see: F2)

MyData 2019 - Conference Presentation

Helsinki Finland Fri 27th Sept 2019 12:15 - 13:30

Exploring Solutions from a Design and UI Perspective – The Challenges of a Sovereign Boundary Mechanism

This research was presented as part of the MyData Conference 2019. Following the presentation, the researcher participated as a panel member discussing the future development of SSI. Fellow Participants Included:

Ken Ebert: Software Architect and Open Standards Engineer, Sovereign Foundation.

Andre Kudra: CIO, ESATUS AG

Henrik Biering: CEO, PEERCRAFT

Sovrin DID UX Working Group Meeting

Tue 3/12/2019 16:00-17:00

Following an invitation from Kaliya Young, this research was presented to the Sovrin DID UX working group. The group which is part of the Sovrin Foundation, is made up of prominent individuals in the Self-Sovereign Identity domain.

ME2B Working Group Meeting

6/12/2019 17.30-18.30

Following an invitation from Lisa LeVasseur, this research was presented to the Me2B Alliance. The working group is a direct spur from project VRM, and as such has prominent contributors from the decentralised and identity domains.

Frontiers in Blockchain - Establishing Self-Sovereign Identity with Blockchain

2 Abstracts Submitted: 23/1/2020

SSI Friction, Value Proposition and Sustainable Adoption

Friction, Complexity and Cognitive Theory within SSI Interactions

Following an invitation from the editors, 2 abstracts have been submitted for this edition of Frontiers. A copy of the abstracts can be found in the appendix of this document. (see: F1)

The Editors Include:

Alan Sherriff: Consultant London, United Kingdom
Kaliya Young: Merritt College Oakland, United States.
Michael Shea: Independent researcher Connecticut, United States
Athina Karatzogianni: University of Leicester, United Kingdom

7.7 A Personal Reflection

The research journey has been both enjoyable and challenging, the commitment required to deliver a part-time doctoral thesis from an intellectual, practical, and emotional standpoint has been considerably more demanding that was ever imagined. The process stages and the carrying of a myriad of conceptual ideas and theories, while constantly attempting to synthesise a degree of new knowledge has been an enlightening experience. The distance travelled and the quantity of knowledge accumulated has been surprising, there are still holes in the scaffold, but the small slice of expertise developed has been worth the sacrifices made. The first steps into a welcoming research community and the prospect of participating in, and adding value to, such a genuine topical domain, is a very positive personal outcome delivering a great deal of satisfaction, building a momentum that I hope to perpetuate through the rest of my career.

Moving forward the focus now is twofold, firstly the further dissemination of the research and findings found within this thesis, and secondly further research and investigation, focusing on the recommendations highlighted within this chapter. The process of engaging with PhD study, has equipped me with the practical research skills, domain expertise, and the confidence to make further knowledge contributions within a subject area for which I have a genuine passion, and vocational commitment.

Appendix A - Artefact Catalogue

A1 Sample of the Artefact Catalogue

			Inception	Position	Location	Employees	Decentrilised	Maturity	Alive	Original	Value Proposition	Mental Model	Identity	Blockchain	Technology Rejuvenation	TOTAL	Individual	Commercial	Societal	
								0	1	2	3	4	5							
SSI TECHNOLOGY																				
CORE	SSI / FULL SYSTEM / INFRASTRUCTURE	SOVRIN	2016	Active	US	No data	5	3	5	5	3	2	5	5	NA		5	5	5	The Sovrin Network is a new standard for digital identity – designed to bring the trust, personal control, and ease-of-use of analog IDs – like driver's licenses and ID cards – to the Internet.
CORE	SSI / FULL SSI SYSTEM / INFRASTRUCTURE	VERES ONE	No data	Active	US	No data	5	3	5	3	2	2	5	5	N/A		5	4	4	Veres One is a fit-for-purpose blockchain optimized for identity on the Web
CORE	SSI / FULL SSI SYSTEM / INFRASTRUCTURE	UPORT	2016	Active	US	1-10	4	3	4	4	1	2	5	5	N/A		3	3	3	Uport is a full spectrum SSI solution built around Ethereum
CORE	SSI/CENTRALISED HUB	BLOCKSTACK	2013	Active	US	11-50	3	3	5	4	4	4	4	5	NA		5	4	4	Blockstack is a decentralized computing platform that puts users in control of their data and identity. Apps built on Blockstack make data breaches and trust violations an antiquated notion.
CORE	SSI/LIMITED IDENTITY MECHANISUM	Blockpass	2017	Active	No data	No data	2	3	5	3	2	2	5	2	3		3	3	2	Particepent can verify their identity by entering details and the submitting password and utility bill information > Once this has been verified I assume by a human being looking at the photgraph and passport. The participent can then us this idenity for
OI	IDENITY / CRYPTO / SOCIAL	Pillar Project	2017	Active	UK	11-50	3	3	3	4	3	3	4	4	N/A		4	2	4	This is an Interesting project in which Idenity is built into a cryto wallet that includes a Social Element.
CORE		Connect Me / Evernym	2013	Active	US	51-100	4	2	4	3	3	2	5	5	N/A		3	3	3	Evernyms Wallet application based around Sover
OI	SSI/LIMITED IDENTITY MECHANISUM	Blockcerts	2016	Active	Canada	101-250	3	2			1	2	5	5	N/A					Seems to be a development community for now… With an explination of SSI for what seems like qualifications.
OI	SSI/LIMITED IDENTITY MECHANISUM	Lifeid	2017	Active	US	1-10	2	1												Very early, mjust seems to be a website…
OI	SSI/LIMITED IDENTITY MECHANISUM	Securekey verified.me	No data	Active	Canada	No data	3	2	4	3	1	2	5	5	N/A		3	3	3	Embrionic SSI system that promises all thiongs, back end system and application, poor reveiws.
CORE	SSI/LIMITED IDENTITY MECHANISUM	Civic	No data	No data	No data	No data	3	2	4	3	1	2	5	5	N/A		3	3	3	Our Secure ID product allows you to securely provide authenticated identity to partner websites for account sign up and access, including technologies that no longer require legacy usernames and passwords.
CORE	SSI/LIMITED IDENTITY MECHANISUM	Shocard	2015	Active	No data	No data	3	2	4	3	1	2	5	5	N/A		3	4	3	Seems to be another SSI system for limited idenity but aimed more at the commercial application.
CORE	SSI/LIMITED IDENTITY MECHANISUM	Dominode	2017	Active	No data	11-50	3	2	4	3	1	2	5	5	N/A		3	3	3	Seems to be another SSI system for limited idenity at a very early stage, aimed it would seem at gaming, gambling…
CORE	SSI / COMPONENT	Archive	No data				3	2	0	4	2	2	5	0	5	15	4	3	2	We are the world's first identity shop…there to help you move through the web more safely and securely than you currently can. No more remembering multiple passwords and usernames. No longer will
OI	SSI / COMPONENT	Pribook	2013	Closed	UK+G81	1-10														A personal cloud network for "personal identity, vendor relationship management and personal network of things"
CORE	SSI / COMPONENT	Monax	2014	Active	UK	11-50	2	3	4	4	3	2	4	4	NA		4	4	4	(Formerly Eris) a platform "designed to support multiple blockchain clients connected to multiple blockchain networks with different smart contract interpreters all seamlessly connected to other

VRM PRINCIPLES																				
OI	VRM PRINCIPLES	MyWave	2014	Active	Asia-Pacific	1-10	1		2	5	2	4	4		0	0	17	2	5	0	Frank' puts the customer in control of getting personalised experiences anytime, anywhere, on any device.
CORE	VRM PRINCIPLES	WebOfNeeds	No data				3		3	1	3	3	3	1	0	3	16	4	2	4	A distributed marketplace driven by customer needs.
OI	VRM PRINCIPLES	Dappre	No data				2		4	5	1	3	3	2	0	2		3	2	1	Allow people and organisations to subscribe to you and share the data you want to share
OI	INTENT CASTING APPLICATION	About2Buy	No data				1		1	0	2	3	1		0	2	9	2	2	0	A Collaborative Commerce System to Align Internet Buyers & Sellers Via Multiple
CORE	INTENT CASTING APPLICATION	Greentoe	2012	Active	US	11-50	1		4	1	3	3	3		0	3	17	1	3	0	Finally...There's a New Way to Shop! Name Your Price & We Negotiate For You.
CORE	INTENT CASTING APPLICATION 'Indervidual'	Indie Dash Button	No data				3		2	2	4	3	2	3	0	1	14	3	3	0	This ... turns traditional advertising on its head, and removes the need for complicated targeting technology. Customers readily identify themselves, creating more valuable sales channels where guesswork is all but eliminated.
CORE	INTENT CASTING APPLICATION 'Indervidual'	Intently	2011	Active	UK	1-10	2		3	2	3	3	3	2	0	2	16	3	3	0	Request any service anywhere with Intently.co.
CORE	INTENT CASTING APPLICATION 'Indervidual'	EmanciPay	No data				2		1	0	2	2	2	0	0	3		3	2	2	Dev project for customer-driven payment choices
OI	INTENT CASTING APPLICATION 'Indervidual'	Pikaba	No data	Active	US	11-50	1		3	0	2	2	2	2	0	2	11	2	2	2	Pikaba is Social Shopping Platform that captures consumer intent to purchase and connects them with the right local business.
OI	INTENT CASTING APPLICATION 'Collective'	Crowdspending	No data				2		2	0	3	3	3		0		11	4	2	4	Gives each of us the power of all of us.
OI	INTENT CASTING APPLICATION 'Functional Trade'	RedBeacon	1995	Active	US	11-50	1		4	4	2	3	4	3	0	3	20	4	3	2	Trusted pros for a better home.
OI	INTENT CASTING APPLICATION 'Functional Trade'	TaskRabbit	2008	Active	US	51-100	1		4	4	2	3	4	3	0	3	20	4	3	2	"Tell us what you need, let us know what we can take off your plate, choose a Tasker, hire one of our fully vetted Taskers to get the job done."
OI	INTENT CASTING APPLICATION 'Functional Trade'	Zaarly	2011	Active	US	11-50	1		4	4	2	3	4	3	0	3	20	4	3	2	Hire local, hand-picked home services. We moderate every job and guarantee happiness at virtually any cost.
OI	INTENT CASTING APPLICATION 'Functional Trade'	Thumbtack	2008	Active	US	501-1000	1		4	4	2	3	4	3	0	3	20	4	3	2	We help you hire experienced professionals at a price that's right.
OI	INTENT CASTING APPLICATION 'Functional Trade'	HomeAdvisor	No data				1		4	4	2	3	3		0	4	20	3	3	3	We help you find trusted home improvement pros.
		OpenBazaar																			
OI	PERSONAL ASSISTANT / AGENT	Magic	No data				1		4	4	3	4	4	2	0	5	24	3	1	0	Text this phone number to get whatever you want on demand with no hassle…
OI	PERSONAL ASSISTANT / AGENT	MyTime	No data				1		4	5	3	3	3	2	0	5	23	2	3	2	"Book appointments for anything."
OI	PERSONAL ASSISTANT / AGENT	Operator	No data				1		2	0	2	2	3	1	0	3	12	2	2	0	Operator is an iPhone app that connects you to experts to help find you what you want, when you want.
OI	PERSONAL ASSISTANT / AGENT	PricePatrol+C54	No data				1		2	0	2	2	2	2	0	2	10	3	2	2	Monitors nearby stores for what you want at the price you want
OI	PERSONAL ASSISTANT / AGENT	yellCast †	No data				4		2	2	3	3	3	1	0	3	16	4	2	4	What you want, where you want it.
OI	PERSONAL ASSISTANT / AGENT	Flamingo	2013	Active	US	11-50	0		3	3	2	0	2	0	0	NA		0	3	0	Service company working toward leveling the playing field between service providers and customers
CORE	PERSONAL ASSISTANT / AGENT	TrackIf / MyAlerts	2013	Active	US	11-50	1		2	0	3	3	2	2	0	3	13	4	3	0	Track your favourite sites for sales, new items, back-in-stock, and more.

PERSONAL INFORMATION MANAGEMENT

OI	COLLATION PERSONAL DATA STORE	Arkkeo	2010	Active	Finland	No data	0		4	5	2	2	3	3	0	2		2	2	0	Automatically stores and archives all the purchase receipts, warranty, insurance, healthcare and travel documents you receive from the businesses and service providers that you deal with."
CORE	COLLATION PERSONAL DATA STORE	CozyCloud +	2012	Active	EU	11-50	1		4	5	2	3	3	2	0	1		3	4	0	A Personal Cloud you can host, customize and fully control
CORE	COLLATION PERSONAL DATA STORE	Digitteria	No data	Active	UK	No data	2		1	0	2	2	2	2	0	3		3	2	0	Puts people at the heart of managing their data usage. Digitteria's products allow people to protect and share their information with third parties on their own terms.
CORE	COLLATION PERSONAL DATA STORE	Digi.me	2009	Active	UK	11-50	2		4	5	3	2	2	2	0	1		1	3	0	A copy of everything you said or shared, every photo you posted, every friend you made, all safe in a personalised library on your computer forever. With digi.me you can see the whole story of you, one that you own and control
CORE	COLLATION PERSONAL DATA STORE	Meeco	2012	Active	Asia-Pacific	11-50	2		4	4	2	2	2	2	0	1		3	3	0	An easy way to manage your digital life and control your information, so you can decide who you want to share it with and be rewarded for being you!
CORE	COLLATION PERSONAL DATA STORE	Pikcio	2015	Active	Swiss	No data	4		3	5	3	3	3	3	3	NA		3	0	0	Pikcio With Pikcio, you are now able to take control over your digital life. Communicate with confidence with people you trust and retrieve your various data in a single and secure application.
CORE	COLLATION PERSONAL DATA STORE	FileThis	2011	Active	US	11-50	1		3	4	2	3	3	1	0	0		2	3	0	Take the paper and the work out of paperwork... Think outside your inbox.
CORE	PERSONAL DATA STORE WITH EXSTENDED CAPABILITIES	Mydex	2008	Active	UK	11-50	3		3	0	3	1	1	3	0	4		3	3	3	Enables individuals to exchange personal data with confidence.
OI	PERSONAL DATA STORE WITH EXTENDED CAPABILITIES	Enigma	No data				5		1	3	4	NA	NA	4	5	NA		NA	NA	NA	A decentralized cloud platform with guaranteed privacy. Personal data is stored, shared and analyzed without ever being fully revealed to any party.
CORE	PERSONAL DATA STORE WITH EXSTENDED CAPABILITIES	Sandstorm	2014	Active	US	1-10	4		4	4	3	2	2	NA	0	3		5	2	5	An open source platform for personal servers.
CORE	PERSONAL DATA STORE WITH EXSTENDED CAPABILITIES	PeerCraft	No data				3		2	3	3	2	3	4	0	NA		3	2	0	A personal agent offering a set of commerce tools to its individual private users. The main purpose for Peercraft is to support Vendor Relation Management - the new trade paradigm allowing consumers to
CORE	PERSONAL DATA STORE WITH EXSTENDED CAPABILITIES	CloudOS	No data				3		2	2	1	2	2	0	0	3		4	1	3	An operating system for personal clouds and PICOs (persistent compute objects).
OI	PERSONAL DATA STORE WITH EXSTENDED CAPABILITIES	Percloud	No data				2		0	0											Personal cloud free software. Currently in fund raising. Explained here.
CORE	PERSONAL DATA STORE BASIC	OwnCloud	2011	Active	US	51-100	3		3	3	1	1	1	0	0	2		3	0	0	Access your data from all your devices, on an open platform you can extend and modify.
CORE	PERSONAL DATA STORE BASIC	OwnYourInfo	No data				1		3	3	1	1	1	0	0	2		3	0	0	Your Personal Data Vault.
CORE	PERSONAL DATA STORE BASIC	The Locker Project	No data				4		1	0	2	1	1	2	0	3		3	0	0	Open source, personal data
CORE	PERSONAL DATA STORE WITH BASIC APPLICATIONS	Owncloud	2011	Active	US	51 - 100	4		4	2	2	3	4	0				4	0	2	Provides universal access to your files via the web, your computer or your mobile devices — wherever you are.
OI	PERSONAL DATA STORE BASIC	Personal.com	No data																		Your cloud. Your data. Your rules. (Note: Personal became TeamData in May 2016, and the Personal.com URL redirects there. The entry persists here because the company maintains its original

DAPPS

OI	DAPP	Dmail	No data				4		2	5	1	2	2	5	5	N/A	5	4	2	This is a decentrilised email communications platform with crypto exchange capabilitys.
OI	DAPP	Forms.id	No data				4		2	5	1	2	2	3	5	NA	4	4	2	Decentrilised, Form, servay and polling data application.
OI	DAPP	Graphite	2018	Active	US	1 - 10	4		4	5	1	2	3	5	5	NA	4	4	1	This is a decentrilsied versiuon of Google Docs.
CORE	DAPP	Bitpatron	No data				3		2	5	3	3	2	5	5	NA	4	1	1	Decentrilised music, Video, Podcast distrobution where you can earn Bitcoin…
OI	DAPP	Recall	No data				4		2	5	1	2	2	5	5	NA	4	1	1	Encrypted version of Google phones.
OI	DAPP	Blackhole	No data				4		2	5	1	2	2	5	5	NA	4	4	1	Decentrilised file transfer system.
OI	DAPP	Lannister Capital	No data				4		2	5	1	2	2	5	5	NA	5	1	1	A decentrilised way to track all of your Financial assets.
OI	DAPP	Scannie	No data				4		2	5	2	2	2	5	5	NA	4	4	1	An application that allows you to scan in your documents and then store them securley via blockstack in an area of your choice.
OI	DAPP	Lander	No data				4		2	5	1	1	2	4	5	NA	4	1	1	The application would seem to offer a free homepage with data storged through blockstack. Not sure just how decentrilised wthis is in terms of the URL and who might host or control that.
OI	DAPP	Sigle	No data				4		2	5	1	1	2	5	5	NA	4	2	0	This is a decentrilised blogging platform,
CORE	DAPP	Encrypt My Photos	No data				4		2	5	1	3	2	5	5	NA	4	3	2	This is an interesting application, nothing new about the storage of photos, it is though free and it would appear unlimited.
OI	DAPP	BlockVault	No data				4		2	5	1	3	1	5	5	NA	4	2	2	It would seem to be a decentrilise password manager, this is good but might be a leap of faith to push all your most presisise data into the ether.
OI	DAPP	DPAGE	No data				4		2	5	1	3	1	5	5	NA	4	2	2	Very simple webpage holder, single page.
OI	DAPP	SpringRole	2014	Active	US	11-50	4		2	5	3	2	1	5	5	NA	4	3	1	This is a decentrilised Linkedin where the personal data is held through block stack. There is also an elemet oif verifiable cridentail.
OI	DAPP	Xor Drive	No data				4		2	4	1	1	1	5	5	3	4	3	1	This would still seem to be ain a developer format requiring build. It's a decentriled storage system.

Appendix B - Public Survey

B1 Research Method Plan Check List

Defined by John Creswell (Creswell, J, 2003)

Purpose of a survey design stated.
The purpose of this survey design is to gain a general sense of the beliefs and attitudes towards internet privacy, personal data collection, secondary data use and data protection. At this time the survey is not being designed to validate any hypothesis, instead its primary goal is to inform a later practice led design process.

Are the reasons for choosing the design mentioned?
A Survey method has been chosen as it offers a consistent, efficient method to establish statistically reliable quantitative data around a number of defined areas of interest.

Is the nature of the survey (cross-sectional vs. longitudinal) Identified?
The nature of this survey is cross-sectional, the interest is in the attitudes and activities of individuals at this moment in time. A longitudinal study would be very interesting and could potentially be developed overtime based on this study, but in the context of this research we need to know what the current position is in order to answer the research questions.

Are the population and its size mentioned?
The target population within this study is that of the digital native, aged 18-44, living within a Western Liberal Democracy, EU, North America. A sample of younger participants 18 and below may also generate significant data, though this has ethical implication and is outside of the scope of this first study.

Will the population be stratified? If so how?
Stratification will be implemented to remove individuals participating who are outside of the EU and North America. The rational for this is tied to the Privacy Views and Cultural differences which might be found in other countries. The Survey will then include these participants and a contrast will be made, this may result in some interesting data and knowledge, though the primary focus within this first survey will be what are considered territories within a Western Liberal Democracy. A balance of gender will also be considered as it is important that a balanced representation is achieved. If there is an imbalance individual from the dominant set will be randomly removed.

How many people will be in the sample? On what basis was this size chosen?
An important consideration in calculating the required sample size is that this survey is being undertaken is not primary to prove a null hypothesis but instead to gain a general sense of a belief or attitude, as such smaller sample size can be drawn. With reference provided by Fowler (Fowler, F, 1998) a confidence range of +- 5% in a sample result of 50% across any given question can be estimated from a sample size of 300 people. A +- 5% error margin within this study is acceptable.

What will be the procedure for sampling these individuals (e.g. random. Non random)?
The Sample will be random in nature, having been draw through broadcast email invitation and open survey, where the subject will happen across the opportunity to participate.

What instrument will be used in the survey?
Who developed the Instrument?
A Internet Mediated Questionnaire Survey method has been chosen as it offers a consistent, efficient method to establish statistically reliable quantitative data around a number of defined areas of interest. Observational methods and semi-structured interviews might have been engaged to generate similar data, but the timescales involved would have been impractical and impossibly labour intensive. An Internet Mediated Questionnaire Method has been chosen because it offers a number of advantages and efficiency's over a physical alternative. The digital survey allows for a broader cast of participant engagement, the survey will be conducted via direct email request, and in the field where participants are physically approached and asked to conduct the interview via a supplied handheld device. The Internet Mediated Option also allows for quick and easy download of data in

a suitable digital format. In this case, following recommendation, cost implication, and information regarding data protection, So-Go-Survey has been chosen as the digital survey tool (sogosurvey.com, 2017).

What are the content areas addressed in the survey? The scales?

The survey will investigate the following areas:

- Understand the participants level of awareness of surveillance capitalism.
- Understand the participants attitude towards personal data collection and use.
- Understand the general level of concern with regard personal data collection.
- Any actions taken by the participant to combat any concerns and protect data.
- Attitudes towards potential agency and adoption of decentralised technologies.

The majority of responses to the questions derived from the themed areas will be ordinal in nature and as such will be unidirectional Likert type.

What procedure will be used to pilot, or field test the survey?
The survey will first be piloted with a small number of academics with knowledge of the chosen field. The survey was then piloted within a wider group of academics. Each time questions will be asked with regards the wording and clarity of the questions, the suitability of responses, and any views on bias.

What are the variables in the study?
A clear list of variables that will be considered within the survey have been systematically listed within a data requirements table that can be found within the appendix of this document.

How do these variables cross-reference with the research questions and items on the survey?
Cross referencing of all of the variables within the defined category's, relative to both research questions and objectives can be found in the data requirements table that can be found within the appendix of this document.

What specific steps will be taken in data analysis?

Stratification
Male and female will be separated into two groups, non-EU–North America subjects will also be separated. A balance between both male and female will be achieved through random sampling.

Descriptive analysis
The Mode will be calculated for each ordinal response type question so the central tendency can be considered. Descriptive analysis will also be undertaken on visualize the data by way of charts so that the overarching responses can be better understood.

Statistical Analysis
Data will be coded to numerical values; these values will then be entered into the Statistic Analysis Package SPSS. The package will be used to identify correlations between the Subjective States within the ordinal questions. A Spearman's Rank-Order Correlation method will be used, and subsequent correlation data will be studied, any significance identification will be investigated, and hypothesis can be draw with regards cause and effect.

2										
3	New	Understanding the degree of Network Activity. (attribute)	Level of Network engagement.	Ordinal Scale		How would you describe your level of internet engagement 'email, browsing, social networks etc.' ?	Extremely engaged/ Very engaged / Moderately engaged / Slightly engaged / Not at all engaged		Forced Binary	
4	New	Understand how many devices the participent has and engages with.. (attribute)	Number of engaged devices.	Numeric		How many different types of devices do you use to access the internet? Laptop, Smart Phone, TV, Games Console.....	1/2/3/4/More		**EFFORT MADE TO PROTECT PRIVACY**	Forced Binary Survay 9
		WHAT DO PEOPLE UNDERSTAND AND ARE AWARE OF IN THE CONTEXT OF PERSONAL DATA AND THE NETWORK	**UNDERSTANDING OF PERSONAL DATA AND ITS VALUE**	Likert scale 7 Likert Objects					**INTEREST IN CONTROLLING PERSONAL DATA**	Forced Binary Survay 4
5	New	Understanding the knowledge level and what the participant considers personal data and information on the network. (knowledge)	Personal Confidence in the understanding of what personal data actually is on the network.	Ordinal Scale		In the context of your internet activity, how Confident are you that you really understand what personal data actually is?	Extremely confident / Very confident / Moderately confident / Slightly confident / Not at all confident			
7	New	Knowledge and Awareness of the practice of personal information collection on he network. (knowledge)	Personal understanding of data gathering on the network.	Ordinal Scale		How aware are you that your personal data might be gathered while you use the Internet?	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
7a	New	Knowledge and awareness that personal brewing patterns are tracked while engaging with the network. (knowledge)	Personal understanding that brewing patterns might be gathered online.	Ordinal Scale		How aware are you that your browsing patterns might be tracked as you move across different parts of the internet?	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
8	New	Knowledge and awareness that personal data has value to a spectrum of different company's, individuals and organisations. (knowledge)	Personal understanding of the commercial value of personal data.	Ordinal Scale		Are you aware that your personal data is valuable to many different types of companies and individuals?	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
9	New	Knowledge and awarenesthat there is a market place for personal data and information. (knowledge)	Personal understanding of the market place for personal data.	Ordinal Scale		How aware are you that collected personal data is actively bought and sold by data brokers on the open market?	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
10	New	Have individuals ever considered the value of personal data other then directly financial, (the value to our overall economy, planning, social benefits, health and well being) (knowledge)	Personal Understanding of the value of personal data outside of marketing.	Ordinal Scale		Are you aware that personal data has great value and potential not only for marketing purposes but also wider society, health care, education, planning etc..	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
11	New	Understanding of existence, prospect and impact of the IOT. (knowledge)	Personal Awareness and understanding the IOT.	Ordinal Scale		How aware are you of the IOT 'Internet Of Things' and the implications this technology might have on the amount of data different objects, items, persons and things will produce?	Extremely aware / Very aware / Moderately aware / Slightly aware / Not at all aware			
		EFFORT MADE TO PROTECT PRIVACY	**COMFORT LEVEL WITH NETWORK ENGAGEMENT**	Likert Scale 5 Likert Objects						
12	New	The view of the balance between data disclosed and the services received. (opinion)	Individual opinion of the fairness of the current relationship between service provider and user.	Ordinal Scale		When you visit or use online services that are free to use 'Facebook, Google, Twitter' it can be argued that you pay for this service with the disclosure of your personal data and information. Do you think the disclosure of your personal data and information is a fair exchange for the services you receive?	Extremely fare / Very fare / Moderately fare / Slightly far / Not at all fare			
13	New	The Sense of control over personal data felt by the participant while online. (perception)	How in control of their personal data a participant feelswhile engaging on the network.	Ordinal Scale		How would you describe the amount of control you feel you have over the amount of data collected about you online?	Extremely in control / Very in control / Moderately in control / Slightly in control / Not at all in control			
14		Confidence level that harvested personal data will be kept secure. (perception)	Confidence in the security of personal data disclosed on the web.	Ordinal Scale		How confident are you that the information you disclose will be kept private and secure?	Extremely confident / Very confident / Moderately confident / Slightly confident / Not at all confident			
15		Views on the capability of 3rd Parties to uncover sensitive personal information if motivated to do so. (opinion)	Opinion of the capabilities of third party's to uncover secret information from data disclosed	Ordinal Scale		If motivated how difficult do you think it would be for a person or organisation to learn details about your past you would rather keep private?	Extremely difficult / Very difficult / Moderately difficult / Slightly difficult / Not at all difficult			
16		Option on weather company's should be achieve on the collect personal data and information on the network. (opinion)	Comfort level in the practice of data gathering.	Ordinal Scale		Overall, how happy are you that organisations and companies collect your personal data.	Extremely comfortable / Very comfortable / Moderately comfortable / Slightly comfortable / Not at all comfortable			

				when your on the internet / Content of your online chats hangouts / Websites you browse / the searches you perform / Apps or programs you use / Times of day you are online.				
				EFFORT MADE TO PROTECT PRIVACY	**Forced Binary Survay 9**			
19	**PEW:**		Do you engage in any of the strategies below in order to be less visible online.	Cleared Cookies and browser history. >>> Deleted / Edited something you posted in the past >>> Set Browser to disable or turn off cookies >>> Not Used website because it asked me for my real time >>>Used temporary username / email address >>> Post comments without revealing who you are >>> Asked someone to remove something posted about you >>> Tried to mask your identity >>> Used public computer to browse anonymously >>> Used fake name / untraceable user name >>> Encrypted your communications >>> Used	Binary Question		Do you engage in any of the strategies below in order to be less visible online ?	YES / NO	
				INTEREST IN CONTROLLING PERSONAL DATA	**Forced Binary Survay 4**			
21			Has the individual ever considered the control their personal data, with some sort of exstended technology. (opinion)	Has the individual ever thought of controlling their personal data?	Binary Question		Have you ever considered controlling access to your personal data, information and identity online?	YES / NO
22			If the means of control was avalible, would the participent be interested in using it. (opinion)	Would an individual like to have the ability to control their data?	Binary Question		If technology was available to allow you to better manage and control access to your own personal data would you be interested in using it.	YES / NO
23			Given the current understanding of the centrilised internet would the indervidual have the motivation to engage with technology. (opinion)	Do current concerns generate sufficant motivate to engege with decentrilsed technologys.	Binary Question		Do you think your concerns about your personal data online 'if any' would motivate you sufficiently to manage part or all of your personal data if the capability existed.	YES / NO / Not at this time /
24			Would the individual be willing to allow a third party to manage there personal information	Would the partcipent feel comfortable allowing a third party to manage presonal data.	Ordinal Scale		Would you feel comfortable in allowing another person or organisation to manage your personal data for you; if your best interests were guaranteed	Extremely comfortable / Very comfortable / Moderately comfortable / Slightly comfortable / Not at all comfortable
25			Would the individual be willing to allow a AI to manage there personal information	Would the partcipent feel comfortable allowing A1 to manage presonal data.	Ordinal Scale		Would you feel comfortable in allowing an Artificial Intelligence to manage your personal data, if you best interests were guaranteed.	Extremely comfortable / Very comfortable / Moderately comfortable / Slightly comfortable / Not at all comfortable
26			Would Individuals have any interest in engaging a technology that protected their personal information.	Would indervidula consider decentrilised technoogy.	Binary Question		If a service was available that would manage your personal data, and the benefits of that were clear 'Better internet security, protection of your identity, genuine recommendations for you, genuine advantages for society, fairer economy, protected democracy' is this something that you would consider paying for?	YES / NO
27	**CATAPULT:**		What sectors people trust most.	What sectors do people trust most with the personal data.	Predefined Choice		Which sector do you trust the most with your personal data?	**Media** (TV, internet companies such as Twitter / Google) **RETAIL** (online, mobile and street stores) **TELECOMS** internet providers, phone companies **PUBLIC SECTOR** (Government, Healthcare, Education) **FINANCIAL SERVICES** (Banks, insurance companies) **UTILITIES** (Energy Company's) **TRANSPORT** (Bus, Train, Plane, Taxi companies) **OTHER** Please State.
28	**CATAPULT:**		Motivation to share personal information	What would convince an individual the most to want to share their personal data.	Predefined Choice		Which one of the following options would most convince you to share your personal data?	If it was going to be used to benefit the UK economy / If it was going to be used to improve society (Healthcare, Education) / If it was going to improve my own personal experience with that organisation / If I was paid for the data /
			Comparison Information					
29			Option of weather any user should ever have complete anonymity online.		Binary Question		Do you think people should have the ability to use the Internet with complete anonymity?	YES/NO
30			Understand of weather a user ever reads terms and conditions.		Binary Question		Do you read the Terms and Conditions when agreeing to a service online?	YES/Sometimes/Never
31			Understing about weather the user actually even cares about there personal data online.		Ordinal Scale		In all honesty, at this moment in time, how concerned are you about the personal data you disclose online?	Extremely concerned / Very concerned / Moderately concerned / Slightly concerned / Not at all concerned
			Has the individual ever suffered a data batch		Binary Question		Have you ever been the victim of what you would consider a fraud, breach or an abuse of personal data?	YES/NO

B3 Survey Academic Pilot Feedback

20th March 2017

Dr Carole O'Reilly

7 What is the purpose of the project section: we aim to COLLECT your views; possible benefits section: in a BROADER more transparent way; confidentiality section: and other MEDIA of academic dissemination.
18 Is the question I work through it clear enough in its implication?
32 question 27 responses are all mis-spelled: FAIR not fare; question 31 has a misspelling of organisations.
34 spellings: organisations and benefit.
48 question 46 mis-spelling energy companies; 47: mis-spelling of organization.

Dr Wilfred Darlington

32 'Fair' typo in 27. Concept of fairness not defined in this Q. Organisation typo in 31

Dr Alex Fenton

14 You might want to consider using age brackets - particularly if that maps with similar studies, you could compare results. It could be worth stating on the opening page why you are capturing personal details. What are the research questions, what is the study trying to find out? Could UK be easier to find in the list?
26 You may not need full stops at the end of the words. You might consider 5 point likert scales for questions. Again, could this map to similar studies?
32 fare should be fair exchange in 27. You're now using 5 point rather than 4?
38. my real time? What about browser incognito?
44 could be 5 point Likert. Internet with capital I? It seems a little bit of a leading question.
51 could be 5 point Likert.

Dr Ivan Garcia

14 A very clear online interface.
26 Responses to 20 are not mutually exclusive - could a participant tick more than one? There is perhaps too much overlap between Qs. 22 and 24. There is a typo in Q. 22, i.e. 'orginisations'.
32 There are typos in Qs. 27 and 31 i.e. 'fare' and 'orginisations'. I would remove 'if motivated' from Q. 30 as I don't think adds much to the question being asked.
34 There is a typo in Q. 33. Also the answers are not mutually exclusive: could you allow for participants to tick more than one answer?
38 I didn't quite understand the fourth answer from the top.
45 I'd remove inverted commas around 'if any' in Q. 41

Professor Karl Dayson

7 I really like the check boxed above. The whole page is clear and easy to follow.
14 Its unusual to request personal details fist. Normally these are the last things you collect, mainly because it takes time way for the substantive material and they may not wish to complete this data. On age it is normal to ask in categories rather than specific age. plus I don't like this rolling text, as i can't easily see what I've typed.
18 You don't define the user categories in Q15, so I had to guess.
26 I'm not sure if asking awareness followed by the other questions really works. It looks like a quiz rather than a survey. Again the questions are quite subjective, what can you realistically learn from the answers unless you can quantify levels of awareness. i.e. I might say I'm 'very aware' but how do I know that?
32 You need to more consistency with the options made available. Likert scale usually allows for 5 options, with a middle one of something like 'neither yes or no'.
34 Q33, why could the respondent only select one choice?
38 Add a category for 'other' and allow the respondent to add their own comments
45 Try to avoid questions with more than one clause. Q41 is just difficult to understand and the others could be broken down into two questions.
48 These seem better questions, but Q47 again assumes I know something that will happen.
53 What I'm unsure of is the order of these questions. Do they make sense to the respondent. They should be able to follow a logically trail and these appear to be disconnected.

Dr Aleksej Heinze

7 Generally, it is a bit long! ;-) If you could reduce the word count and perhaps not repeat same messages - e.g. need for the title and the "invitation paragraph" - most will have read this in an email already? Also, your contact details perhaps are needed so that people can contact you if they have any questions about this work?
14. Do you really need me to re-enter my email address? Also, name and last name could be one field to speed up the process?
18 Q15 could mean different technical abilities - how the hardware or software works?
26 Some answer options have full stops and others don't? I don't think you need these really
32 Q 31 and Q 30 - perhaps add a comma to make the question text easier to read?
38 perhaps add "other" open entry box - I think there are other ways to make it more private...
48 I am not sure the first answer needs "UK" in there - economy in general would be easier to justify?
53 I wonder if as part of Q52 there should be "if yes please give some more details option?"

Dr Ben Sherley

under the description of project I'd suggest a fairly informal tone and explaining in lay persons terms. Not necessarily the detail. E.g. "we are interested in how people use the internet and their views on their personal data and how it is used". The less you tell them about why you want the data the less risk of bias.
under the confidentiality section I would suggest, unless there is a really good reason not to do so, that you state all data will be anonymized and will not be identifiable to any individual. I can't imagine any scenario where you would identify individuals so your current text may alarm some people. Given that you are interested in their views on personal data you risk rejecting people with strong views that want their data confidential.
under the type of information section I'd suggest again that a more lay persons language would be appropriate. I would have to Google "Re-decentralised Internet" for example. The less you tell participants, the less opportunity to bias any results.

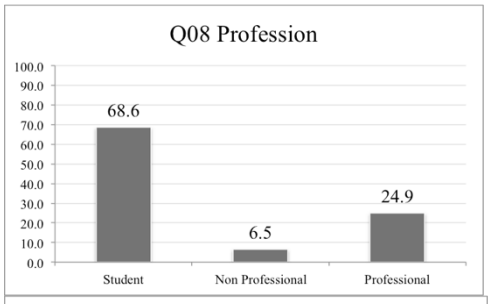
Professor Seamus Simpson

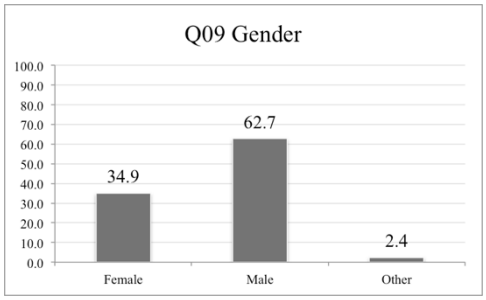
I've had a read through the survey, It seems mostly fine to me and likely to yield some interesting data.
I spotted some typos:
'aim to collection your views'
'majority of the questions'
'in a border more transparent way'
Q27 'fare'
Q33 and Q 47 'orginisation'
Other comments:
might be worth considering more what the possible benefits of participation are; or alternatively leaving this section out entirely
I think the line 'benefits you are entitled to in any way' is confusing and I'd remove it.
Also, the 'Re-decentralised Internet' is an unusual term that could do with some brief explanation (or don't mention it at all).
More generally, I think you have to be realistic about the percentage of responses you can get from an exercise like this. We did one recently and received a relatively low response rate - so worth having in mind the minimum number of responses you need and cast your net as wide as you need to from there.

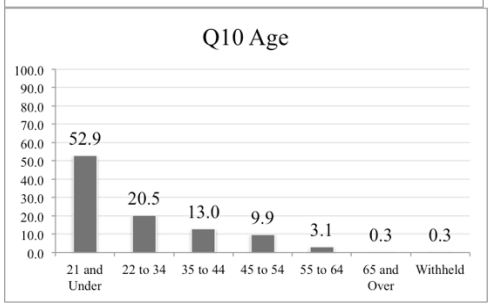
B4 Additional Survey Data

Descriptive Statistics - Individual Question Results
In the following section the raw data from all questions in the Public Survey are reported in bar chart format. The charts contain percentage values for both Likert elements and forced binary responses.

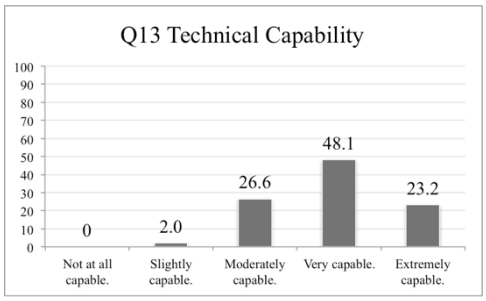
Demographic
Q8/9/10 Profession / Gender / Age



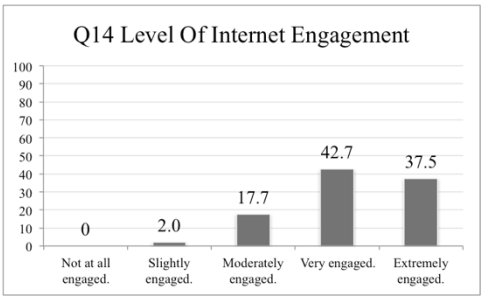




Digital Fluency
Q13 As a Computer User how would you describe your technical capability?

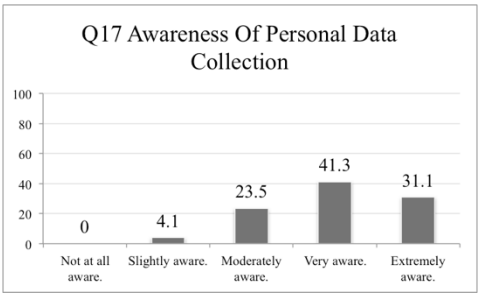


Q14 How would you describe your level of internet engagement 'email, browsing, social networks etc?'

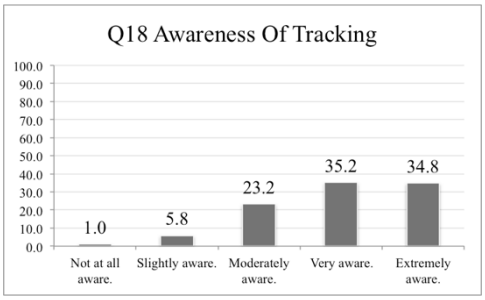


Understanding of Personal Data and Its Value

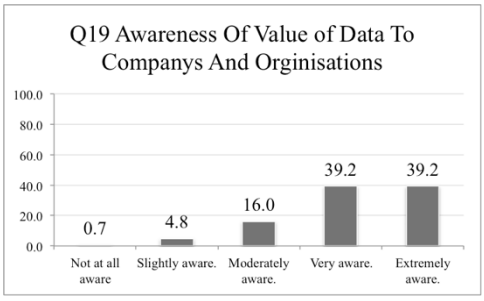
Q17 How aware are you that your personal data might be gathered while you use the Internet?



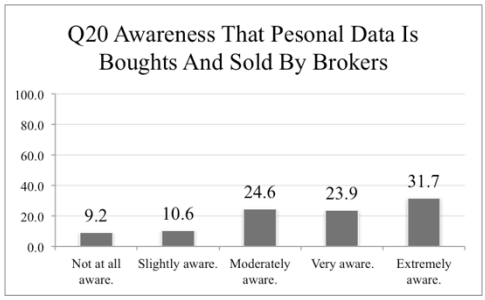
Q18 How aware are you that you're browsing patterns might be tracked as you move across different parts of the Internet?



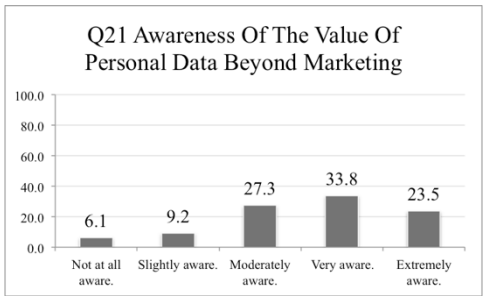
Q19 Are you aware that your personal data is valuable to many different types of companies and individuals?



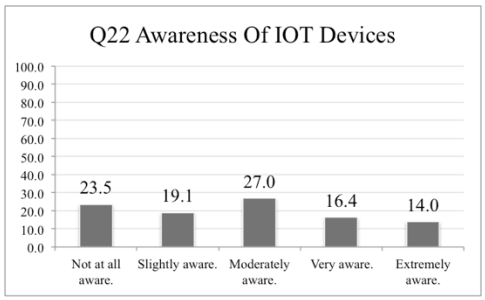
Q20 How aware are you that collected personal data is actively bought and sold by data brokers on the open market?



Q21 Are you aware that personal data has great value and potential not only for marketing purposes but also wider society, health care, education, planning etc..

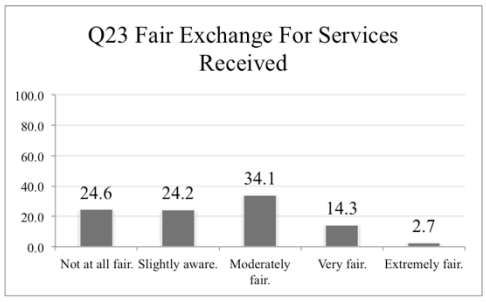


Q22 How aware are you of the IOT 'Internet of Things' and the implications this technology might have on the amount of data different objects, items, persons and things will produce?

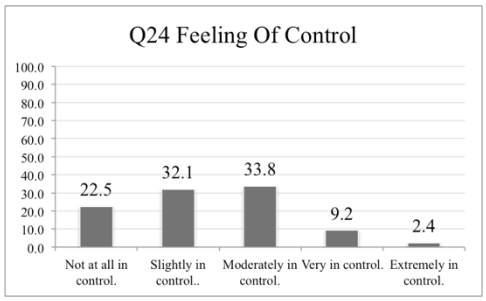


Comfort Level with Network Engagement

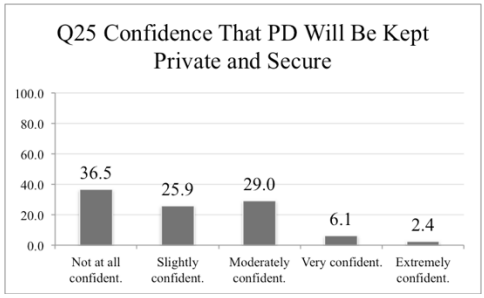
Q23 When you visit or use online services that are free to use 'Facebook, Google, Twitter' it can be argued that you pay for this service with the disclosure of your personal data and information. Do you think the disclosure of your personal data and information is a fair exchange for the services you receive?



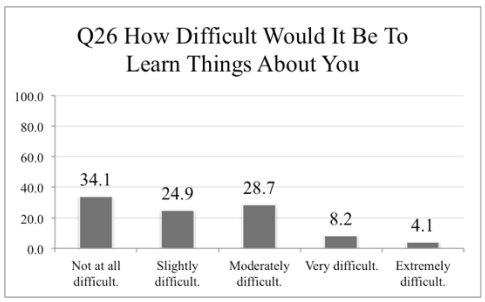
Q24 How would you describe the amount of control you feel you have over the amount of data collected about you online?



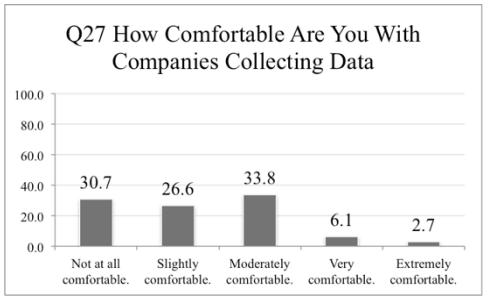
Q25 How confident are you that the information you disclose will be kept private and secure?



Q26 If motivated how difficult do you think it would be for a person or organisation to learn details about your past you would rather keep private?

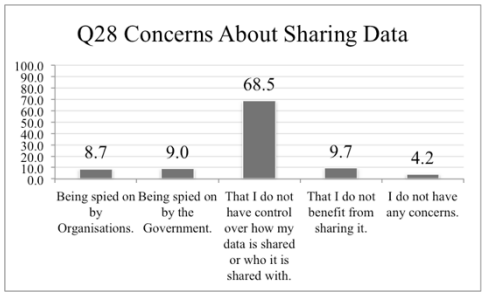


Q27 Overall, how happy are you that organisations and companies collect your personal data.



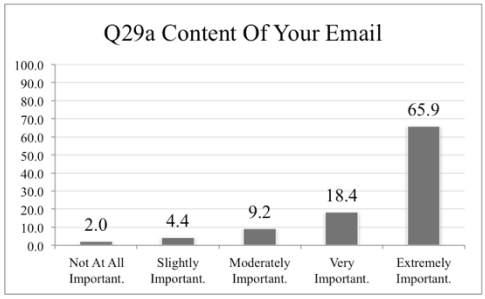
Specific Concerns

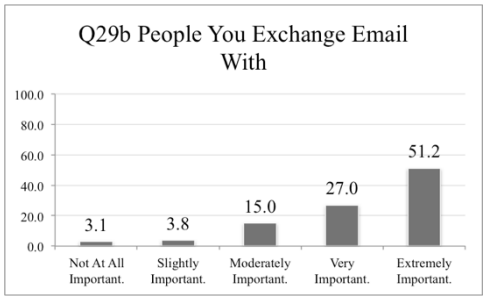
Q28 What most concerns you about sharing your personal data?

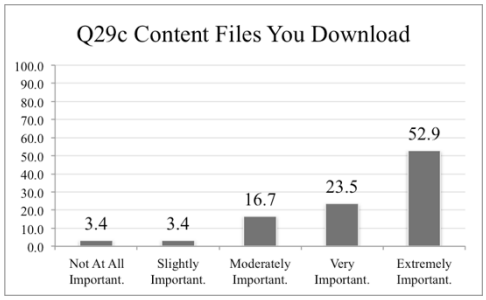


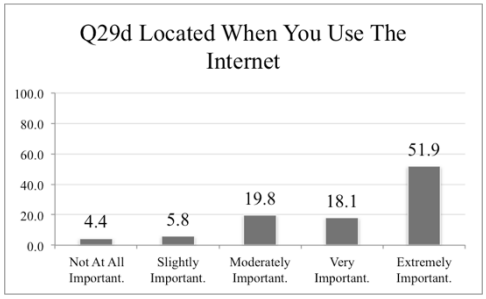
Perception of Importance

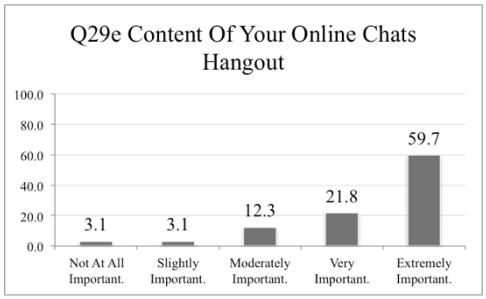
Q 29 How much do you care that only you, and those you authorise, should have access to the following information? a,b,c,d,e,f,g,h,i.

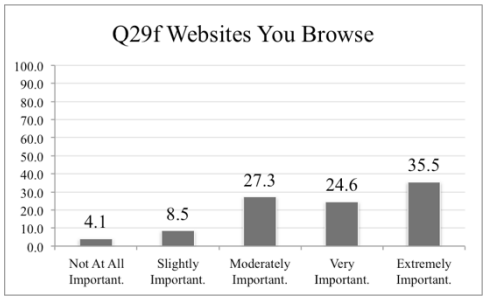


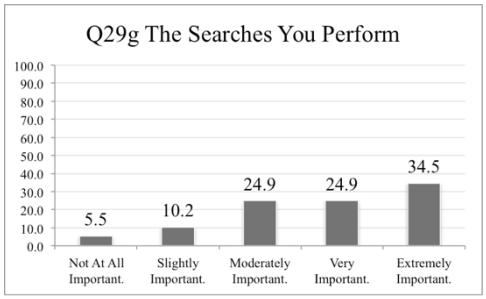


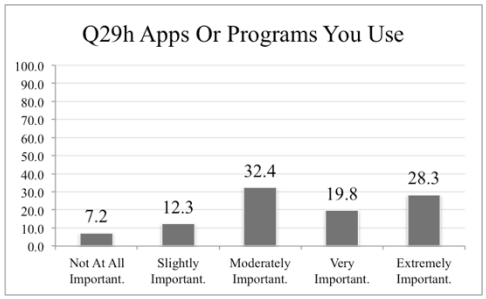


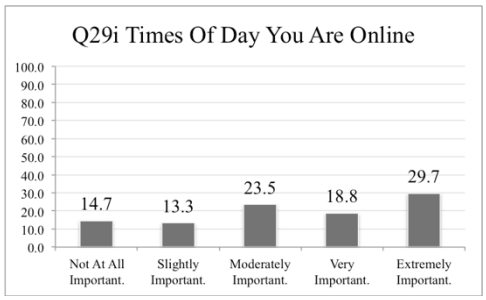






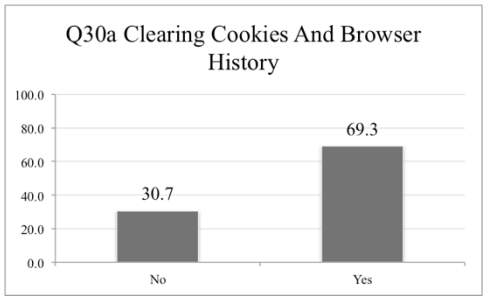


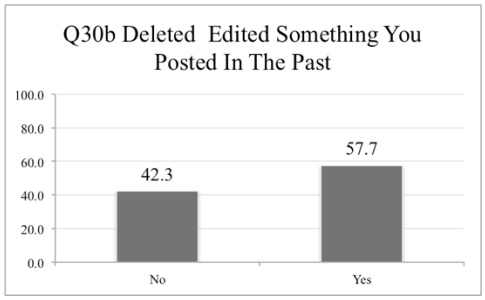


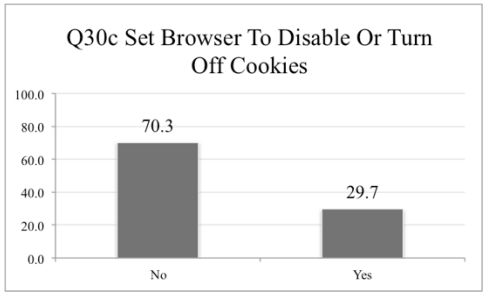


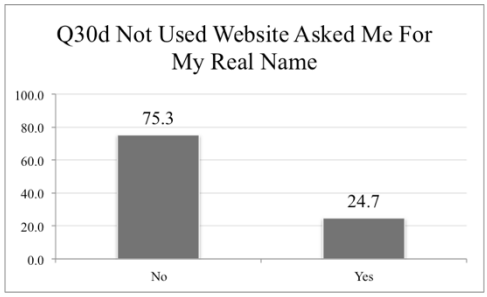
Effort Made to Protect Privacy

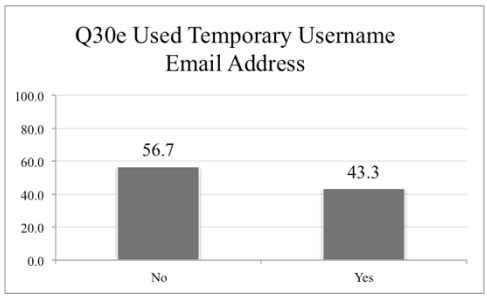
Q30 Do you engage in any of the strategies below in order to be less visible online ? a,b,c,d,e,f,g,h,i.

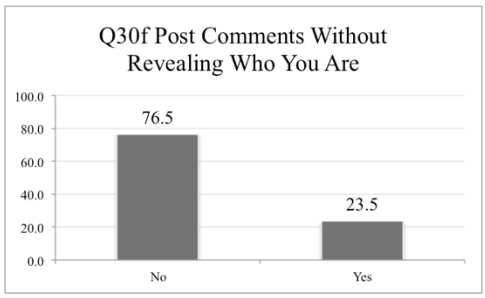


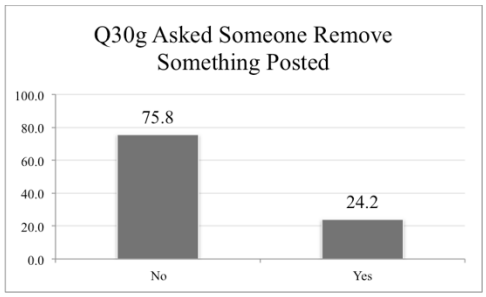


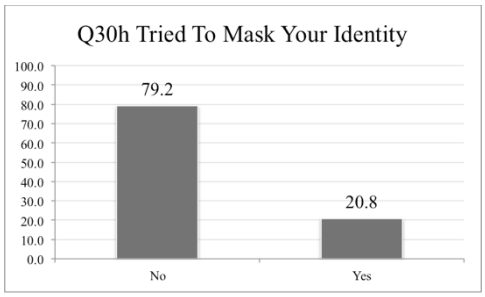


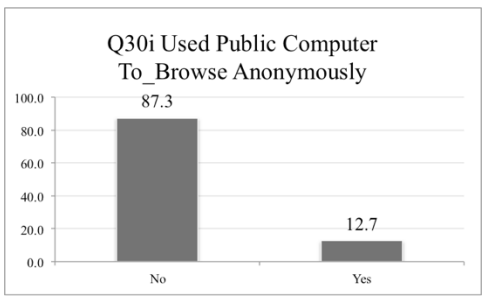


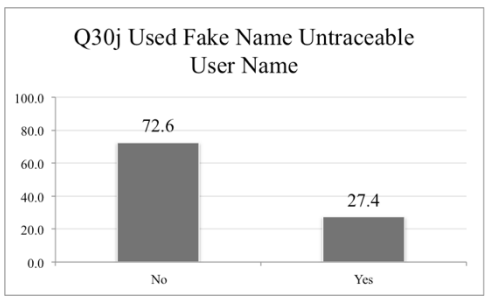


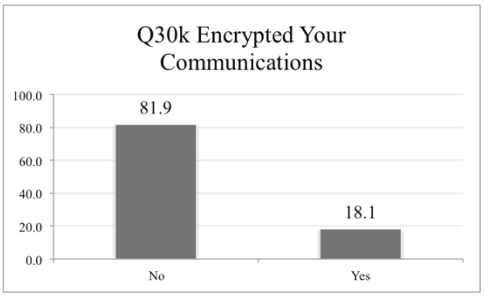


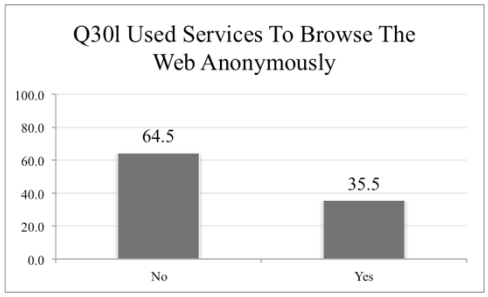


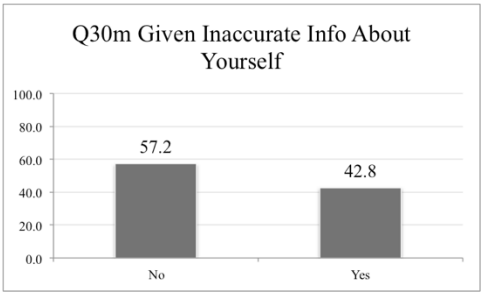






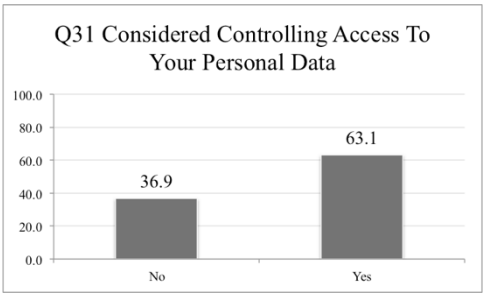




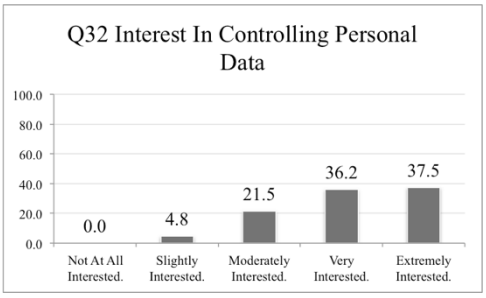


Understand What Participants Actually Do at This Time to Safeguard And Protect Their Personal Data

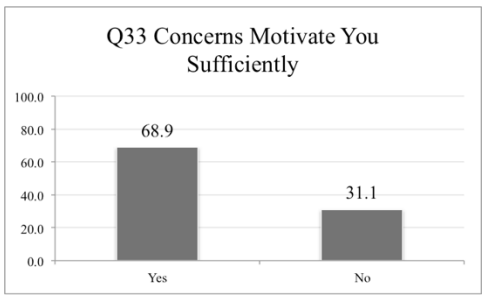
Q31 Have you ever considered controlling access to your personal data, information and identity online?



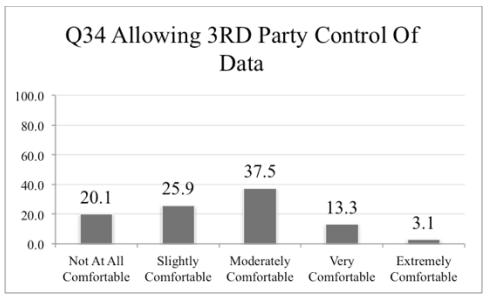
Q32 If technology was available to allow you to better manage and control access to your own personal data would you be interested in using it?



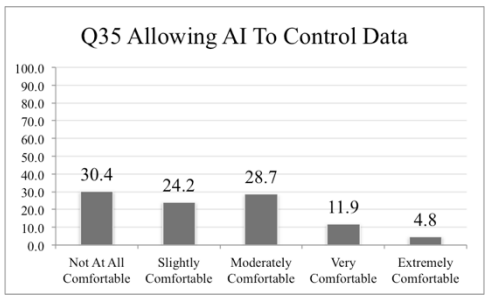
Q33 Do you think your concerns about your personal data online 'if any' would motivate you sufficiently to manage part or all of your personal data if the capability existed?



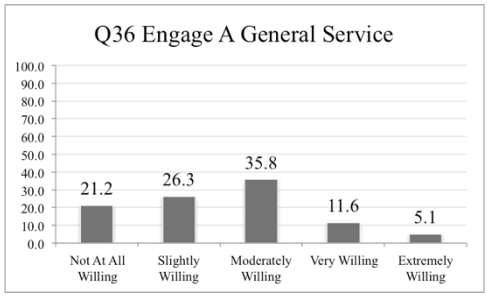
Q34 Would you feel comfortable in allowing another person or organisation to manage your personal data for you; if your best interests were guaranteed?



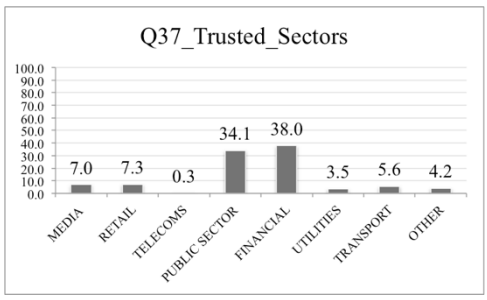
Q35 Would you feel comfortable in allowing an Artificial Intelligence to manage your personal data, if you best interests were guaranteed.



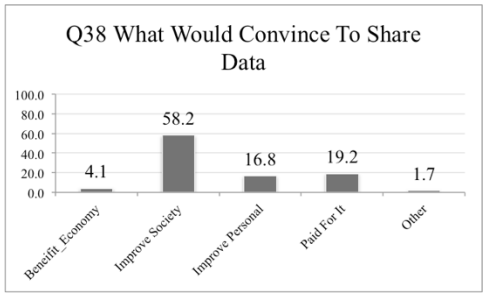
Q36 If a service was available that would manage your personal data, and the benefits of that were clear 'Better internet security, protection of your identity, genuine recommendations for you, genuine advantages for society, fairer economy, protected democracy' is this something that you would consider paying for?



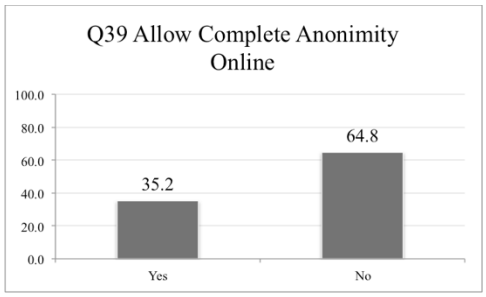
Q37 Which sector do you trust the most with your personal data?
Media / Retail / Telecoms / Public Sector/ Financial / Utilities / Transport / Other



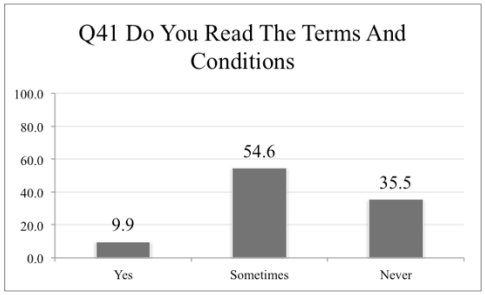
Q38 Which one of the following options would most convince you to share your personal data? Benefit Economy / Improve Society / Improve Personnel / Paid For It / Other



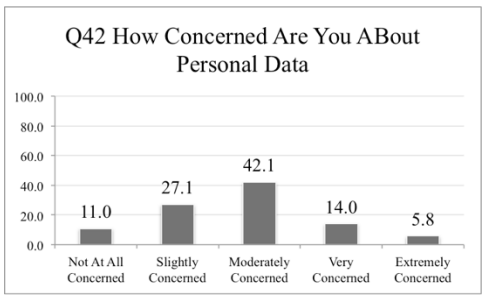
Comparison Information
Q39 Do you think people should have the ability to use the Internet with complete anonymity?



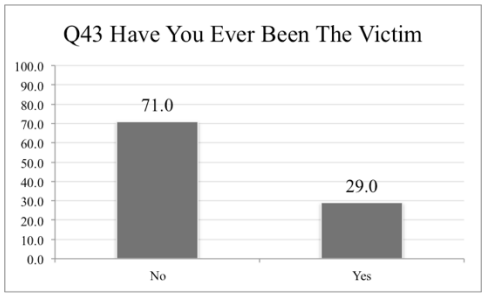
Q41 Do you read the Terms and Conditions when agreeing to a service online?



Q42 In all honesty, at this moment in time, how concerned are you about the personal data you disclose online?



Q43 Have you ever been the victim of what you would consider a fraud, breach or an abuse of personal data?



Appendix C – Expert Interviews

C1 Phase 1 Interviews Participants

David Alexander (D **Alexander** 2019, Personal Communication, 2nd Dec 2016)
Co-Founder Chief Executive and Platform Architect Mydex CIC /
https://www.linkedin.com/in/davidealexander/

Adrian Gropper MD (A **Gropper** 2016, Personal Communication, 23rd Nov 2016)
CTO at Patient Privacy Rights / Key figure in the Internet Identity Workshop / Harvard / MIT / https://www.linkedin.com/in/adrian-gropper-6916651/

John Laprise PhD (J Laprise 2019, Personal Communication, 2nd Dec 2016)
Principle Consultant - Internet governance, public policy, technology strategy, emerging technologies, technology forecasting.
https://www.linkedin.com/in/jplaprise/

Michael Linton (M **Linton** 2016, Personal Communication, 2nd Nov)
Originator of the term 'Local Exchange Trading System' / Prominent figure in the World of community currencies / Joint Initiator - Open Money Development Group /
https://www.linkedin.com/in/michael-linton-5ab50/

Jeremy Rushton (J Rushton 2016, Personal Communication, 27th Oct)
Founder Tiddy Wiki / EX-Head of Open Source Innovation BTExact / CTO CRTLio / Founder at Federation.
https://www.linkedin.com/in/jermy/?originalSubdomain=uk

Philip Sheldrake (P Sheldrake 2016, Personal Communication, 19th October)
Author of the Business of Influence (Sheldrake, 2011), Partner Euler Consultants Ltd, Founder of The Human Interface Project 'Hi-Project' AKASHA Foundation / Digital Life Collective / Current PhD candidate investigating Network Agency under Wendy Hall at Southampton University.
https://www.linkedin.com/in/philipsheldrake/

David Irvine (D Irvine 2016, Personal Communication, 9th Nov)
Founder - Co Founder of the Maidsafe Network.
https://www.linkedin.com/in/dirvine/

C2 Phase 2 Interviews Participants

Colin Wallis (C Willis 2018, Personal Communication, 11th Oct)
Executive Director – Kantara Initiative Inc
https://www.linkedin.com/in/colin-wallis-7921274/

Darrell O'Donell (D O'Donnell 2019, Personal Communication, 9th Feb)
CTO & Advisor – CULedger / Technology & Strategy Advisor – Continuum Loop
https://www.linkedin.com/in/darrellodonnell/

Gregor Zavcer (G Zavcer 2018, Personal Communication, 4th Feb)
Protocol & Foundation – Co-Founder Datafund
https://www.linkedin.com/in/holodeck/

Ian Forester (I Forester 2018, Personal Communication, 11th Nov)
Senior Producer at BBC R&D / Founder DataPortability.org
https://www.linkedin.com/in/cubicgarden/

Jonny Howle (J Howle 2018, Personal Communication, 4th Feb)
Product Designer at uPort (ConsenSys) / UX-UI Designer IBM
https://www.linkedin.com/in/jonnyhowle/

Kaliya Young (K Young 2018, Personal Communication, 21st Nov)
Identity Woman / CO-Founder Internet Identity / Leader in the field of Self-Sovereign Identity or Decentralised Identity Workshop
https://www.linkedin.com/in/kaliya/

Micheal Becker (M Becker 2018, Personal Communication, 28th Sept)
Founder, managing Partner – Identity Praxis, Inc.
https://www.linkedin.com/in/digitalsovereignty/

Phil Windley (P Windley 2018, Personal Communication, 12th Nov)
Chair at Sovrin Foundation / Fonder and Organiser – Internet Identity Workshop / Author of Digital Identity: Unmasking Identity Management Architecture (IMA) (2005)
https://www.linkedin.com/in/335indley/

Sunil Malhotra (S Malhotra 2018, Personal Communication, 5th Oct)
Founder CEO – Idea-farms – New Delhi – India
https://www.linkedin.com/in/sunil-malhotra/

C3 Phase 1 Questions with Decentralised Experts

PHASE 1 Interviews with Decentralised Experts
SEMI-STRUCTURED INTERVIEW
Version 1 (21.4.2016)

Interviewer introduction and proposed questions.
Hi, my name Is [*****], I am currently conducting research focusing on the Decentralised Internet, more specifically the design of user journey and interaction methods around establishing a personal data set, managing a data set, access control and system status. I have invited you to participate in this research given your expertise and knowledge of the field. I intend to conduct a semi-structured recorded interview. Please find a list of initial questions I would like to ask below.

Your background and work within the field.

Main question	Additional questions	Clarifying questions
Please could you tell me want your experience is with regards the Decentralised Internet and it's associated technologies?	Please could you tell me about your specific area of interest within the decentralised internet field?	Could you tell me anything about your perceptions of emerging areas of interest within the field, or areas you think might becoming topical in the near and medium term?

The direction of decentralised technologies and associated application.

Main question	Additional questions	Clarifying questions
What do you consider to be the direction of travel within the Decentralised field in terms of objectives and areas of current focus?	Looking forward, where do you think the decentralised field will be in the short to medium term, what do you think it will be achieved and what will be its impact on the interactions and perceptions within the main stream.	What do you perceive as being the long-term manifestation of the decentralised internet and its associated philosophy's and technology's?

The perceived barriers to adoption and the challenges ahead.

Main question	Additional questions	Clarifying questions
In terms of the direction of travel and the projects and technologies being developed, what do you see as the main barriers to entry for mainstream adoption?	What areas of the decentralised internet do you think will be adopted first and become mainstream?	Are there any strategies that you think should be adopted to stimulate the adoption of decentralised technologies?

User Journeys within a decentralized personal data framework.

Main question	Additional questions	Clarifying questions
Do you have any thoughts on the architecture of any decentralised system, how personal data is constructed, maintained, organised and distributed from the perspective of the user?	Do you have any thoughts on the way personal data is disseminated and organised as a user utilises any number of applications or components available to them?	Have you considered how any decentralised system might be structured in such a way as to give the user feedback on the system state and give confidence that the information and data is indeed secure?

User interaction and interface Issues.

Main question	Additional questions	Clarifying questions
In a decentralised system what are the core components of interaction a user might engage with during common tasks?	How do you imagine the interface should be constructed or designed to facilitate interactions within such a system?	What are the core areas you think should be considered, tested and developed first in order to develop a system that can be adopted by the mainstream?

C4 Phase 2 Questions with Decentralised SSI Experts

PHASE 2 Interviews with Decentralised Experts 'Self-Sovereign Identity'

SEMI-STRUCTURED INTERVIEW

Version 1 (21.4.2016) Revision for May 2018

Interviewer introduction and proposed questions.

Hi, my name is Mick Lockwood, I am currently conducting research focusing on the Decentralised Internet, more specifically the design of user journey and interaction methods around establishing a personal data set, managing a data set, access control and system status. I have invited you to participate in this research given your expertise and knowledge of the field. I intend to conduct a semi-structured recorded interview. Please find a list of initial questions I would like to ask below.

'As of May 2018, these questions have been simplified and are now focused on the investigation of *Self-Sovereign Identity*

Your background and work within the field.

Main question	Additional questions	Clarifying questions
Please could you tell me want your experience is with regards the Decentralised Internet and it's associated technologies?	Please could you tell me about your specific area of interest within the decentralised internet field?	Could you tell me anything about your perceptions of emerging areas of interest within the field, or areas you think might becoming topical in the near and medium term?

User interaction and interface Issues.

In the context of *Self-Sovereign Identity* and the process of:

Complexity….?

Establishing connections, Claims, Defining Zero Knowledge Claims, Contracts and Smart Contracts.

Main question	Additional questions	Clarifying questions
In a decentralised system what are the core components of interaction a user might engage with during common tasks?	How do you imagine the interface should be constructed or designed to facilitate interactions within such a system?	What are the core areas you think should be considered, tested and developed first in order to develop a system that can be adopted by the mainstream?

The perceived barriers to adoption and the challenges ahead.

In the context of Self-Sovereign Identity:

Adoption, Use Cases, Value Propositions, Potential Spearhead Trust Frameworks…..

Main question	Additional questions	Clarifying questions
In terms of the direction of travel and the projects and technologies being developed, what do you see as the main barriers to entry for mainstream adoption?	What areas of the decentralised internet do you think will be adopted first and become mainstream?	Are there any strategies that you think should be adopted to stimulate the adoption of decentralised technologies?

Additional Discussion >>>>

Conceptual Models / Mental Models, and the Complexity of Interaction.

Adoption, Use Cases, Business Models, Routes to Adoption.

C5 Thematic Analysis Code Book

Name	Description	Files	References
A Definite Use Case	Any example of decentralised technology being used.	12	41
A Political Issue Or Consideration	Any comment or conversation relating to political issues relating to the decentralised Internet.	2	3
Adoption	Any issues relating to the Adoption of Decentralised Technologies	0	0
Barriers To Adoption	Any opinion or examples of issue preventing adoption of decentralised technology.	8	36
Pathways To adoption	Any indication as to how decentralisation might be adopted in the near future.	16	118
AI Reference	Any reference to Artificial Intelligence, automation or agents.	6	8
Attitudes Towards The Subject Area	Any comments regarding the attitudes towards The Decentralised domain held by the general public	9	23
Big Data	Any reference to Big Data, process size, missed opportunity or other.	1	1
Blockchain	Any discussion around Blockchain or Distributed Ledger Tech	2	5
Business Models	Any reference to commercialisation or business models.	14	43
C Commons_ Terms_Machine Readable - Smart Contracts	Anything to do with agreements or mechanisms to deal with agreements, machine readable or smart contracts.	6	18
Cognitive Load	Any reference to the cognitive load, the required effort or problem solving.	3	5
Data Ownership	Any conversation regarding data ownership, individual, shared or 3rd party.	8	20
Definition Or Description	Any statement that describes a concept or an idea.	2	5
Gender Issues	Any comments relating to Gender.	1	1
General Challenges Or Problems	References to general challenges or Problems within the decentralised space or wider tech or privacy field.	14	114
General Technology Solution Or Approach		10	55
Indervidual_Understanding		1	1
Mental Model	Any reference to the users mental model.	12	26
Metaphor And Understanding	Any conversation referring to the understanding of the user or Metaphor.	15	42
Interface Description Or Suggestion	Any statement relating to an Interact with technology or User Interface	7	26
IOT Reference	Any reference to the Internet Of Things	1	1
Legal Law Enforcement	Any reference to state law, legal practice, enforcement….	1	1
Peer Or Node	Any reference to a peer or node based model or Interaction.	1	2
Privacy By Design	Any reference to Privacy By Design or decentralised design philosophy.	4	4
Privacy Issues	General conversation around Privacy in all its forms.	11	36
Privacy Paradox	Anything in and around the Privacy Paradox.	3	8
Probability Decentralisation Is Achievable	Any comments or views around the probability of a decentralised Internet becoming a reality.	7	12
Reference To Identity	Any reference to identity, digital identity or identity as part of the decentralised Internet.	5	16
Reference To Persona	Any reference to individual or multiple personas and their context…	1	1
Reference To Theory	Any reference to academic theory.	4	14
Requirements For Adoption	What is required to make decentralisation or SSI a reality in terms of adoption.	10	24
Scalability	Any comment or discussion considering scalability of a decentralise systems.	6	16
Seamless Transaction	Any reference to seamless transaction of data, and it's benefits across the board.	3	6
Semantic Technology	Any reference to the need for or implementation of a semantic layer.	1	2
Significate Statement	Any statement thats profound or has real impact.	16	136
SSI	Any reference or statement regarding Self Sovereign Identity.	5	13
Static And Dynamic Data	References to the types of data a decentralised internet may have to deal with.	4	8
Surveillance	Any commit or discussion around Surveillance.	1	1
Technology Challenges	Any technology issues that need further development or that might stand in the way of development or adoption.	12	29
Technology Progression Development And Trajectory	Statements regarding development and the direction of travel.	6	24
The Mechanics Of A Concept	Any indication as to how a technology mechanic may function.	6	24
Timeframe Towards Realisation	The time it's going to take for decentralised technology to become a reality, and any milestones along the way.	3	9
Trust Framework	Any mention of a Trust Framework or the Web Of Trust.	10	45
Usability Obstacles	Any clear thoughts on obstacles when it comes to usability.	11	32
Value Proposition	Identification of Value Proposition of any kind.	15	93
Visibility of Technology	Whether the technology is hidden or made visible to the user.	8	21
Zero Knowledge Proof	Any reference to Zero Knowledge Proof and related technology…	3	7

Appendix D - Conceptual Modelling

D1 Case Studies

Use Cases		
Title	**Description**	**Steps**
CONNECTION		
Establishing A Pairwise Identifier with an Individual in the field, or generate and send.	A means of generating a new DID pairing with another entity in the real world using a model device running the agent.	Log in Using Biometrics. Access DID Creation Object. Enter Metadata. Generate New DID. Share as QR Code or DID Address.
Establishing A Pairwise Identity through a Website	A method of establishing a Pairwise relationship with a Web Entity. This then can define a comms channel, the future interaction using this Pairing can be bound by restriction. This method would ocurre when a normal interaction would request a sign-up. This would be the beginning of Proof exchange bound by shared understanding of restrictions.	Enter Web Site. A point requiring data exchange is reached. Site Initiates Pairing and offers DID or QR. Users copies or scans and enters into agent DID Creation Object Agent initiates pairing.
Establish A Pairwise Relationship from a public DID	Instead of going instigating an exchanging with a Unique DID from a third party either through a web interaction of in the field. A well known public DID could form the starting point for relationship building with a established organization or institution.	Source DID. Users copies or scans and enters into agent DID Creation Object. Agent initiates pairing.
AUTHENTICATION		
Sign Into A Website With a Pairwise Relationship	As a means of replacing passwords a DID wise relationship with a Web entity can be used to verify identity through DID aurth. On entering a web site or application a request for ID can be made. The user inputs or conveys the DID through which the connect should be made. The 3rd party looks to the DID, confirms its on record, looks up DID end point and uses DID auth to request authentication. Authentication request is pushed the user's agent. User is notify and decides whether or not to confirm. Conform and there in.	Browse to Website / Application. Site requests login. Users Enters DID. 3rd Party requests authentication through DID Auth. User receives authentication request. User authenticates. User then has access to site.
Authenticate a Credit Card Transaction Online	As Above / User and Credit Card company have a reciprocating relationship through a DID for the Card In Question / If the Ard Is Used for an online transaction, the credit company requests Authentication in real time, via the agent before the transaction is complete.	User uses Credit Card or mobile Payment Application. As part of the payment process provided requests authentication. through DID relationship via DID Auth. User receives authentication request. User authenticates. Payment is processed.
Authenticate a Credit Card Transaction Within A Retail Environment	As above / Replacing a Pin Number or in addition to it authentication is used requested whenever the user issues a transaction in whatever form, with a physical card or Via QR banking application.	User uses Credit Card or mobile Payment Application. As part of the payment process provided requests authentication. through DID relationship via DID Auth. User receives authentication request. User authenticates. Payment is processed.
Authenticate Identity With A Physical Space	Via RFID, QR, or facial recognition individual declares entry into a physical space, current proof status can be assessed automatically or authentication request made in real time.	User transmits DID relative to location, is detected, or chooses to disclose. As part of the proximity / security process system requests authentication through DID relationship via DID Auth. User receives authentication request. User authenticates. Payment is processed.
Authenticate In The Field to Access a door	QR lock can be accessed as part of an application and DID, Proof payment process can be initialised before door is unlocked. If a prior agreement / access is inplace, locked door requests authentication in real time before opening.	Door locking application accessed. Door in question is scanned. Pairwise relationship to established. Proof request is made by door System. User compiles and sends Proof. Authentication is requested. User receives authentication request. User authenticates. The Door Opens.
Authenticating A Ticket At A Gate	Digital ticket is presented at a gate, system then requests Authentication in real time before individual is allowed to progress. AT this point addition proofs may be required.	User presents digital ticket through boarding application. As part of the security process system requests authentication through DID relationship via DID Auth. User receives authentication request. User authenticates. Please proceed.

SHARING DATA		
Establish relationship and terms of data use when visiting website	At the point when now a Website asks a user to accept cookies, the user will instead form a relationship with the Web enitity to authenticate trust, establish secure communications channel and argees terms by which data can be gathered, used, redistributed.	Web entity offers DID. User automatically pairs through agent specifying terms of connection. User requests Proofs. Web entity provides Proofs. If the user is happy Authentication login process is followed as above.
Prove your eligebility to hire a vechical, age, licence capbility to pay, additional insurances	As part of a booking process with a car hire company, the user is asked to provide a number of number of cridentchelas to prove they are eliganble for the hire terms, payment capability and additional ensurances.	Pairwise relationship is established and user accesses service. At a point in the process where cridentails are required request is sent by company for complex proof to user. User accesses agent and considers request. Proof Builder Autmatically populates with Zeros. User decides is Zeros are approprate. User specifies terms. If not reassign, If happy provide Proof. Service cheaks the Proofs. Transaction continues.
Supply a Number Of Proofs For An Employment Licence	As part of an application process with an authority a number of cridentaical are required.	As Above
Apply for Credit Providing proof of employment, address and income	As part of an application for Credit, via card, loan or overdraft, a number of cridentails are required.	As Above
Share realtime Data for an IOT health data device with medic	Data from an IOT health devive is to be shared with a Medical Practionaire, this requires the wiring and authentication of the data stream from the IT device and within this process terms of use need to be applied.	
Share Purchase History and Finiacal Position With An Intent Casting Application	Share past perchase history and evidance of ability to pay in order to recived the bast offers on good s and services from an intent casting service.	Provide a Data Wrapped claim
Provide A Claim For A Personal Reference	Respond to a request for a personal reference by generating a claim which may or maynot include proofs of standing and reputation.	
Provide a Claim To Allow Child On School Trip	Respond to request for permission from a School to allow child to go on a school trip through a standard claim.	
Data Gathering		
Request A Claim Of Education Record	Request a Claim for a Record of Past Education from a School or University.	
Establish a Repository For IOT Data	As part of the connection process establish a reposity for Data for an IOT device.	
Download and Redistribute Social Network Data	Download data generated from an exsisting relationship with a Social Network Provider and Redistribute when moving to another provider.	

D2 Task Analysis

Task Analysis			
Task Category	**Task Title**	**Description**	**Steps**
CONNECTIONS			
Connections Establishing DID's - Generation and Send All of these scenarios involve the generation and transmitting of DID Addresses. The did is then picked up by the opposite party via either direct address input or scanning QR. Once DID information is picked up by Agent, a reciprocating DID can be sent in response	**Generate new DID and Send**	User generates a new DID and assigns metadata to it so it can be recognised by the agent. The address is then passed to the external party ' via email, text, message' with metadata for basic identification by their agent.	Accesses Connections Object Access New DID creation Enters basic metadata for later identification in agent Generation of new DID Access to means of dissemination 'Address, QR' Dissemination elements can be copied and sent
	DID Pairing	User retrieves a 3rd Party Pairing DID, this may come in the form of a DID address generated by a website to allow pairing. It may be picked up in the Field Via a QR code. The DID is retrieved and metadata is associated before a Pairing response is sent via the DID's Public key.	Accesses Connections Object Access Pairing Pastes DID address or enters via QR Looks up the DID Adds Metadata for Agent identification Pairs
Needs reviewing	**Publically visible DID**	The user may have posted a publically visible DID in which case a request for pairing may come via a direct message, through the agent, and alerted within the System State Notifications.	
AUTHENTICATION			
	Pushed Authentication	A DID formed as a Pairwise relationship may be used to establish access in a number of forms. Once that DID is specified the reciprocating side needs to verify that the issuer has control over the private key. To do this a verification request is pushed back to the agent in real time. The agent notifies the user. The user accesses the device, checks the verification request and responds accordingly. Accessing Physical Location Using DID as Login Verification of Data Transaction Verification of Wealth Transaction	Receives Authentication request from existing pairing User can choose to Pend / Reject / Authenticate

	CONNECTIONS ***'Historys and Management'***			
	Accessing information about existing pairwise connections. Prerequisite for the following tasks.	**Accessing Pairwise Information and Interaction History**	Locating an existing Pairwise relationship. Selecting the Pair and Access History Log	Selects Connections. Filter Selection Type 'Individual, Organisation, Device' Locates the Relavent Connection Accesses minimum information Drills down into the detail Filters the Log Type Browse through the details Link off to required actions communicated in the log
	Subset Action once Pairwise connection is accessed.	**Accessing Pirewise connection details and subsequent actions**	Once a Pairwise connection is selected further detailed information and action can be accesed. At this point this offers a very limited amount of management over the relationship. Refresh Revoke	Cog Icon is selected Detailed info opens Option to Update or Revoke is avalible
	Subset Action once Pair-Wise connection is accessed.	**Sending direct Message to Pairwise DID**	Once a Pairwise connection is selected the user can send that entity a direct message. At this stage that letter is ASCII and not machine readable.	Selects New Selects Message Opens Message Dialog Types Messages Cancels or Sends Message
	Subset Action once Pair-Wise connection is accessed.	**Requesting a Claim**	Once a Pairwise connection is selected the user can choose to request a claim. The claims available and offered by the Pairwise relationship will be listed. The user Selects the Clam they require and makes the request.	Selects New Selects Claim Opens Claim Dialog Selects from list of available Claims Makes Claim Request
	Subset Action once Pair-Wise connection is accessed.	**Requesting a Proof**	Once a Pairwise connection is selected the user can choose to request a Proof. The Proofs available and offered by the Pairwise relationship will be listed. The user Selects the Clam they require and makes the request.	Selects New Selects Proof Opens Proof Dialog Selects from list of available Proofs Makes Proof Request
	Subset Action once Pair-Wise connection is accessed.	**Data Binding to DID**	Once a PAIR_WISE connection is selected the user can choose to allocate a data repository to it or Bind data.	Selects New Selects Data 'If no repository exists' Opens data Binding Dialog User specifies data type 'IPFS ?' User specifies conditions / restrictions User generates Binding
	Subset Action once Pair-Wise connection is accessed.	**Claim Request**	Once a Pairwise connection is selected the user can choose to request a claim from that relationship.	Selects New Selects Claim Opens Claim Dialog A list of Claims offered by the third party Select Claim Request Or Cancel
	Subset Action once Pair-Wise connection is accessed.	**Proof Request**	Once a Pairwise connection is selected the user can choose to request a Proof from that relationship.	Selects New Selects Proof Opens Proof Dialog A list of Proofs offered by the third party Select Proofs Request Or Cancel
	Subset Action once Pair-Wise connection is accessed.	**Claim Send**	Once a Pairwise connection is selected the user can choose to send a Proof to that relationship.	Selects New Selects Claim Send Opens Claim Send Dialog A list of Claim templates available Select Claim Populate Claim Send Or Cancel
	Subset Action once Pair-Wise connection is accessed.	**Direct Message**	Once a Pairwise connection is selected the user can choose to send a Direct message.	Selects New Selects Direct Message Opens Direct Message Dialog Populate Message Send Or Cancel

D3 Objects and Operations Analysis

Object / Operations Analysis			
Objects	**Description**	**Attributes**	**Operations / Actions**
SSI AS A CONCEPT			
Core Identity	At the core of a SSI system sits the capability to independently control, define and maintain a digital identity. An individual may hold more the one identity, the identity my belong to a business, organisation, authority or an object. The identity will be linked to other forms of information and identities, raw data, claims, contracts and wealth. The core identity is anonymous. The authenticity of an identity can be seeded by institutional claims from which trust can cascade or might be developed over time as interactions building trust and reputation. The relationships are collections of pairwise sudonosum relationships, these relationships are based on trust genertaed through verifiable claims that can be linked to static or dynamic data. Realtionship and cliams are managened through a Wallet or Agent, that manages the components of identity, and provate and public keys..		
IDENTITY MANAGEMENT TOOLS 'Constituting a Wallet'			
System State 'Notifications / Activity / Data Flow'	A central object or dashboard from which other objects will be managed. The area will deal with notifications and warnings, and act as a jumping off point to other areas of interaction. It will be a place to discover identity, services and applications. It will be a place to understand more about the position of an individual's identity.	Visualize all incoming warnings / communications / requests / and data flows. View notices / concerning activity / outdated or required renewal agreements, proofs, claims. View communication notices. View data requests incoming & outgoing and gain basic information.	Manage SSI Configurations. Review of communications / notifications. Link to action notifications. Notification management. Link to other manager objects.
ADMIN AND AUTH			
Administration	An Object from which general administration configuration can be accessed.	Parameters for the following: Notifications / Back-ups / Biometrics / Passcode / Notifications / Key Recovery / Dynamic Data Storage / Version Information.	Capability to select from the options and drill down into details and configuration functionality.
Push Authentication	An Object from which pushed requires for authentication from relationships can be accessed and actioned.	The header defines the connection from which the requested authentication originates. Basic authentication request title and details. The connection name is listed together with DID information. The DID triggering authentication is listed. The Time and Data of the authentication request highlighted.	Capability to: Pend, Reject, Authenticate.
CONNECTION 'Creation and Distribution'			
DID Creation	This Object allows the creation of a new DID. At this point metadata can be associated with it so it can be recognised and managed by the agent application.	Name of the connection once entered. Name of the reference once entered.	Capability to: Enter metadata, both connection name and more detailed reference information. To cancel the process. To generate the connection.
DID Distribution	This Object allows for the external distribution of a newly created DID's They can be copied and pasted into email or transferred via a Generated QR code.	DID metadata, name and reference. DID Address. DID Address as QR Code. DID Doc Info Methods / End Point / Timestamp / Signature. Possibly shielded Private Key.	Capability to: Copy Address. Copy QR Code. Show Private Key. Refresh. Revoke.
CONNECTION 'Incoming and Pairing'			
Inputing DID 'By String'	This Object handles the Inputting of initial, or introductory DID's. This may have been generated as a web introduction or signup. Directly cut and pasted, as part of a notification to a publically available DID, or scanned from a QR code. This object has the capability to jump to a lookup.	DID address once pasted. QR, once scaned and activated.	Capability to: Enter DID Address. Scan QR. Activate Lookup.
Review DID Doc and Create Pairing	Once a DID is entered into the Inputting DID String Object it looked up on the Block and the DID doc data is displayed. At this point there is the option to pair with this DID to create a Pairwise Relationship.	Metadata, name and detailed reference.	Capability to: Add metadata reference for agent and idenitfication. Option to Cancel. Option to Pair.

CONNECTION 'Established Connections Management'			
Listing Of Established Connections	Within this Object ALL existing connections manifesting as DID's can be viewed filtered and accessed. They can be scrolled and opened for limited information. An indication of Action Required can be seen, and if the connect has a associated data repository, this is also communicated via icon.	The page lists connections established through DID's. These can be listed by a number of categories All / Indervidual / Orginisation / Device / Pending and New. The elements can be opened to reveal limited information. Limited information will draw attention to any actions that are needed.	Capability to: Limit view of connections through filters. Scroll up and down through the Connections. Alphabetically jump through type. Open connection up with limited information. Select to open view and move to the connection history page. The New connection button will move the user to the New connection options. Possible search capability.
History of Interactions	Selecting an existing pair-wise connection in the Established Connections Object, opens the History of Interactions Object. This displays a log of all interactions with this DID pairing. Notifications / Requests / Transactions / Messages. If any action is required this is highlighted. Within this Object there is the functionality to generate Direct Messages / Request Claim / Request Proof / Send Claim / Data Binding. This Object can be consider together with the Notification Object to be a Core Hub Object.	The page lists all recorded interactions between established incoming and or pairwise DIDS. These include Notification / Communications / Requests / Transactions. Each inclusion lists the date and type and indicates if action is required.	Capability to: Limit view of interaction records through filters. Scroll up and down through the interaction records. Alphabetically jump through type. Select to open view to move to the connection history page. The New connection button will move the user to the New connection options. Selecting the COG Icon to be taken to the details of the connection.
Detailed Pairwise DID Information	Within this Object the details of the Paired DID's can be reviewed in detail. The DID document on both sides. Auth Methods / Endpoints / Signature..... There will be an option to Refresh and Revoke.	The page offers details to the signal or Pairwise DID connection. Opening Incoming or Outgoing panels shows the information held within the DID Doc. 'DID Address, Public Key, Methods, Endpoints, Timestamp, Signature.	Capability to: Refresh both incoming and outcoming. Revoke incoming or outgoing.
History Log Detail	If a specific element in a pairing history is selected a detailed drill down of this interaction can be understood within this Object.	Detailed information regarding the selected Interaction.	Capability to: Move back to the full list of interaction linked to the selected Connection.
CONNECTION SECTION 'Functions'			
Direct Message	This Object is selected from a New option within the History of Interactions Object, this function is relative to the selected pairing. From here a direct text based message can be sent. At this stage that letter is ASCII and not machine readable. At this stage this a very basic and will no doubt evolve as the system develops.	The only feedback to the user is the simple message being written.	Capability to: The capability to cancel the message. The capability to send the message.
Claim Request	This Object is selected from the New option within the History of Interactions this function is linked to the selected pairing. The Predetermined Claims available and offered by the pairwise relationship will be listed. The user can then select the Clam they require and make the request.	When a Claim request is made to an existing connection the avalible types of claims are listed. Details about the claims offered can be opened and inspected.	Capability to: Selected the desired claim from a list. Activate the request action. Cancel the request.
Proof Request	This Object is selected from the New option within the History of Interactions this function is linked to the selected pairing. The Proofs available and offered by the pairwise relationship will be listed. The user can then select the Proofs they require and make the request.	When a Proof request is made to an existing connection the avalible types of Proofs are listed. Details about the Proofs offered can be opened and inspected.	Capability to: Selected the desired proof from the list Activate the request action. Cancel the request.
Claim Send	This Object is selected from the New option within the History of Interactions this function is linked to the selected pairing. The Claims available and will be listed. The user can then select the the Claim they require and make the request. ***There needs to be an area to populate the CLAIM***	Entered Search Term. When a search is made existing predifined Claim types are listed.	Capability to: Trigger a search of prefined Claims. Selected the Claim. Activate the send Claim action. Cancel the panel.
PROOF BUILDING			
Auto Population and Confirmation	This Object is launched from a Proof Request. Once opened the details of the Proof Requests are visible. The Zeros are auto Populated and can be seen. Within in each element there is the option to select an alternative.	The header defines the requesting contact and the Proof request type. The elements of Proof required are listed. Drilling down specifies auto populated Zeros from Existing Claims. Drill down offers the prospect of alternatives. Conditions offers the Conditions which can be placed on the Proof.	Capability to: Drill down into the detail of the proof claim zeros. To activate alternative. Select the conditions Build the proof.
Selecting Alternative Zeros	Within a Proof Zero Element, there is the option to select at alternative if the auto selected is not deemed suitable. Within this Object a new Claim can be selected from which then Zero can be selected.	A list of Claims which may contain zeros as alternative source of proof.	Capability to: Select Claim. Use Proof Cancel

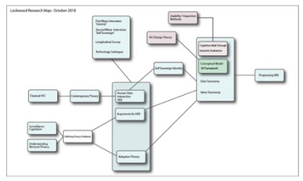
DATA MANAGEMENT ***'Proof / Claim'***			
Sent Proof Listing	Within this Object ALL existing sent Proofs can be viewed filtered and accessed. They can be scrolled and opened for limited information. The State of the Proof is Indicated. Live / Pending / Revoked. Selecting the Proof opens the Sent Proof Dig Down Object.	Sent Proofs are listing alphabetically. Proof details can be opened displaying minimal information. The state of the proof is communicated Live / Pending / Revoked.	Capability to: Double click proof to reveal further details.
Sent Proof Dig Down	Within this Object details of the Proof can be understood. Dependent on the State the Proof can be Sent / Amended / Revoked.	Detailed information about sent proof. Information regarding proof state.	Capability to: 'Depending on Proof State' Review / Amend. Revoke. Submit. Review Archive.
Received Proofs	Within this Object ALL existing Received Proofs can be viewed filtered and accessed. They can be scrolled and opened for limited information. The State of the Proof is Indicated. Live / Expired / Revoked. Selecting the Proof opens the Received Proof Dig Down Object.	Received Proofs are listing alphabetically. Proof details can be opened displaying minimal information. The state of the proof is communicated Live / Expired / Revoked.	Capability to: Double click proof to reveal further details.
Received Proofs Dig Down	Within this Object details of the Proof can be understood. From here there may be an option to Dig down further into the real details of the Proof and the Zero Elements	Detailed information about received proof. The state of the proof is communicated Live / Expired / Revoked.	Capability to: Look at even greater depth.
Received Claims	Within this Object ALL existing Received Claims can be viewed filtered and accessed. They can be scrolled and opened for limited information. The State of the Claim is Indicated. Live / Pending / Revoked / Expired / Rejected Selecting the Claim opens the Received Claim Dig Down Object.	Received Claims are listing alphabetically. C	Capability to: Selection Claim to reveal further details.
Received Claims Dig Down	Within this Object details of the Claim can be understood. Depending on the State the Claim can be either Accepted Or Rejected.	Detailed information regarding claim. Issued data, Claim type, Issued By, When Offered and Accepted Breif description. Detailed claims data.	Capability to: Depending on State: Accept / Reject / Refresh.
Sent Claims	Within this Object ALL existing Sent Claims can be viewed filtered and accessed. They can be scrolled and opened for limited information. The State of the Claim is Indicated. Selecting the Claim opens the Sent Claim Dig Down Object.	Sent Claims are listing alphabetically. Claim details can be opened displaying minimal information. The state of the Claim is communicated Live / Pending / Revoked / Expired / Rejected. Indication is made as to whether the claim has associated data	Capability to: Double click proof to reveal further details.
Sent Claim Dig Down	Within this Object details of the Claim can be understood. Depending on the State the Claim can be either Updated / Amended / Revoked	Sent Claims are listing alphabetically. Claim details can be opened displaying minimal information. The state of the Claim is communicated Live / Pending / Revoked / Expired / Rejected. Indication is made as to whether the claim has associated data	Capability to: Update / Amend / Revoke
Dynamic Data			
Data Binding	This Object is selected from the New option within the History of Interactions Object linked to the selected pairing. Once a Pairwise connection is selected the user can choose to allocate a data repository to it or Bind Data.	Name of DID to link data repository is to be linked. Capability to dig further into the DID document. Repository name once entered. Conditions on which the data can be used.	Capability to: Dig down into DID document. Enter Meta Label for Data Repository. Define Condition of data use. Generate / Cancel
Existing Data Binding Listings	Within this Object ALL existing Data Bindings can be viewed filtered and accessed. They can be scrolled and opened for limited information. The State of the Claim is Indicated. Selecting the Claim opens the Sent Claim Dig Down Object.	Existing Data Bindings Are listed Alphabetically Details can be opened displaying minimal information. The state of the Binding is communicated - Live / Suspended / Revoked.	Capability to: Open for further details. Select to move to a Dig Down view
Data Binding Dig Down	Details of a Data Binding can be accessed from this Object. From here the Data Binding can be Downloaded 'The Raw Data' / Suspended / Revoked.	Detailed information regarding date established, offered, accepted / category / data type. A basic description. Data Statistics	Capability to: Download raw data. Revoke

Introduction

The management of our digital identities and personal data represents a formidable challenge for the 21st century; issues of privacy, inference profiling, surveillance capitalism, GAFA monopoly, democratic interference and the lost opportunities of Big Data are significate. Many envisage a decentralized alternative to the centralized network, one that places the human at the center of data streams, facilitating transparency, agency and negotiated access to personal data. Self-Sovereign-Identity is a collection of concepts and standards that promise to emancipate the everyday user from the A-Symmetric relationships currently observed across todays Internet. This will be achieved by establishing and controlling persistent digital identity's, exchanging digital credentials, building trust networks, while managing access to sovereign data repository's .

The decentralized nature of the technology and the subsequent responsibility placed on the user to manage and protect what is a decentralized key management system, pose interesting interaction design challenges. This paradigm shift introduces new concepts and and ideas, raises issues of missing metaphor and mental models. As these concepts move from theory towards tangible artifact, it is important to shift the thinking anyway from a purely technological problem space, towards a that of a user centric design prospective. If SSI is to become accessible and sustainable, we must investigate and understand SSI applications, case-studies, interactions, and identify the foundation design patterns needed to manage and transact personal data through digital identity.

This ongoing work highlights efforts being undertaken to develop a conceptual model for SSI and subsequent preliminary UI. This work provides a framework for the design community to better understand, explore and investigate user tasks and interactions for SSI.

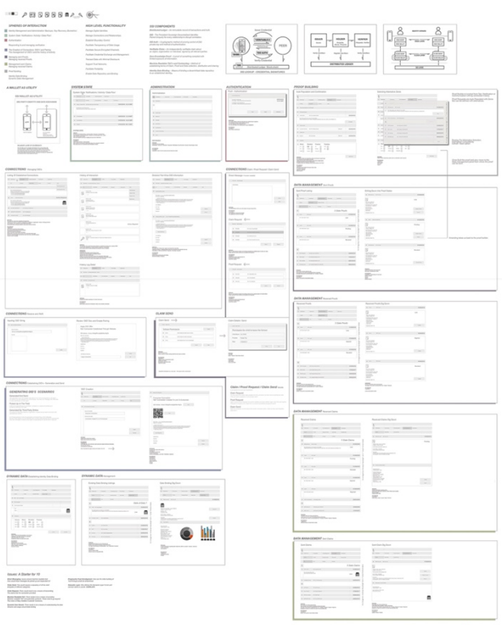


This work forms part of a wider study into the decentrilised internet. The focus on a conceptual model for SSI and subsequent UI development, sits is towards the end of a study which takes into account the theory's of Privacy, Surveillance Capitalism, Adoption Theory, Classical and Contemporary Human Computer Interaction and Public Option. The study hopes to make a contribution to the field of Human Data Interaction, while moving the focus of the Decentralised Internet community towards design, interaction, adoption and sustainability issues.

Wireframe

The following Wire-Frame is a first pass UI following the conceptual model. This is very complex, and there is no claim that this is the optimum solution. It is though a starting point, and represents an interpretation of a minimum traditional interface requirement for the principles of SSI using the identified components.

Download



Nick Lockwood
01-01-2019
n.g.lockwood@salford.ac.uk

Self Sovereign Identity

Self-Sovereign Identity solves one of the most challenging issues of the Internet. The capability to establish, own and control a persistent sovereign identity online. SSI can be defined as the following: A Digital Identity that is owned and controlled by an individual, company or machine, it has no reliance on any centralized authority. The identity is persistent and can never be taken away from its owner. The identity is part of a wider eco-system, where relationships can be built, trust can be developed, identity attributes, and data, can be exchanged under the complete control of the Sovereign Identity. The concept of a wider eco-system is important, once a user can control an identity, the concepts of SSI then allows the user to establish independent unique relationships and communication channels with peers on the network, it allows for the requesting, issuing and distribution of verifiable credentials. This in turn empowers the user to control their personal data and information in a transparent manner. The eco-system allows for the development of trust networks that are judged appropriate dependent on the situation. The realisation of SSI is considered to be Web 3.0, and to many is inevitable, (Tobin, A, 2016) Web 3.0 represents the next evolution of the Internet. Where the common user is no longer reliant on a loaned or leased identity, owned and controlled by a third party or centralized authority. The user is no longer bound by contracts of adhesion, data exploitation, and resulting A-Symmetric power imbalances.

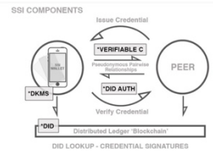
The route principles of SSI can be traced to the seminal work of Kim Cameron in his 'The Laws of Identity' (Cameron, K, 2005). These principles of digital identity where then evolved further in the context of SSI by Christopher Allen (Allen, C, 2016).

For an online identity to be self-sovereign, it should satisfy the following principles:

- **Existence.** Users must have an independent existence.
- **Control.** Users must control their identities.
- **Access.** Users must have access to their own data.
- **Transparency.** Systems and algorithms must be transparent.
- **Persistence.** Identities must be long-lived.
- **Portability.** Information and services about identity must be transportable.
- **Interoperability.** Identities should be as widely usable as possible
- **Consent.** Users must agree to the use of their identity.
- **Minimalization.** Disclosure of claims must be minimized.
- **Protection.** The rights of users must be protected.

Building Blocks for Self-Sovereign Identity

Below we discuss the fundamental building blocks of SSI in order to understand the concepts, component parts and interactions required. The following graphic illustrates a balanced peer to peer relationship between two entities across the network, and the exchange of data and credential, supported by distributed ledger based verification and the current developing standards.



Distributed Ledger Technology

Recent developments in Self Sovereign Identity have been made possible by the advent of Blockchain Technology and its capability to provide a public, immutable, distributed ledger of transactions between peers. In most cases the infrastructure is designed as a public utility. And acts as a decentralized foundation, supporting identity development, authentication, verifiable credentials, trust and data management. (Wagner, 2018). The system uses encryption and digital signatures to establish a decentralized Public Key Infrastructure, through which identity, communications, credential, trust and authentication can be managed.

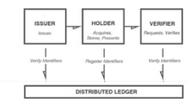
DID's 'Decentralized Identifiers'

Decentralized Identifiers 'DIDs' are digitally verifiable identity's that can be created by an individual, organisation or object. DIDs are fully under the control of the DID subject and are independent from any centralised authority, identity provider, or certificate authority. A DID is a Key Value Pair. A DID address can be looked up on a blockchain and returns a DID document which contains information about the DID. That information is composed of Proof Purposes, Verification Methods, A Public Key and Service End Points. Through the public key the identity holder can establish a private communication channel, prove they control the identity and register, resolve, update or revoke details for the DID on the hosting blockchain. There is no limit to the number of DID's that can be created and controlled, this capability enables Privacy By Design principles, preventing correlation and tracking. A principle of SSI, promotes the concept of Synonymous Pairwise Identifiers, which means that for every peer to peer relation across the network a unique pair of DID's are created, allowing authentication, communication and transaction.

In the design of any Self Sovereign Identity system creating, sharing and managing DID is a primary Sphere of Interaction.

Verifiable Credentials

We use verifiable credentials and subsequent claims in our everyday life's, a driving license, educational certificates, a password, are all verifiable credentials that we can used in a spectrum of contexts and combinations to make claims about ourselves. A verifiable credential in the context of SSI is a digitally signed document that is produced by an issuer and accepted by a subject or holder. Verification signatures proving the integrity of the document, its issuers identifier and the acceptance of the credential by the holder are recorded on the blockchain. The holder of the credential can then disclose the credential to another peer in full, or in part, and this third party or verifier can confirm the integrity of the credential by looking it up on the blockchain, confirming origin, signature and integrity. This in turn allows identities to prove claims about themselves to develop webs of trust.



Credentials can be shared in part through the concept of Zero Knowledge proofs, which allows the holder to make a claim about an identity without revealing the whole of the credential. Credentials or claims can be created and sent by any identity holder, it might be a government, organization, company, individual or object. The level of trust required will varied dependent on the circumstance. Credentials and Claims can be combined in different circumstances, they can be static or dynamic in nature and they can be rejected or revoked at any time. In the design of any Self Sovereign Identity system creating, sharing and managing verifiable credentials, derived claims and zero knowledge proofs is a primary Sphere of Interaction.

Verifiable Claims Working Groups https://www.w3.org/2017/vc/WG/

DID-Auth 'Decentralized Identifier Authentication'

DID Auth as a method by which the owner of an identity, through a wallet or agent, can prove to a relying party that they are in control of a decentralized identifier 'DID' The demonstration of control of the DID is accomplished through a mechanism specified in the DID Document. Depending on context this authentication might be done while establishing a connection, as part of a transaction or as part of a request. The notion of proving control over the private keys associated with any given DID is an integral part of a SSI system and various concisions authentication interactions.

Web Of Trust Info – DID Auth https://github.com/WebOfTrustInfo/rwot6-santabarbara/blob/master/final-documents/did-auth.md

DKMS 'Decentralized Key Management Structure'

Self-Sovereign-Identity employs the power of distributed ledger technology to provide security and immutability to provide scalable key distribution, verification methods and key recovery. This allows for a personal decentralized public key infrastructure free from any centralised authority or control and makes the benefit of a public key infrastructure available to everybody. As individuals engage in the practice of SSI they will be generating a considerable number of encryption keys that they will have to take responsibility for and manage. When a user manages their relationships, credentials and data through SSI they will do so with a wallet on the edge of the network, this will typically be on a mobile device; most keys and credentials will be held on that device. The wallet is complemented by an agent which will automatically management many of the automated mechanisms needed to engage with SSI. For the identity to be available 24/7 and to also manage backup and key recovery, a limited representation of the edge level Wallet / Agent is available.

This allows for a personal decentralized public key infrastructure free from any centralised authority or control and makes the benefit of a public key infrastructure available to everybody. As individuals engage in the practice of SSI they will be generating a considerable number of encryption keys that they will have to take responsibility for and manage. When a user manages their relationships, credentials and data through SSI they will do so with a wallet on the edge of the network, this will typically be on a mobile device; most keys and credentials will be held on that device. The wallet is complemented by an agent which will automatically management many of the automated mechanisms needed to engage with SSI. For the identity to be available 24/7 and to also manage backup and key recovery, a limited representation of the edge level Wallet / Agent is available.

http://bit.ly/dkmsv3

DKMS is an Emerging open standard that deals with the interoperability of both agents and wallets at both the cloud and edge layers. Within in this work we are focused on the interactions and capabilities of the user's edge wallet though the agent and cloud layer confirmation may for part of the configuration interaction within the edge layer wallet UI.

D5 Focus Group Planning

SSI - WORKSHOP PLANNING - March 2018

Agenda

13.00 - 13.30 I'll share some of my work to date.
13.30 - 14.00 General discussion around data, privacy, the decentralized internet and SSI.
14.00 - 14.15 Coffee.
14.15 - 15.15 Focus Groups around questions of complexity and value proposition.
15.15 - 15.30 General Round up discussion.
15.30 Quick half in the Dock Yard.

Discussion Topics 13.30 - 14.00

- Privacy, personal data and the problem space.
- The general progression of ideas leading to SSI.
- The understanding of the concepts and components.
- Notion of separation of SSI as a boundary interaction.
- The complexity of interactions.
- Cognitive loads and mental models.
- Balancing Value Proposition against Cognitive Load.
- Top-down, bottom up adoption across the world.

Focus Group 14.15 - 16.15

Putting This Technology In The Hands Of The Masses from a Task and UI Prospective.

- How understandable are the concepts?
- What are your views on the complexity of these interactions?
- What are your views on the UI in front of you?
- How do we build something that's assessable?
- Any thoughts on the mental models, and metaphors?
- Can could any of this be automated? SEP
-

Where is the value proposition that may Driven Adoption?

- Balancing Value Proposition. against the Cognitive Load?
- Where's the value?
- Views of privacy around the World and across cultures!!!
- Bottom up – Top Down?
- How much of the complexity does the user need to see?
- Trust frameworks?
- Opportunity for business...
- Compliance.....Quality data, society and Big Data...

D6 Focus Group Participants

Focus Group Participants
Egg Space – University of Salford – Media City - 15th March 2019

Attendees:
Paul Wilshaw
Head of Innovation / Barclays UK
https://www.linkedin.com/in/paulwilshaw

Ian Forester
Senior Producer at BBC R&D / Founder DataPortability.org
https://www.linkedin.com/in/cubicgarden/

Suzanne Clark
Senior UX Designer BBC
https://www.linkedin.com/in/suzanne-clarke-31b720b2

Tom Cheeseman
Applied Futurist
https://www.linkedin.com/in/tomcheesewright

Jasmine Cox
Product Designer BBC
https://www.linkedin.com/in/jasmine-cox-b07663a3

Le T Hoai
MA Digital Experience Design at Hyper Island
https://www.linkedin.com/in/hoai-le-56565760

Marc Tobia
MA Digital Experience Design at Hyper Island
https://www.linkedin.com/in/marctobia

Tim Panton
Co-founder and CTO at Pipe
https://www.linkedin.com/in/timpanton

Lisa Ortega
UX Expert and Digital Psychologist at Keep It Usable /
https://www.linkedin.com/in/lisaduddington

Nathen Broadbent
Head of Creative Technology at Trunk. Agency
https://www.linkedin.com/in/nathanbroadbent/

D7 Focus Group Transcript 'Extract'

Focus Group
Egg Space – University of Salford – Media City
15th March 2019

Extract 01 00.00 -

Paul Wilshaw: 00.00
You know, you tell them you're wearing a watch it it gives off, you know millions of bits of data every hour they go seriously and they don't understand that complexity of what that data is or what that data is doing or what that data is for.

Lisa Ortega: 00.20
It's about the word data as well but if you say its information about you might get a different response when you talk about, I think the word goes over their head, there's nothing interesting about me. There's nothing interesting to see me in my data so, it doesn't exactly...

Paul Wilshaw: 00.38
yeah and un-be-knowing to you know, the the three thumbs up on Facebook people don't know that that's the engine behind it and they don't care, you know, it's a balance and also I think a lot of these interfaces what the problem with this is that because it's an interface on a device but a manufacturer that manufacturer that device has information from here, so you cannot control that day to yourself because it's unless you can make your own Hardware. unless you have your own need your own OS. Yeah, it's very difficult to control, because you know Apple every time I'm on something even if I use Chrome on my Apple device, it still knows where I'm going what I'm doing and has a paging file millions and millions of pages long, and there's a good story out there that someone's experience on Tinder and they did an experiment for a month on Tinder and then ask for the data back on it and they had but an 800 page document that from Tinder and every single conversation. They had every single swipe and it said they knew them better than they did themselves and it's yeah so imagine wait Mick what you were saying earlier if you have to authorize can like, you know, some people putting on school trip or things like that. That becomes a massive overhead for those kind of people in those.

Lisa Ortega: 02.18
So how about this element of control as well, so I know you said this is about control giving people back that control. But what about people that need help to understand what the best choice should be for them? Yes is this going to just overwhelm people and they're not going to end up making the right decisions really, you know, I know in the presentation you said about not defaulting options anymore, but sometimes that can actually be of benefit to people who default them to something that's in their best interest.

Mick: 02.46
I thought a lot about the contracting part of this, so once you his former relationship with somebody and you generate a data repository, you have to specify how that data can then be used and these contracts, you know, we're remembering our session that we ran when we were at my data we talked, there was a guy that came in and talked about contracts, and people, who deals with contracts you might we'll do the contract when you get a job and buy a house and then you never do with the contract at any other time in your life. And he was explaining the contracts aren't necessarily what you agree to do. But what happens if you don't do what you're expected to do, you know, and and then and then we moved on to the idea that contracts can be developed by community so that you know, you have your Champions and individuals that you trust that that specifying people have a library of contracts that they can use, you know, so instead of making all the decisions yourself. There's also almost a community that you sign up to [00:03:46] of or follow and they guide you through what you should be doing, which is

Ian 03.52
This is one of the things that we talked about was more of a distributed model because right now a lot of this is it's kind of like completely decentralized or emphasis on you and actually, you know maybe I'm time poor, maybe I

don't understand the language more do whatever reason happy to let others do some of that work for me. And so that could be an established organization like the BBC maybe or it could be the niece or nephew, or it could be a cousin you to make certain decisions rather than it always be on you. It's a bit like when you said about how do you kind of regain control of your identity, you go to these different people rather than it be all you're all nothing which unfortunately a lot of the SSI stuff. I've seen I've come from like a very American kind of you own everything because you don't trust anybody and it's kind of like.

Tim: 05.00
I was going to ask does anyone know what the relationship is between the kind of self-sovereign? Thinking and the right-wing militia groups who also call themselves Sovereign are they just conceptually the same background or or is that just like they read the same books or what?

Mick: 05.23
It's not all US though, half of the top people in Sovrin are UK based though they might have been from America.

Paul: 05.34
I think we in this room can acknowledge that there is an issue with identify that probably needs solving at some point. But after watching some of the videos you sent over, I think industry is pretty bad at articulating what that problem actually is. And the point about privacy verses control and I think that most consumers don't understand it in any real sense yet.

Tom: 05.58
If you look at the actual behavioural change post Cambridge Aynalitica generality. Yeah, almost no, behaviour change. There was a long slow decline in Trust of the Facebook's and the Google's from Generation by generation, but it doesn't feel like it's some sort of Crisis probably they're linked to a like this decision. It's like yeah, I'm a great example. I know I shouldn't store a bunch of my passwords in Chrome. I store a bunch of my passwords in chrome, because it comes back to that balance you put up there of cognitive load versus value proposition probably should, like I can't be arsed with the cognitive load of not doing it, so the value propositions need to be really strong.

Tim: 06.36
Chrome's really interesting in that respect because it's the level of trust that you can put in trim Chrome is to do with the fact that it is an open source project and if it is inspected by people who you might or might not trust it's not just Google there's another leg on it.

Tom: 06.57
But who things that yeah

Tim: 06.59
They do. Right, right.

Tim: 07.06
An they've also had this thing about like how many questions can you ask a user before? They just say? Yes all the time like the the you know, am I going to let you use the microphone question should we play auto play videos? All of those sorts of things they spend ludicrous amounts of time trying to get the right defaults for you without asking you the question and then fixing it if you Grumble like

Tom: 07.31
I review every single app my kids and store for permissions. I'm getting very close to just saying no more time, rather than just saying yes or time is it just can't be arsed reviewing collisions all the time.

Ian: 07.42
This is the GDPR Thing right, on the face of it is fantastic. But then what they've done is they've gone right? Let's reveal all the cookies. So you have to go through and go. Yes, no, no, no, no, no, no, no see made it harder. So the actual the spirit of it is that yes you have , they get about okay, we'll give you complete control so that you just feel so overwhelmed that you just go. Yes, whatever.

Paul: 08.15
It becomes more of an annoyance. Yeah.

Ian: 08.21
It deliberate though and that the thing.

Paul: 08.22
I Noticed a lot of that GDPR our stuff has made websites unusable. On devices.

Le T Hoai: 08.35
I think at some point sharing like people sharing their personal data, they kind of take it at a stranded for example, put it at the bank buying context or like flight ticket booking stand. They take it as granted that they have to share this kind of information and then for me taking control of data is not an absolute concept. Because for example at some point the user they can give you the fact information. For example, they need your their email to access to Wi-Fi and I can just create the really fake information to be able to do that. And then at some point for the company, I don't think that they take control all the data because if they can generate the personal like content they need to have enough personal information to be able to do that and then to be able to do that. They actually have to come by all the data that they have from personal, that user, and then they have to come by related information from personal from other sources. So basic actually, they also need to do the next step to actually filter the data to make it relevant. So basically But we like consider quality data of the company that they had from a news. I don't think that's all of them is like usable for the company to do that. So for me control is not an absolute concept for both side user and the company.

Tim: 10.19
There's a really nice example of that, the New York Times in Europe ran ran without the decided their way to deal with without with gdpr was just not to personalize adverts, just Geo Target them and that's it, and they got better results. Their advertising sales went up. So it's like the whole thing is completely pointless. I mean, you have to be a brand like New York Times for it to work and with a particular demographic and whatever but but I think a lot of basic what you're saying is that it doesn't work and it doesn't.

Tom: 10.59
I was chatting to be the chief data scientist for part of Salesforce for Commerce cloud, and he was saying I don't want you to shoe size or what football team support what newspaper with absolute so they can do 90% of the personalization just basically last ten clicks anything beyond that is under the regulations link GDPR. It's just eating without them. It's just it is an expense rather than value to.

Ian: 11.24
It's a head ache.

Tom: 11.24
It's a headache its massive head idea.

Mick: 11.28
Can I ask these specific questions? Yes, and then just get some feedback on whatever people think. So this is putting this technology in the hands of the masses from a task in the UI perspective, how understandable are the concepts.

All: 11.43
Lots of laughing.

Lisa: 11.45
Do you need us to answer that.

Ian: 11.48
So I personally understand them only because I've done some research in that. I've also I also own a number of digital wallets. So, I understand the concept also and stand private and public key encryption. I still, I was trying to

find the example with said Anna some of this is like where there's like a there's a thing that the stores talk about when into a public and private key encryption try to find that but yeah, I just don't think that people will understand that.

Mick: 12.29
But in terms of, if you where to talk to a normal person when they come across any of this before not necessarily the public-private key, but the concept of a wallet and close relationships, individual relationships and Gathering claims and then re Distributing them. Is that makes any sense?

All: 12.47
No,
Jasmine Cox: 12.53
Not in the digital world or the real world? I don't even talk about having a relationship with the corner shop and buy a paper or the airline where buy a ticket or Facebook who hosts my social Network. I don't have relationships and establish those at all don't kind [00:13:11] of take personal information and kind of plant access to get get access to the system. So yeah, I think I think this vocabulary needs, changing and I think I think the underlying concept of having a verified identity from somebody you trust and being able to prove your identity using it is completely sound, but that's as far as it goes. Don't think I need to know anything else.

Mick: 13.42
Any other thoughts about concepts in general.

Paul: 13.44
I agree with that because well, so I worked on Barclays Pingit back in 2012, and that was that was hard to Market because of the concept of a digital wallet people didn't get it and I yeah, I remember going backwards and forwards of the marketing department. We did this first video with a giant rugby ball with screens flashing on and off. Nobody got it. Nobody got it and nobody really got at that time as well. You could transfer money Via mobile phone number and I'll be it was an alien concept probably the best part of like three years for the concept get out and then like other other brands doing it. And yeah, it's really tricky for consumers to grasp even care like the concept of your mobile number can do something ring a person.

Tim: 14.49
We've been tackling this but in a very niche way and it is possible to get these messages across if you put them into particular concept, context, so to the workplace, we found really works is things like security cameras, baby monitors where people do understand that their image the image of their kids is something that they want to protect and they want to keep within a close to user group and they want to manage that in a way that's understandable. And so where we've done, it's on Tiny fraction of what you're talking about here, but the where we've done that, and we're it's resonated is that but it's about ownership and it's about my family relationships. You have to get it at that real concrete level that people care about otherwise, you can talk about abstract all day and people say yes. Yes, and they don't do anything about it.

Mick: 15.43
Okay,

Tom: 15.46
People understanding of identity is application and content specific.

Mick: 15.51
Again, please give me your honest answers to this. What are your views on the complexity of the interactions?

All: 15.55
Laughs

D8 Wireframe Link

The following link can be used to download a PDF version of the SSI Wireframe

https://www.dropbox.com/s/f55j2jqk3b3r460/LOCKWOOD_SSI_WIREFRAME_2020_b.pdf?dl=0

https://tinyurl.com/wh9cu28

Appendix E – Ethical Considerations

E1 Participant Invitation Letter

PARTICIPANT INVITATION LETTER
version 1 (01-06-16)
{name}
{address}

Dear {name}

Nodeism – The decentralisation of private data – adoption, usability and the required digital interface. 'An investigation into how a suitable interface and user journey should be designed to allow the participation and mass adoption of a decentralized Internet'.
My name is [*****] and I am currently a PhD candidate at the School of Computer Science at the University of Salford. My research is concerned with the required user interface to allow the adoption of what is know as the Decentralised internet, I am contacting you to invite you to be part of that research.

This phase of my study aims to explore some of the current emerging software in the field to better understand how new users relate to it and interact with it. Participants will be asked to attend a testing session where software will be presented to them. Participants will be asked to interact with the software to achieve certain defined goals. A facilitator will be present and may ask questions or guild the participant through the process. Following the process additional questions maybe asked about the experience. No pervious experience is necessary.

A recording of the testing session will be made consisting of visuals and video of the screen which the participant will be engaging with. The whole process will last no longer then 30mins.

If you decide to take part in this study, a suitable time and date will be agreed and you will be invited into our facility accordingly.

I have included a detailed information sheet about the study with this letter.

It is expected that this project will make a substantial contribution to the fields of software usability, the Decentralised Internet, and the semantic web. The study aims to help inform the development of systems which may eventually lead to better individual security of personal information and data.

I hope that you will choose to become part of this study, and I look forward to hearing from you soon.

Thank you. Yours sincerely, [*****]

E2 Participant Information Sheet

PARTICIPANT INFORMATION SHEET
Version 1 (01-06-16)

Title of Research Project
Nodeism – The Decentralisation of private data – adoption, usability and the required digital interface.

Invitation paragraph
You are being invited to participate in this research project. Before you decide to take part, it is important for you to understand why this research is being done and what it will involve. Please take your time in reading this information sheet carefully and discuss it with others if you wish. Please feel free to ask for further information if there is anything that is unclear or if you would like to have more information. Take time to decide whether
or not you wish to take part.

What is the purpose of the project?
The objective of this section of the research is two-fold, firstly to gather your views and options of the Decentralised Internet, its direction, issues, advances and arguments. And secondly to request your involvement in a later participatory design exercise where prototype interface and user journey development will be defined for emerging sovereign technology.

Why have I been chosen?
You have been chosen on both counts because of your understanding and expertise of the decentralised and related fields. It is hoped that this understanding can act to bring depth and quality to the research outputs of this project.

Do I have to take part?
It is completely up to you whether or not to take part in this project. If you decide to take part you will be asked to sign a consent form. You can withdraw at any time without it affecting any benefits that you are entitled to in any way. You do not have to give a reason for your withdrawal.

What will I be expected to do if I choose to take part in the research?
In the first part of this research phase you will be invited to take part in at interview, the interview will be conducted in person, via telephone or via conference call. You will be asked a number of semi-structured questions regarding the Decentrilised Internet field, the conversation will be recorded and later transcribed, coded, indexed and analysed. In the second part, at a later date, you will be invited to take part in a participatory design process, this process will consist of you being asked to test and comment on the design of a sovereign technology system at various stages.

What are the possible disadvantages and risks of taking part?
In participating in this project in terms of the interview there a possible risk as your options become public, as these may form part of a published paper or thesis, if you wish your identify to remain private your comments will be anonymised.

What are the possible benefits of taking part?
Participation will help to develop research in your field and help in the development of sovereign technology.

Will my taking part in this project be kept confidential? / What will happen to the results of the research project?
This research does not involve the collection of personally sensitive data. In order to contextualize your feedback, we would like your permission to collect the following information:

Your name.
Your research / industry profile.
Your feedback and comments regarding the interview questions.
Your feedback and views on any tested prototype technology.

However, if you prefer to be an anonymous participant, we will keep all your personal information confidential. If this is the case you will appear in all the research outputs with a pseudonym, and we will take all possible measures to keep your data secure and protected.

Once the survey is finished, the data collected in your interview and or feedback on platform develop will be transcribed, coded, indexed and analysed and integrated in the research. All the data collected during the interview, either raw or processed, will be used in academic contexts, such as publications, conferences and other mediums of academic dissemination.

What type of information will be sought from me, and why is the collection of this information relevant for achieving the research project's objective?
Given your background and understanding of the subject area, your option, view and knowledge is very important to this research, it will allow us to better understand the trajectory of the subject area, the academic arguments and the technology developments. Your views will form the foundation of the research direction; the area of depth literature review, and the latter part will inform the development of the practice-based element of this project.

Who is organising / funding the research?
This research is a PhD projected associated with the University of Salford, as such it is bound by their ethics. This research is the solo endeavour of an individual at PhD level, and as such receives no funding.

Contact for further information
Mick Lockwood
School of Art and Media
MC326, Salford University
Media City UK
Salford Quays
M50 2HE
0161 295 4879
07812173855
m.g.lockwood@salford.ac.uk

E3 Interview Consent Example

CONSENT FORM
Interview / Participatory Design Process
Version 1 (1.6.2016) Phase 4 / 5

Participant Identification Number:

Nodeism – The decentralisation of private data – adoption, usability and the required digital interface. 'An investigation into how a suitable interface and user journey should be designed to allow the participation and mass adoption of a decentralized Internet'.

Name of researcher: Mick Lockwood

- I confirm that I have read and understand the information sheet dated 01-06-16 for the above study. I have had the opportunity to consider the information and ask questions.
- I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.
- I understand that the information given by me may be used in future reports, articles or presentations by the research team.
- I understand that my name may appear in any reports, articles or presentations, unless I request that my personal information be anonymised.
- I understand that all data will be stored securely and is covered by the data protection act.
- I agree that the researcher can contact me on the email address or telephone number below to arrange a suitable time to undertake an interview and to take part in a participatory design process.

Please anonymous my personal details.

☐

Name of Participant:

Date:

Signature:

Contact telephone number:

Email Address:

Appendix F – Dissemination of Research

F1 Abstracts Submitted to Frontiers

F1.1 SSI Friction, Value Proposition and Sustainable Adoption

Mick Lockwood [0000-0001-7795-1144]

m.g.lockwood@salford.ac.uk
University of Salford, M5 4WT, UK.

Abstract. This research presents work undertaken to investigate the balancing of interactive friction, against the value proposition required for sustainable adoption of Self-Sovereign Identity (SSI) technology. The work explores the potential friction posed for a full agency engagement with a User-Centred Data Ecosystem (UCDE) through SSI, utilising what is described as a Sovereign Boundary Mechanism (SBM). An SBM is a standardised collection of SSI interactions, which can collectively be described as a metaphorical ring of sovereignty between the participant and the wider network. Within this model participants control identity, relationships, data streams, and access control. This research argues that an interaction with an SBM presents a significant friction, and that the required Value Proposition to attract participant adoption would need to be compelling. This paper considers theory relating to Privacy, Surveillance Capitalism, and Adoption, alongside Thematic Analysis of interviews with experts in the decentralised field, and results from a Public Survey. This research argues that the required value for decentralised technology is unlikely to come from the direct perceived protection of privacy. And that the decentralised Internet cannot be marketed solely on the fact that it is decentralised. Instead value will emerge from the decentralised Internets capability to supersede the centralised model, offering innovation and reduced transactional friction across the individual, business and wider society. It argues that a collective design and development strategy is required within the decentralised community, that a cohesive communications strategy and narrative needs to be developed, and that value needs to be considered across cultural context. This research concludes that an SSI driven UCDE, requires a collective and targeted approach to identify areas of accessible and high commercial value, in order to seed and develop a decentralised ecosystem, to increase the probability of critical mass.

Keywords: Self-Sovereign-Identity, User-Centred Data Ecosystem, Adoption, Value Proposition

F1.2 Friction, Complexity and Cognitive Theory within SSI Interactions

Mick Lockwood [0000-0001-7795-1144]

m.g.lockwood@salford.ac.uk
University of Salford, M5 4WT, UK.

Abstract. The mechanisms and evolving standards collectively known as Self-Sovereign Identity (SSI) offer the prospect of a decentralised Internet by providing a central pillar for a User-Centred Data Ecosystem (UCDE). Once established these technologies promise to afford participants the same agency in the digital realm as individuals experience in the real world. Investigation suggests that the domain is now sufficiently mature to practically realise the principles of SSI, but in order to achieve sustainable adoption significant design focused work needs to be undertaken at the interface layer. Within this paper we present recent practice led research designed to project current SSI prototypes to scale, through conceptual modelling, preliminary user interface, and critical analysis. This research introduces the term Sovereign Boundary Mechanism (SBM), a standardised collection of SSI interactions, which can be described as a metaphorical ring of sovereignty between the participant and the wider network. Within this model participants control identity, relationships, data streams, and access control. This research identifies the domains of interaction and the minimum required objects for a full-scale SSI engagement through an SBM. It defines the components parts and functionality of a wider UCDE which require further consideration, and it identifies original concepts for which a participant may lack metaphor, mental model and understanding. The research considers Human Computer Interaction (HCI) theory across Internalised, External and Distributed cognition, arguing that the current trajectory of SSI requires significant internalised representations, prior knowledge and participant responsibility. This research argues that these elements are problematic and pose a significant friction and barrier to sustainable adoption. In conclusion this research argues that the decentralised community need to recognise the obstacle potentially posed at the interface layer, and engage in collective standardisation, strategy and design thinking to increase the probability of sustainable SSI adoption.

Keywords: Self-Sovereign-Identity, User-Centred Data Ecosystem, Human Data Interaction, Decentralised Internet, HCI Cognitive Theory

F2 MyData-2018 Un-Conference Session

30th Aug 2018 - 1HR

Facilitators:
Mick Lockwood - University of Salford UK
Ian Forester - BBC R&D

Title:
DESIGN Adoption, Usability, Mental Models, Interface, Bridging the Gap.
Participents 16+

A session was proposed as part of the MyData 2018 un-conference. The broad title being Design, Adoption, Usability, Mental Models, Interface, Bridging the Gap. The session was described as an opportunity to discuss the importance of design within the MyData community, the strands some believed where missing, and the importance of engaging design thinking in the context of adoption of the MyData core principles. The session was well attended with some 15 participants and some great debate and discussion ensued.

The issue of adoption ran through the session with an acknowledgement that the redesign of existing web-based services under the guise of decentralization is unlikely to drive adoption. A quote by Philip Sheldrake 'The Decentralised internet cannot be marketed' was raised and discussed. The Privacy Paradox was debated, Instant Gratification Bias and the opinion that a decentralized Internet needs to offer something, better, faster and unique if it is to find mainstream adoption. It was suggested that peak Facebook had now pasted and that following Cambridge Analytica a real consciousness of privacy issues now existed in the general public, and the appetite is now there for a change of direction. Debate continued with regard whether this was enough of a catalyst for change. Regardless of the current level of demand, it was acknowledged that the technology has now matured to facilitate true decentralization and that its practical realization requires concurred design thinking in order to progress.

Conversation continued and the examples found at project VRM was raised, it was acknowledged that we are now in a second cycle of innovation, where many of the concepts and ideas found within the Project VRN archive where now potentially technically viable, and that the archive should be combed for the brightest and best ideas which might now be revitalized. It was suggested that MyData should establish its own archive of decentralised innovations and that the VRM database might well act as a solid historic foundation.

The conversation moved to privacy and the general lack of understanding as a concept. A quote by David Irvine 'MaidSafe' was debated 'What we need to start talking about is control, it's not about your privacy being taken off you, it's about being controlled.' It was suggested that we should move to describe the issue as one of control instead of privacy, and that every effort should be made to replace the notion of privacy with the word control at every opportunity. It was debated that the emotional impact and understanding of the notion of being controlled, impacted more

profoundly and clearly on the individual, and may well go some way to raising awareness of MyData concepts and de-energize the phenomenon of instant gratification bias.

The general conversation about language moved to a discussion regarding design themes, investigating potential value propositions by identifying areas of focus, whether that be developing communication strategy's and educational materials or homing in on broader issues such as, equality, personal data wealth, seamless transaction and so on. It was felt that Identifying themes might be beneficial within the community to focus minds, drive efficiency's and ultimately generate practical solutions, benefit and application.

Some debate around Self Sovereign Identity continued and the concept of a standard Design Patterns where discussed. The user journey through Decentralised identifiers, Verifiable Claims, Zero Knowledge Proofs, Conditions and Ultimately Smart Contracts. Project IF [1] was raised as an example of people already working on Decentralized Data Tool design patterns, and this was followed by some interesting debate around the notion of a contract and the focus on the consequence of infringement in normal negotiations.

The session rounded off with a discussion able what recommendations might be made to the wider community from those with a focus on Design issues. These are listed below.

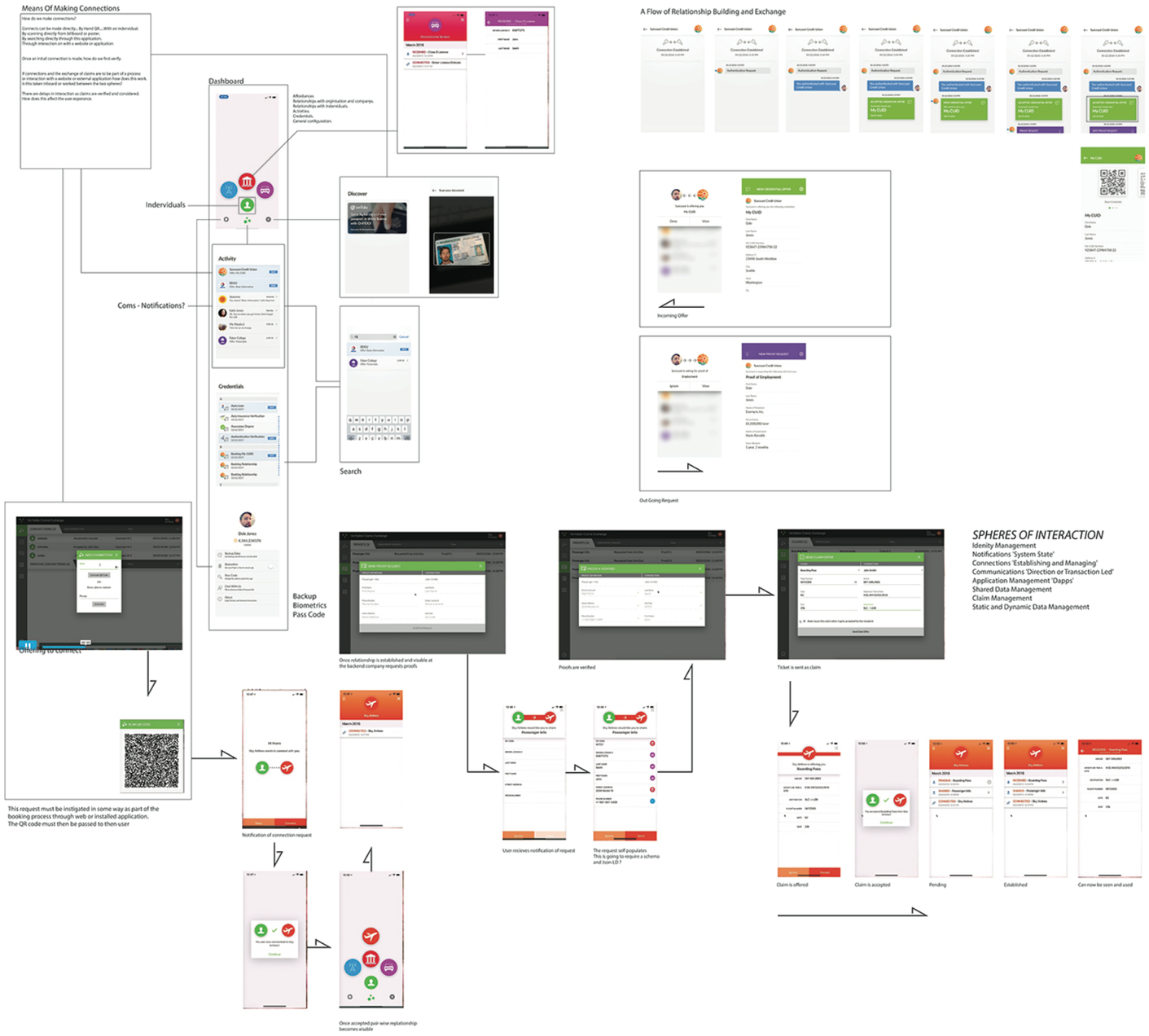
Recommendations from session.

- Establish a substantial design track and recognise its value and necessity within the MyData community.
- Design for Adoption and uncover the real value propositions.
- The need to consider the complexity of interactions and simplification for the user.
- Define a Design Agenda 'what do we want to achieve and where are the priorities.'
- Establish a design based achieve and resource 'a joint effort with Project VRM might super charge this.'

1: https://catalogue.projectsbyif.com/

Appendix G – Evernym Prototype

G1 Dissection & analysis of Evernyms *Connect Me* SSI wallet prototype



References

Aboukhadijeh, F. (2016). *What Is the Decentralized Web?* iSchool at Syracuse University. Available at: https://ischoolonline.syr.edu/blog/what-is-the-decentralized-web/ (Accessed: 1 October 2016).

Acquisti, A, John, L. & Loewenstein, G. (2013). *What is privacy worth?* The Journal of Legal Studies, (pp. 1–57).

Acquisti, A, Taylor, C, & Wagman, L. (2016). *The Economics of Privacy*. Journal of Economic Literature, p.54(2), (pp. 442–492). Available at: http://pubs.aeaweb.org/doi/10.1257/jel.54.2.442. (Accessed: 28 Sept 2017).

AdBlock. (2009). AdBlock [software]. Available at: https://getadblock.com/ (Accessed: 27 Oct 2019).

Albrechtslund, A. (2008). *Online Social Networking as Participatory Surveillance*. First Monday, p.13(3).

Alexa. (2019). *Amazon Alexa Official Site: What is Alexa?* [online] Available at: https://developer.amazon.com/en-US/alexa (Accessed: 19 Nov 2019).

Allen, C. (2016). *The Path to Self-Sovereign Identity.* Available at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html. (Accessed: 1 Jan 2020).

Allen, R. (1997). *Mental models and user models*. Handbook of Human-Computer Interaction, (pp. 49–63), Elsevier.

Anthes, G. (2015). *Data Brokers Are Watching You*. Communications of the ACM.

ARC OS. (2017). *ARC OS* [software]. Available at: https://arkos.io/. (Accessed: 4 July 2017).

Archer, B. (1995). *The Nature of Research*. Co-design, interdisciplinary journal of design, (pp. 6–13).

Barlow, J.P. (1996). *A Declaration of the Independence of Cyberspace*. Available at: https://w2.eff.org/Censorship/Internet_censorship_bills/barlow_0296.declaration. (Accessed: 28 September 2017).

Barnard, P. (1991). *Bridging Between Basic Theories and Artefacts of HCI, Designing Interaction:* Psychology at the Human Computer Interface. Cambridge University Press.

Barnes, S, B. (2006). *A Privacy Paradox*: *social networking in the United States*. First Monday, p.11(9).

BCG. (2012). *The Value of Our Digital Identity*, The Boston Consulting Group. Available at: https://2zn23x1nwzzj494slw48aylw-wpengine.netdna-ssl.com/wp-content/uploads/2017/06/The-Value-of-Our-Digital-Identity.pdf. (Accessed: 22 March 2017).

Beresford, AR, Kubler, D, Preibusch, S. (2012). *Unwillingness to pay for privacy: a field experiment.* Econ Lett, (pp. 25–7).

Berners-Lee, T. (2014). *A Magna Carta for The Web.* Available at: http://www.ted.com/talks/tim_berners_lee_a_magna_carta_for_the_web?language=en. (Accessed: 10 March 2015).

Blandford, Ann. (2013). *Semi-structured qualitative studies.* In: Soegaard, Mads and Dam, Rikke Friis (eds.). *"The Encyclopedia of Human-Computer Interaction, (2nd Ed)."*. Aarhus, Denmark: The Interaction Design Foundation.

Blockstack.org. (2019). [online] Available at: https://blockstack.org/ (Accessed: 13 Oct 2019).

Bogner, B. (2009). *Interviewing Experts.* ECPR Research Methods, Springer.

Brave Browser. (2018). *Secure, Fast & Private Web Browser with Adblocker | Brave Browser* [software]. Available at: http://brave.com/ (Accessed: 27 Oct 2019).

Brown, B. (2001). *Studying the Internet experience.* HP Laboratories Technical Report (HPL-2001-49). Available at: http://www.hpl.hp.com/techreports/2001/HPL-2001-49.pdf (Accessed: 28 September 2017).

Brown, I. (2013). *The Economics of Privacy, Data Protection and Surveillance.* Handbook on the Economics of the Internet.

Bryant, A. (2002). *Regrounding Grounded Theory*. Journal of Information Technology Theory and Application, (pp. 25–42).

Burns, A. (2006). *Towards Produsage, Futures for User-Led Content Production.* Proceedings Cultural Attitudes towards Communication and Technology, (pp. 275–284).

Cafaro, F. (2010). *RFID Localization for Tangible and Embodied Multi-User Interaction with Museum Exhibits*. Proceedings of the 12th ACM International.

Cambridge Analytica. (2017). *Cambridge Analytica* [online] Available at: https://cambridgeanalytica.org/ (Accessed: 30 September 2017).

Cameron, K. (2005). *The Laws of Identity.* Available at: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf. (Accessed: 1 Jan 2020).

Campbell, J.E. et al. (2010). *Online Surveillance and the Commodification of Privacy*. Available at: Panopticon.com. (Accessed: 9 Jan 2015).

Card, S.K., Moran, T.P. and Newell, A. (1983). *The Psychology of Human - Computer Interaction.* Hills-dale, LEA, Hillsdale, NJ.

Carrascal, J. (2013). *Your Browsing Behaviour for a Big Mac: Economics of Personal Information Online.* Proceedings of the 22nd International Conference on World Wide Web. WWW '13. Rio de Janeiro, Brazil, (pp. 189–200).

Carroll, J.M. (1991). *Designing Interaction: Psychology at the Human Computer Interface*. Cambridge University Press.

Catapult. (2016). *Trust in personal data: A UK Review.* Available at: https://www.digicatapult.org.uk. (Accessed: 9 Jan 2015).

Cathy, O. (2016). *Weapons of Math Destruction.* Crown Publishing Group.

CES. (2010). *Consumer Electronics Show.* [Exhibition]. Las Vegas Convention Center. 7 Jan 2010-10 Jan 2010.

Cavoukian, A. (1995). *Privacy-Enhancing Technologies: The Path to Anonymity.* Information and Privacy Commissioner/Ontario Canada. .[online] Available at: http://www.ontla.on.ca/library/repository/mon/10000/184530.pdf (Accessed: 5 Jan 2020).

Choe, E. K. (2014). *Understanding quantified-selfers' practices in collecting and exploring personal data.* In Proceedings of the 32nd annual ACM conference on human factors in computing systems (pp. 1143-1152).

Citizen Me. (2017). Citizen Me Available at: http://www.citizenme.com. (Accessed: 30 September 2017).

Civic. (2020). *Civic Secure Identity Ecosystem - Decentralized Identity & Reusable KYC*. [online] Available at: https://www.civic.com/ (Accessed: 17 Jan 2020).

Cohen, J. (2012). *What is Privacy?* Harvard Law Review, pp.1–24. Available at: http://www.businessdictionary.com/definition/privacy.html. (Accessed: 28 Sept 2017).

Coles-kemp, L. & Lai, Y. (2010). *Privacy on the Internet: Attitudes and Behaviours.*

Craik, K. J. W. (1943). *The Nature of Explanation*. University Press, Macmillan.

Creswell. J.W. (2003*). Research Design: Qualitative, Quantitative and Mixed Methods* Approaches. London: Sage. (2nd ed).

Crowdspending. (2019). *Crowdspending | Join the Crowd, Change the World.* [online] Available at: http://www.crowdspending.com/ (Accessed: 19 Nov 2019).

Cyber Ghost. (2018). *Cyber Ghost Vpn* [software]. Available at: https://www.cyberghostvpn.com (Accessed: 27 Oct 2019).

CYSMN. (2003). *Can You See Me Know*? Available at: https://www.blasttheory.co.uk/projects/can-you-see-me-now/. (Accessed: September 23rd 2019).

Databox Project. (2019). *Databox Project – EPSRC Project on Privacy-Aware Personal Data Platform.* Available at: https://www.databoxproject.uk/ (Accessed: 27 Oct 2019).

Datacoup. (2017). *Datacoup.* Available at: https://beta.datacoup.com. (Accessed: September 30 2017).

DID-Auth. (2020). *WebOfTrustInfo/rwot6-santabarbara*. [online] Available at: https://github.com/WebOfTrustInfo/rwot6-santabarbara/blob/master/final-documents/did-auth.md (Accessed: 17 Jan 2020).

DID. (2020). *Decentralized Identifiers (DIDs) v1.0*. [online] Available at: https://www.w3.org/TR/did-core/ (Accessed: 17 Jan 2020).

Digital-Me. (2019). *digital-me*. [software] Available at: https://digital-me.nl (Accessed: 19 Nov 2019).

DKMS. (2020). *hyperledger/indy-sdk*. [software] Available at: http://bit.ly/dkmsv3 (Accessed: 17 Jan 2020).

Dominode. (2020). *Dominode - Trusted Professional Identity Solutions for a Global Economy*. [software] Available at: http://dominode.com/ (Accessed: 17 Jan 2020).

D'Souza, G, Phelps, J. (2009). *The Privacy Paradox: The Case of Secondary Disclosure.* Review of Marketing Science, De Gruyter, vol. 7(1), pages 1-31, December.

Duhigg, C. (2012). *How Companies Learn Your Secrets.* The New York Times. Available at: http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html. (Accessed: 28 September 2017).

Eben, M. (2010). *Freedom in the Cloud*. Available at: https://www.softwarefreedom.org/events/2010/isoc-ny/FreedomInTheCloud-transcript.html. (Accessed: 28 September 2017).

Elmqvist, N. (2011). Embodied Human-Data Interaction. In Proceedings of the CHI Workshop on Embodied Interaction: Theory and Practice in HCI. (pp. 104–107).

Estrin, D. (2014). *Small data, where n= me.* Communications of the ACM, (4), (pp. 32–34).

EU. (1950). *Convention for the Protection of Human Rights and Fundamental Freedoms*, Available at: http://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680063765. (Accessed: 28 September 2017).

Evans, J.D. (1996). *Straightforward Statistics for the Behavioral Sciences*. Brooks/Cole Publishing; Pacific Grove, Calif.

Evernym. (2019). *Evernym | The Self-Sovereign Identity Company.* [online] Available at: https://www.evernym.com/ (Accessed: 13 Oct 2019).

Express VPN. (2016). *High-Speed, Secure & Anonymous VPN Service | ExpressVPN* [software]. Available at: https://www.expressvpn.com/ (Accessed: 27 Oct 2019).

Finch, H, Lewis, J. (2003). *Focus Groups, Qualitative Research Practice,* SAGE Publications Ltd.

Fitts, P. (1954). *The information capacity of the human motor system in controlling the amplitude of movement.* Journal of Experimental Psychology, Vol. 47, No. 6.

Foucault, M. (1975). *Discipline and Punish: The Birth of the Prison.* Penguin Social Sciences.

Fowler, J. (1993). *Survey Research Methods*. Sage Publications.

Freedombox. (2010). *FreedomBox* Foundation. [software]. Available at: http://freedomboxfoundation.org/ (Accessed: 27 Oct 2019).

Friedman, B., Kahn, P.H. and Borning, A., 2008. Value sensitive design and information systems. *The handbook of information and computer ethics*, (pp. 69-101).

Fuchs, C. (2012). *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*. Routledge, New York.

Future Everything. (2019). *Future Everything*. Available at: https://futureeverything.org (Accessed: 23 September 2019).

Gandy, O. (1996). *Coming to Terms with the Panoptic Sort*. In: Computers, Surveillance, and Privacy (Lyon D and Zureik E, eds), (pp. 132-155). University of Minnesota Press: Minneapolis.

Gaver, W. (1991). *Technology affordances*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91), Scott P. Robertson, Gary M. Olson, and Judith S. Olson (Eds.). ACM, New York, NY, USA, (pp. 79-84).

Gavison, R. (1984). *Privacy and the limits of law*. Philosophical Dimensions of Privacy. An Anthology, 89(3), (pp.346–402).

Getmagic.com. (2019). *Magic - Personal Assistant on Demand*. [software] Available at: https://getmagic.com/ (Accessed: 19 Nov 2019).

Ghostery. (2018). *Ghostery* Makes the Web Cleaner, Faster and Safer! [software]. Available at: http://ghostery.com/ (Accessed: 27 Oct 2019).

Giddens, A. (1986). *The Constitution of Society: outline of the theory of structuration*. Cambridge, Polity.

Glaser, B.G. & Strauss, A.L. (1967). *The Discovery of Grounded Theory*. Available at: http://www.amazon.com/dp/0202302601. (Accessed: 30 September 2017).

Grassegger , H, Krogerus, M. (2016). *The Data That Turned the World Upside Down.* Motherboard. Available at: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win. (Accessed: 30 Sept 2017).

Grassegger, H. (2017). *The Data That Turned the World Upside Down.* Das Magazin. Available at: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win, (Accessed: 28 September 2017).

Greentoe.com. (2019). *Greentoe.com - Name your price on Cameras, TVs, Optics, Electronics and more!*. [online] Available at: https://www.greentoe.com/ (Accessed: 19 Nov 2019).

Group, B, C. (2012). *The Value of Our Digital Identity*. Liberty Global.

Haddadi, H. (2015). *Personal Data: Thinking Inside the Box*. eprint arXiv:1501.04737.

Hall, W. (2016). *Trust in Personal Data*. A UK Review, Catapult.

Harrison, S., Tatar, D. & Sengers, P. (2007). *The Three Paradigms of HCI*. CHI 2007. Grudin, J. *Three Faces of Human-Computer Interaction.* "IEEE Annals of the History of Computing. Vol. 27, no. 4, Oct.-Dec. 2005. (pp 46 – 62).

Harvard. (2017). *Project VRM*. Available at: https://cyber.harvard.edu/projectvrm/Main_Page. (Accessed: 30 September 2017).

Hi-project.org. (2019). *The Hi:Project | EMPOWERING YOU, EMPOWERING US ALL, WITH A MORE HUMAN WEB*. [online] Available at: https://hi-project.org/ (Accessed: 19 Nov 2019).

Himmelfarb, G., (1968). *The Haunted House of Jeremy Bentham*. In: Victorian Minds, (pp 32-81). Alfred A. Knopf, New York.

Homeadvisor.com. (2019). [online] Available at: https://www.homeadvisor.com/ (Accessed: 19 Nov. 2019).

Hornung, H., Pereira, R. & Baranauskas, M.C.C. (2015). *Challenges for Human-Data Interaction* – A Semiotic Perspective, 1, (pp. 37–48).

Huffpost. (2010). *Google CEO On Privacy*. Huffington Post. Available at: http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-if_n_383105.html. (Accessed: September 22 2017).

Hyperledger. (2019). *Hyperledger -The Linux Foundation* [online]. Available at: hyperledger.org (Accessed: 1 Nov 2019).

IBM Watson. (2019). *IBM Watson*. [online] Available at: https://www.ibm.com/thought-leadership/smart/uk-en/watson/index.html (Accessed: 19 Nov 2019).

IIW. (2019). *Internet Identity Workshop*. [online] Available at: https://internetidentityworkshop.com/ (Accessed: 7 April 2019).

Intently.co. (2019). *Intently.co - Request any service anywhere*. [online] Available at: https://intently.co/ (Accessed: 19 Nov 2019).

Jeff, J. (2011). *Public Parts.* Simon & Schuster.

Jenkins, H. (2006). *Convergence Culture*. New York University Press.

Johnson-Laird, P. N. (2001). *Mental Models and Deductio*n. Trends in Cognitive Sciences, vol. 5, no. 10, (pp. 434–442).

Johnson, B. (2010). *Privacy no longer a Social Norm, says Facebook founder*. The Guardian. Available at: https://www.theguardian.com/technology/2010/jan/11/facebook-privacy. (Accessed: 28 September 2018).

Jones, M. & Alony, I. (2011). *Guiding the use of grounded theory in doctoral studies - An example from the Australian film industry*. International Journal of Doctoral Studies, 6, (pp. 95–114).

Jones, W. (2010). *Keeping Found Things Found*. Morgan Kaufmann Publishers.

Kessler, G. (2013). *James Clapper's "least untruthful" statement to the Senate*. The Washington Post. Available at: https://www.washingtonpost.com/blogs/fact-checker/post/james-clappers-least-untruthful-statement-to-the-senate/2013/06/11/e50677a8-d2d8-11e2-a73e 826d299ff459_blog.html?utm_term=.c78196ad2c96. (Accessed: 28 September 2017).

King, L. (2001). *Information, Society and the Panopticon.* The Western Journal of Graduate Research, 10(1), (pp. 40–50).

Kirkpatrick D (2010) The Facebook Effect: The Inside Story of the Company that is Connecting the World. New York: Simon and Schuster.

Kirish, D. and Maglio, P. (1994). *On Distinguishing Epistemic from Pragmatic Action.* Cognative Science 18: 513-549.

Kirsh, D. (1997). *Interactivity and multimedia interfaces.* Instructional Science 25, (pp. 79-96).

Koffka, K. (1935). *Principles of Gestalt psychology.* Oxford, England: Harcourt, Brace

Kokolakis, S. (2017). *Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon*. Computers and Security, 64, (pp. 122–134). Available at: http://dx.doi.org/10.1016/j.cose.2015.07.002. (Accessed: 10 June 2018).

Kosinski, Michal & Stillwell, David & Graepel, Thore. (2013). *Private traits and attributes are predictable from digital records of human behaviour.* Proceedings of the National Academy of Sciences of the United States of America. 110. 10.1073/pnas.1218772110.

Landau, S. (2013). *Snowden: What's Significant in the NSA Surveillance Revelations*. IEEE Security & Privacy, (pp. 54–63).

Landsberger, H. (1959). *Hawthorne Revisited*. Oxford University Press.

Landuer, T.K. (1991). *Let's Get Real Designing Interaction.* Psychology at the Human Computer Interface. Cambridge University Press.

Legard, R. (2003). *In-Depth Interviews.* Qualitative Research Practice*,* SAGE Publications Ltd.

Lehman, J, & Phelps, S. (2005). West's encyclopedia of American law. Encyclopedia.com: Available at: http://www.encyclopedia.com/law/encyclopedias-almanacs-transcripts-and-maps/adhesion-contract. (Accessed: 28 Sept 2017).

Levine, R. (1999). *Clue Chain Manifesto*. Basic Books.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22 140,* 55.

Lockwood, M & Lin, Y & Martin, R., (2011). *'Modernist Map Launch Event* '. exhibited at: Cube Gallery, Manchester, UK, on 23/06/2011.

Lockwood, M & Martin, R. (2010). *Maxamundo, v1.6.* Maxamundo Limited, Manchester, UK. Available at: www.maxamundo.com. (Accessed: 10 Dec 2014).

Lockwood, M. (2010). *Maxamundo IOS application, v1.01*. Maxamundo Limited, Manchester, UK Available at: https://itunes.apple.com/gb/app/maxamundo/id411687715?mt=8. (Accessed: 5 June 2014).

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). *Thematic Analysis: Striving to Meet the Trustworthiness Criteria*. International Journal of Qualitative Methods.

Lyon, D. (1993). *An Electronic Panopticon*. A Sociological Critique of Surveillance Theory. The Sociological Review 41: (pp. 653-678).

Malone, T. (1963). *How Do People Organize Their Desks?* ACM Transactions on Office Information Systems, 1.

Manes, S. (2000). *Full Disclosure: Private Lives? Not Ours!* Computerworld. Available at: https://www.computerworld.com.au/article/20046/full_disclosure_private_lives_ours/. (Accessed: 28 Sept 2017).

Manes, S. (2000). *Private Lives? Not Ours. PC World*. 18 (6): 312.

Mason, R.O. (1986). *Four Ethical Issues of the Information Age*. Southern New Hampshire University.

Meyer, G. (2012). *The API of Me*. Available at: https://gregmeyer.com/2012/11/24/the-api-of-me/ (Accessed: September 30 2017).

Miller, G.A. (1956). *The magical number seven, plus or minus two: Some limits on our capacity for processing information*. Psychological Review, 63, (pp. 81-97).

Moglen, E. (2010). *Freedom in the Cloud*: Software Freedom, Privacy, and Security for Web 2.0 and Cloud Computing. Software Freedom Law Centre. Available at: https://www.softwarefreedom.org/events/2010/isoc-ny/FreedomInTheCloud-transcript.html (Accessed: 10 June 2017).

Moglen, E. (2013). *The Tangled Web We Have Woven*. Communications of the ACM, 56(2). Available at: http://dl.acm.org/citation.cfm?doid=2408776.2408784 (Accessed: 19 Jan 2014).

Molich, R., and Nielsen, J. (1990). *Improving a human-computer dialogue*. Communications of the ACM 33, 3 (March), (pp. 338-348).

Monboit, G. (2017). *Big data's power is terrifying. That could be good news for democracy*. The Guardian. Available at: https://www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy. (Accessed: 28 Sept 2017).

Mortier, R. Haddadi, H. & Henderson, T. (2013). *Challenges & Opportunities in Human-Data Interaction*. Proceedings of DE2013: Open Digital - The Fourth Annual Digital Economy All Hands Meeting, (pp. 4–6). Available at: http://de2013.org/wp-content/uploads/2013/09/de2013_submission_15.pdf. (Accessed: 28 Sept 2017).

MyData.org. (2019). *Declaration - MyData.org*. [online] Available at: https://mydata.org/declaration/ (Accessed: 13 Oct 2019).

Mydex. (2017). *Mydex CIC* [software]. Available at: http://mydex.org/ (Accessed: 27 Oct 2019).

MyTime. (2019). *Online Booking Solution for Businesses | MyTime*. [online] Available at: https://www.mytime.com/ (Accessed: 19 Nov 2019).

Newell, A, Rosenbloom, P. (1980). *Mechanisms of skill acquisition and the law of practice*.

Herbert, A. (1955). *A Behavioral Model of Rational Choice.* The Quarterly Journal of Economics, Volume 69, Issue 1, Pages 99–118.

Nielsen, J. (1993). *Usability Engineering (Interactive Technologies),* Morgan Kaufmann.

Nielsen, J. (2005). *10 Usability Heuristics for User Interface Design.* Available at: http://www.nngroup.com/articles/ten-usability-heuristics/. (Accessed: January 10 2014).

Nielsen, J. (2005). *Usability Inspection Methods.* John Wiley & Sons, Inc.

Nissenbaum, H. (2004). *Privacy as Contextual Integrity*. Washington Law Review 79(1): (pp. 119–158).

Norberg, P, Horne, D, & Horne, D. (2007). *The Privacy Paradox: Personal Information Disclosure Intentions verses US Behaviour's*. The Journal of Consumer Affairs, 41(1), (pp. 100–126).

Norman, D. (1986). *Cognitive Engineering*. In Draper, S. and Norman, D, User Centred System.

Norman, D. (1986). *Theory of Action*. Design. Lawrence Erlbaum Associates, Hillsdale, NJ.

Norman, D. (1987). *Some observations on mental models*. In Human-computer interaction, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (pp. 241-244).

Norman, D. (1993). *Cognition in the head and in the World*. Cognitive Science 17(1), (pp. 1-6).

Norman, D. (2013). *The Design of Everyday Things*: *Revised and Expanded Edition*. New York, Basic Books. http://ebook.yourcloudlibrary.com/library/BCPL-document_id-edxok89.

Norman, D. & Draper, S. (1986). *User Centred System Design*: New Perspectives on Human-computer Interaction, CRC Press.

Nowell, Lorelli & Norris, Jill & White, Deborah & Moules, Nancy. (2017). *Thematic Analysis: Striving to Meet the Trustworthiness Criteria*. International Journal of Qualitative. 16.

Oauth. (2020). *OAuth 2.0 - OAuth*. [software] Available at: https://oauth.net/2/ (Accessed: 17 Jan. 2020).

Ohara, K. (2013). *Are We Getting Privacy the Wrong Way Round?* IEEE Internet Computing.

Ohara, K. (2016). *The Seven Veils of Privacy*. IEEE Internet Computing, 20(2), (pp. 86–91).

Ohm, P. (2010). *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*. UCLA Law Review, 57(6), (pp. 1701–1777.). Available at: http://www.uclalawreview.org/broken-promises-of-privacy-responding-to-the-surprising-failure-of-anonymization-2/. (Accessed: 28 Sept 2017).

Olson, J, & Olson, G. (1991). *The Growth of Cognitive Modelling Since GOMS*. Hum-Computer. Interact. 5, 2.

OpenBazaar. (2019). *OpenBazaar*. [online] Available at: https://openbazaar.org/ (Accessed: 19 Nov 2019)

OpenID. (2020). *OpenID Foundation website*. [online] Available at: https://openid.net/ (Accessed: 17 Jan 2020).

Openstreetmap. (2019). *Open Street Map*. Available at: https://www.openstreetmap.org (Accessed: 23 Sept 2019).

OSM. (2019). *Open Street Map.* [online] Available at: https://www.openstreetmap.org (Accessed: 1 Jan 2019).

OwnCloud. (2017). *OwnCloud - The Leading OpenSource Cloud Collaboration Platform* [software]. Available at: http://owncloud.org/ (Accessed: 27 Oct 2019).

Payne, S. (2003). *Users Mental Models: The Very Ideas*. In HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science.

Pentland, A. (2012). *Reinventing Society in The Wake of Big Data*. Edge.org. Available at: https://www.edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data. (Accessed: 28 Sept 2017).

Posner, R. (1977). *The Right of Privacy*, Review, G.L.A.W. & Lecture, J.A.S. p393.

Post, R.C. (2001). *Three Concepts of Privacy*. Faculty Scholarship Series. Paper 185. http://digitalcommons.law.yale.edu/fss_papers/185.

Poster, M. (1990). *Foucault and Databases: Participatory Surveillance*. In: The Mode of Information, (pp 69-98). The University of Chicago Press, Chicago.

Pugh, S. (1981). Concept Selection *A* Method *that Works.* Proceedings of International Conference on Engineering Design, Heurista, Zürich, 1981, pp. 497-506

ProjectVRM. (2019). Project VRM [online] Available at: https://cyber.harvard.edu/projectvrm/Main_Page (Accessed: 7 April 2019).

Qiyfoundation.org. (2019). *Qiy Foundation*. [online] Available at: https://www.qiyfoundation.org/ (Accessed: 19 Nov 2019).

Quantifiedself.com. (2017). *Quantifiedself.com*. Available at: http://quantifiedself.com/. (Accessed: 30 Sept 2017).

Rainie, L. (2013). *Anonymity, Privacy, and Security Online*. Pew Research Centre.

Re-Decentralize. (2017). *Re-Decentralize*. Available at: http://redecentralize.org/. (Accessed: 30 September 2017).

Reilly, T.O. (2009). *What Is Web2.0 Software?* Available at: http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html. (Accessed: 10 Dec 2014).

Richards, N. & King, J. (2013). *Three Paradoxes of Big Data*. 66 Stanford Law Review Online 41.

Robbins, M. (2017). *The Myth that British Data Scientists Won the Election for Trump*. Little Atoms. Available at: http://littleatoms.com/news-science/donald-trump-didnt-win-election-through-facebook. (Accessed: 30 Sept 2017).

Robins, K. & Webster, F. (1988). Cybernetic Capitalism: Information, Technology, Everyday Life. In: The Political Economy of Information (Vincent Mosco and Janet Wasko, eds), (pp. 44-75). The University of Wisconsin Press, Madison.

Rogers, E. (2012). *Diffusion of Innovations*. New York, Free Press of Glencoe; 5th edition.

Rogers, E. (1958). *Categorizing the adopters of agricultural practices.* Rural Sociology, 23(4), pp.346-354.

Rogers, Y. (2012). HCI *Theory, Classical, Modern, and Contemporary*. Morgan & Claypool Publishers.

Rosenberg, M, et al. (2018). *How Trump Consultants Exploited the Facebook Data of Millions*. New York Times. Available at: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html. (Accessed: September 0, 2019).

Russell, B. (1999). *Headmap Manifesto*. Available at: http://technoccult.net/wpcontent/uploads/library/headmap.pdf (Accessed: 10 June 2014).

Rust, C. (2005). *The Nature of Research.* Review of practice-led research in art, design & architecture. UK, Arts and Humanities Research Council.

Samuel, W. & Brandeis, L. (1890). *The Right to Privacy*. Harvard Law Review.

Schneider, N. (2019). *Decentralization: an incomplete ambition*, Journal of Cultural Economy, 12:4, 265-285.

Schneier, B. (2015). *Data and Goliath*, W. W. Norton & Company.

Schwartz, P. (2004). *Property, Privacy, and Personal Data,* Harvard Law Review, *117*(7), (pp. 2056-2128). doi:10.2307/4093335

Scopus. (2017). *Scopus.* Available at: https://www.elsevier.com/solutions/scopus.

(Accessed: 30 Sept 2017).

Searls, D. (2012). *The Intention Economy*: *When Customers Take Charge*, Linux Journal.

Seiners, P. (1993). Cartoon Captures Spirit of the Internet, The New York Times.

Sharot, T. (2011). *The Optimism Bias.* Current Biology, 21(23), R941-R945.

Shneidermans, B. (1983). *Direct manipulation: A step beyond programming languages*. IEEE Computer.

ShoCard. (2020). *ShoCard: The Premier Blockchain-Based Mobile Identity Platform*. [online] Available at: https://shocard.com/ (Accessed: 17 Jan 2020).

Silverman, J. (2017). *Privacy under Surveillance Capitalism*, United States, Rand Corporation.

Siri. (2019). *Siri*. [online] Available at: https://www.apple.com/uk/siri/ (Accessed: 19 Nov 2019).

Solid. (2017). *Solid*. [online] Available at: https://solid.mit.edu/ (Accessed: 16 June 2017).

Solove, D. (2008). *I've Got Nothing to Hide, and Other Misunderstandings of Privacy*. Solove Post, (May), (pp. 745–772).

Solove, D. (2008). *Understanding Privacy*, Harvard University Press.

Souza, G. & Phelps, J. (2009). *Review of Marketing Science The Privacy Paradox* : The Case of Secondary Disclosure. De Gruyter, vol. 7(1), (pp. 1-31), December.

Sovrin. (2019). *Sovrin Foundation*. [online] Available at: https://sovrin.org/ (Accessed: 13 Oct 2019).

Spiekermann, S. (2017). Towards a value theory for personal data. Available at: https://doi.org/10.1057/jit.2016.4. (Accessed: 30 Sept 2017).

Steel, E. (2013). *Companies scramble for consumer data*. FT. Available at: http://ig-legacy.ft.com/content/f0b6edc0-d342-11e2-b3ff-00144feab7de. (Accessed: 30 Sept 2017).

Such, L.A. (1987). *Plans and Situated Actions*. Cambridge University Press, Cambridge, UK.

Tashakkori, A. Teddlie, C. and Johnson, B. (2015). Mixed Methods. *International Encyclopedia of the Social & Behavioral Sciences*, (pp. 618-623).

Thomson, J.J. (1975). *The Right to Privacy*. Philosophy and Public Affairs, 4(4), (pp. 295–314). Available at: http://www.jstor.org/stable/2265075 (Accessed September 28, 2017).

Torproject.org. (2017). *The Tor Project | Privacy & Freedom Online*. [software]. Available at: https://www.torproject.org/. (Accessed: 27 Oct 2019).

Tsai JY, Egelman S, Cranor L, Acquisti A. (2011). *The effect of online privacy information on purchasing behaviour: an experimental study*. Inform System Res 2011;22(2): (pp. 254–68).

Turnhout, K. (2014). *Design Patterns for Mixed-Method Research in HCI.* HAN University of Applied Science.

Utz, C, et al. (2019). *(Un)informed Consent: Studying GDPR Consent Notices in the Field.* In 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19).

UK-Gov. (2016). *Investigatory Powers Act 2016*. [online] Available at: http://www.legislation.gov.uk/ukpga/2016/25/contents/enacted (Accessed 18 Jan 2020).

UMA (2019). *User-Managed Access (UMA) - ForgeRock*. [online] Available at: https://www.forgerock.com/privacy/user-managed-access (Accessed: 17 Dec 2019).

United Nations. (1948). *Universal Declaration of Human Rights, General Assembly resolution 217 A*. Available at: http://www.un.org/en/universal-declaration-human-rights/. (Accessed: 28 Sept 2017).

Uport.me. (2019). *uPort - Tools for Decentralized Identity and Trusted Data*. [online] Available at: https://www.uport.me (Accessed: 13 Oct 2019).

US-Gov. (1973). *Records Computers and the Rights of Citizens*. Department of Health, Education and Welfare.

Van Kleek, M. & O'Hara, K. (2014). *The Future of Social is Personal: The Potential of the Personal Data Store*. Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society, (pp. 125–158). Available at: http://eprints.soton.ac.uk/363518/1/pds.pdf. (Accessed: 28 Sept 2017).

VCWG. (2020). *W3C Verifiable Credentials Working Group*. [online] Available at: https://www.w3.org/2017/vc/WG/ (Accessed: 17 Jan 2020).

Veres.one. (2019). *Veres One - A Globally Interoperable Blockchain for Identity*. [online] Veres.one. Available at: https://veres.one/ (Accessed: 13 Oct 2019).

W. E. Hick. (1952). *On the Rate of Gain of Information*. Quarterly Journal of Experimental Psychology, 4:1, (pp. 11-26).

Wakefield, R. (2013). *The Influence of User Affect in Online Information Disclosure*. Journal of Strategic Information Systems. (pp. 157–174).

Walker, K. (2000). *Where Everybody Knows Your Name: A Pragmatic Look at the Costs of Privacy and the Benefits of Information Exchange*. Stanford Technology Law Review, 190.

WEF. (2011). *Personal Data: The Emergence of a New Asset Class*. An Initiative of the World Economic Forum, Available at: http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf. (Accessed: September 1 2018).

WEF. (2011). *Personal Data: The Emergence of a New Asset Class.* An Initiative of the World Economic Forum, Available at: http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf. (Accessed: 13 Oct 2019).

Weston, A. (1967). *Privacy and Freedom*, Ig Publishing.

Whitten, A, & J, Tygar. (1999). *Why Johnny can't encrypt: A usability evaluation of pgp 5.0.* In Proceedings of the 8th USENIX Security Symposium, August 1999.

WOT. (2017). *Web of Trust.* [online] Available at: https://www.weboftrust.info/ (Accessed: 7 April 2019).

Young, K. (2018). *Domains of Identity*.[online] Available at: https://identitywoman.net/domains-of-identity (Accessed: 5 Jan 2020).

Zuboff, S. (2015). *Big Other: Surveillance Capitalism and the Prospects of an Information Civilization.* Journal of Information Technology 30: (pp. 75–89).